

ABSTRACT

Title of Document: PROTEOMIC ANALYSIS OF HUMAN
URINARY EXOSOMES.

Patricia Amalia Gonzales Mancilla, Ph.D., 2009

Directed By: Associate Professor Nam Sun Wang, Department
of Chemical and Biomolecular Engineering

Exosomes originate as the internal vesicles of multivesicular bodies (MVBs) in cells. These small vesicles (40-100 nm) have been shown to be secreted by most cell types throughout the body. In the kidney, urinary exosomes are released to the urine by fusion of the outer membrane of the MVBs with the apical plasma membrane of renal tubular epithelia. Exosomes contain apical membrane and cytosolic proteins and can be isolated using differential centrifugation. The analysis of urinary exosomes provides a non-invasive means of acquiring information about the physiological or pathophysiological state of renal cells. The overall objective of this research was to develop methods and knowledge infrastructure for urinary proteomics. We proposed to conduct a proteomic analysis of human urinary exosomes. The first objective was to profile the proteome of human urinary exosomes using liquid chromatography-tandem spectrometry (LC-MS/MS) and specialized software for identification of peptide sequences from fragmentation spectra. We unambiguously identified 1132 proteins. In addition, the phosphoproteome of human urinary exosomes was profiled using the neutral loss scanning acquisition mode of LC-MS/MS. The phosphoproteomic profiling identified 19 phosphorylation sites corresponding to 14 phosphoproteins. The second objective was to

analyze urinary exosomes samples isolated from patients with genetic mutations. Polyclonal antibodies were generated to recognize epitopes on the gene products of these genetic mutations, NKCC2 and MRP4. The potential usefulness of urinary exosome analysis was demonstrated using the well-defined renal tubulopathy, Bartter syndrome type I and using the single nucleotide polymorphism in the ABCC4 gene. The third objective was to study the normal variability between proteomes of female and male urinary exosomes, and to implement a normalization method to analyze urinary exosome samples. Only 19 proteins had a 2-fold change representing 4.9% of the total number of proteins identified which shows that there is high concordance between proteomes of urinary exosomes isolated from males and females. The normalization method, timed urine collection did not correlate as expected with the intensity signal of MVB markers, TSG101 and Alix. This research shows that the proteomic analysis of human urinary exosomes can be the basis for future biomarker studies as well as physiological studies.

PROTEOMIC ANALYSIS OF HUMAN URINARY EXOSOMES.

By

Patricia Amalia Gonzales Mancilla

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:

Associate Professor Nam Sun Wang, Chair

Dr. Mark A. Knepper, Special Graduate Faculty Member

Associate Professor Sheryl Ehrman

Professor William E. Bentley

Professor Catherine Fenselau

© Copyright by
Patricia Amalia Gonzales Mancilla
2009

Dedication

I dedicate this work to my husband, my daughter, my parents and my brother.

Acknowledgements

I would like to thank my advisors: Dr. Nam Sun Wang, Associate Professor in the Department of Chemical and Biomolecular Engineering at the University of Maryland-College Park and Dr. Mark A. Knepper, Chief of the Laboratory of Kidney and Electrolyte Metabolism (LKEM) at the National Heart, Lung, Blood Institute (NHLBI) of the National Institutes of Health (NIH) for their continued guidance and support during my graduate career. The research experience that I gained as a graduate student is invaluable. Also, I would like to thank my committee members for their advice and support in the completion of this work. I am very grateful to Dr. Trairak Pisitkun for his endless help and mentoring while conducting my dissertation research. Thanks to Dr. Robert E. Kleta for his continued collaboration to applying urinary exosomes analysis to clinical medicine. Thanks to Dr. Robert A. Star and Dr. Hua Zhou for their valuable collaboration in biomarker discovery pertaining to urinary proteomics. Thanks to Dr. Jeffrey Kopp for his collaboration in the gender proteomics study. To all the members of the Laboratory of Kidney and Electrolyte Metabolism, thank you for all your support. A special thanks to the Clark School's Future Faculty Program for providing additional training that will serve me in the future. To all my friends in college and graduate school, I will treasure your support and friendship forever. To my parents and brother, thank you for your love, trust, understanding, guidance and continuous support in all my life endeavors. To my husband, thank you for your unconditional love, friendship, and support. To my daughter, thank you for inspiring me.

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vi
List of Figures.....	vii
1 Introduction.....	1
2 Background.....	5
2.1 Human urinary exosomes.....	5
2.1.1 Introduction.....	5
2.1.2 Urinary exosome biogenesis.....	7
2.1.3 Summary.....	8
2.2 Proteomics.....	9
2.2.1 Introduction.....	9
2.2.2 Tandem Mass Spectrometry.....	9
2.2.3 Analysis of tandem mass spectrometry data.....	11
2.2.4 Summary.....	12
2.3 Development of urinary clinical proteomics.....	12
3 Large Scale Proteomic Analysis of Human Urinary Exosomes *.....	14
3.1 Introduction.....	14
3.2 Methods.....	15
3.2.1 Large-scale proteomic analysis.....	15
3.2.1.1 Urinary Exosome Isolation.....	15
3.2.1.2 In-Gel Trypsin Digestion.....	17
3.2.1.3 Nanospray LC-MS/MS.....	18
3.2.1.4 Analysis of Data.....	18
3.2.1.5 Target-Decoy.....	18
3.2.1.6 InsPecT.....	19
3.2.1.7 Minimizing False-Discovery Peptide Identifications.....	20
3.2.1.8 Elimination of Ambiguous Protein Identifications.....	20
3.2.2 Phosphoproteomic Analysis.....	22
3.2.2.1 Phosphopeptide Enrichment.....	22
3.2.2.2 Analysis of Phosphopeptide Data Sets.....	23
3.3 Results.....	24
3.3.1 Large-Scale Proteomic Profiling of Human Urinary Exosomes.....	24
3.3.2 Phosphoproteomic Analysis of Human Urinary Exosomes.....	25
3.4 Discussion.....	28
3.4.1 Large-Scale Proteomic Profiling of Human Urinary Exosomes.....	28
3.4.2 Relevance to Renal Biology.....	29
3.4.3 Phosphoproteomic Analysis of Human Urinary Exosomes.....	30
3.5 Conclusion.....	30
4 Application of urinary exosomes analysis using polyclonal antibodies.....	43
4.1 Introduction.....	43
4.1.1 Bartter syndrome type I *.....	43
4.1.2 Mutation of ABCC4 gene.....	44

4.1.3	Designing anti-human polyclonal antibodies.....	45
4.2	Methods and Materials.....	47
4.2.1	Patients with Bartter Syndrome Type I.....	47
4.2.2	Patients with mutation in ABCC4 gene.....	47
4.2.3	Polyclonal antibodies.....	48
4.2.3.1	Multidrug resistant protein 4.....	48
4.2.3.2	Thiazide-sensitive co-transporter.....	49
4.2.3.3	Sodium potassium chloride co-transporter 2.....	50
4.2.4	Affinity purification of polyclonal antibodies.....	50
4.2.5	Characterization of polyclonal antibody.....	51
4.3	Results.....	51
4.3.1	Bartter syndrome type I.....	51
4.3.2	Mutation of ABCC4 gene.....	54
4.4	Discussion.....	59
4.4.1	Bartter syndrome type I.....	59
4.4.2	Mutation in ABCC4 gene.....	59
4.5	Conclusion.....	61
5	Normal variability and normalization of urinary exosomes: Gender Proteomics	63
5.1	Introduction.....	63
5.2	Methods and Materials.....	65
5.2.1	Tandem mass spectrometry study.....	65
5.2.1.1	In-Gel Trypsin Digestion.....	66
5.2.1.2	Nanospray LC-MS/MS.....	67
5.2.2	Analysis of Data.....	67
5.2.3	Label-free quantification.....	68
5.2.3	Immunoblotting study.....	69
5.2.3.1	Urine collection.....	69
5.2.3.2	Urinary exosome isolation.....	70
5.2.3.3	Immunoblotting.....	70
5.3	Results.....	71
5.4	Discussion.....	73
5.5	Conclusion.....	74
6	Summary and Conclusion.....	84
7	Future Works.....	89
7.1	Isolation Method.....	89
7.2	Tamm-Horsfall.....	89
7.2	Quantification.....	89
8	Bibliography.....	91
Appendix A	97
Electron Micrograph Negative Staining Protocol with Uranyl Acetate.....	97	
Appendix B	99
Complete List of Proteins in Human Urinary Exosomes.....	99	

List of Tables

Table 3.1	Disease Related Proteins in Human Urinary Exosomes.....	32
Table 3.2	Solute and water transporters.....	37
Table 3.3	Proteins of the ESCRT Complex.....	39
Table 3.4	Vacuolar H ⁺ -ATPase units in human urinary exosomes.....	40
Table 3.5	Human urinary exosomes phosphopeptides.....	41
Table 4.1	Kyte-Doolittle Hydropathy Index for MRP4 protein.....	62
Table 5.1	Proteins identified in male and female urinary exosomes proteomes.....	76
Table 5.2	Proteins that changed at least 2-fold.....	83
Supplementary Table 1	Complete list of proteins identified in human urinary exosomes.....	99

List of Figures

Figure 1.1	Diagram of the bisected kidney.....	1
Figure 1.2	Diagram of the nephron.....	2
Figure 2.1	Electron micrograph (EM) of human urinary exosomes.....	5
Figure 2.2	Diagram of membrane trafficking process.....	6
Figure 3.1	Differential centrifugation procedure for the isolation of urinary exosomes from urine.....	16
Figure 3.2	Spectrum generated by InsPecT for CHMP1A protein (NP_00279).....	20
Figure 3.3	Criteria to disambiguate data set.....	21
Figure 3.4	Novel phosphorylation site in the NCC protein.....	27
Figure 3.5	Detection of AQP2-S256 phosphorylation in urinary exosomes.....	28
Figure 4.1	Disease-related protein NKCC2 and Bartter syndrome Type 1.....	52
Figure 4.2	Genetic mutation of ABCC4 results in short and normal MRP4 protein..	54
Figure 4.3	Characterization of rabbit polyclonal anti-MRP4 antibody (316) generated with synthetic peptide 1 (LPVYQEVKPNPLQDANLC).....	56
Figure 4.4	Characterization of rabbit polyclonal anti-MRP4 antibody (319) generated with synthetic peptide 2 (KIGHKRRLEEDDMYSVLPEDRS).....	57
Figure 4.5	Genetic mutation of ABCC4 results in production of short and normal MRP4 protein.....	58
Figure 5.1	Immunoblots of urinary exosomes samples from three male healthy volunteers (M2, M3, M4) and three female healthy volunteers (F1, F2, F3).....	72

1 Introduction

The sequencing of the human genome had a great impact in biomedical research because access was gained to large amounts of data to expand the analysis of genes, gene products and their function [1]. Proteins are the products of genes and are composed of a sequence of amino acids. Proteins are the functional components of eukaryotic cells. The majority of human protein sequences are known and this information can be used for proteomic studies, i.e. the “systematic analysis of proteins for their identity, quantity and function” [2]. As such, the development of proteomic technology in renal physiology and pathophysiology has the potential to greatly impact clinical medicine [3]. In order to implement proteomic technology in clinical medicine, it is important to develop: methods infrastructure, cost-effective, and non-invasive techniques for disease biomarker identification and quantification in body fluids such as urine that will allow early disease detection.

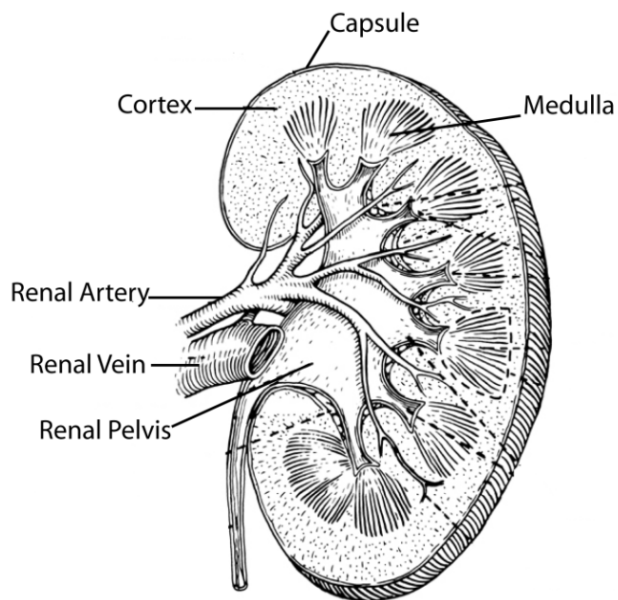


Figure 1.1 Diagram of the bisected kidney [4].

The kidneys are organs located in the posterior part of the abdomen on each side of the vertebral column [4]. These organs maintain the water and sodium balance of the human body [5]. Each kidney surrounded by a fibrous capsule has a renal artery and renal vein [4]. As shown in Figure 1.1, two major regions can be identified in the bisected kidney, the cortex and the medulla. The medulla can further be divided into the outer and the inner medulla. The outer medulla divides into an inner and an outer stripe.

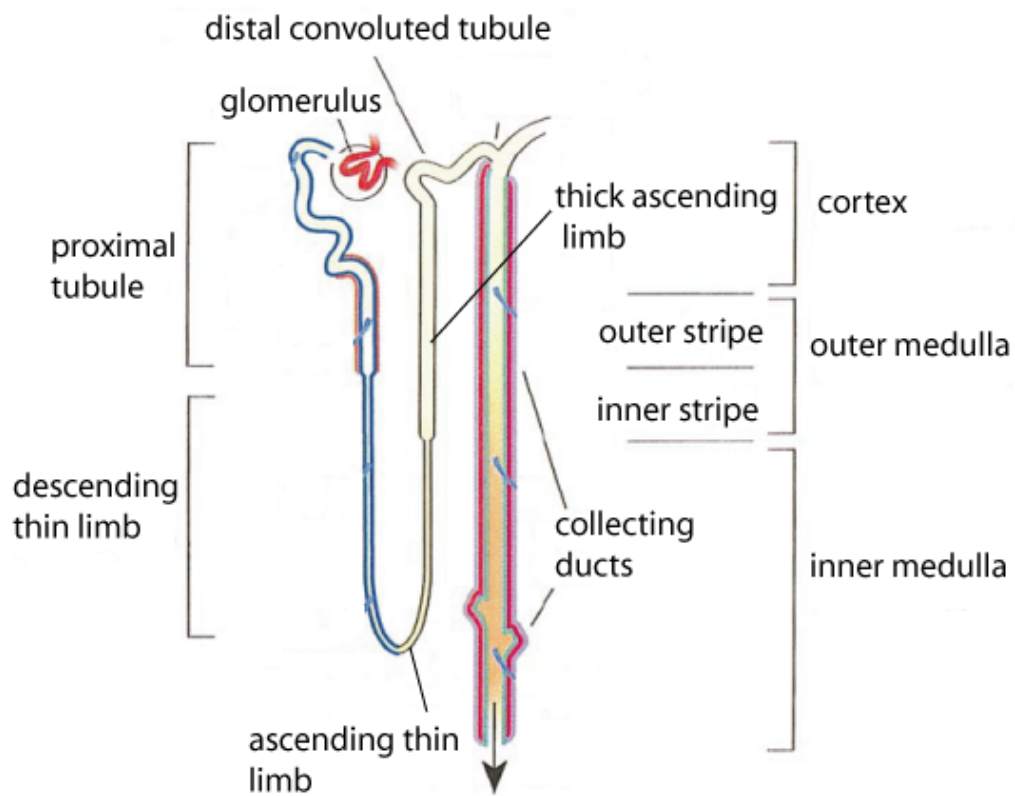


Figure 1.2 Diagram of the nephron [10].

The functional unit of the kidney is the nephron and there are approximately 1 million nephrons in each human kidney. The nephron shown in Figure 1.2 is composed of the glomerulus, proximal tubule, Loop of Henle, the distal convoluted tubule, and the collecting ducts. These tubule segments are found in the three regions of the kidney, the

cortex, the outer medulla and the inner medulla as shown in the figure. The proximal tubule is further divided into the proximal convoluted tubule (PCT) and the proximal straight tubule (PST). The loop of Henle is composed of the thin limb (TL) and thick ascending limb (TAL). There are two regions of the collecting duct, the outer medullary collecting duct (OMCD) and the inner medullary collecting duct (IMCD) [4]. Kidney diseases result from damage to the nephrons and/or by genetics. Some of these renal disease processes are glomerular diseases such as nephritic syndrome and renal tubular disorders such as polycystic kidney disease, Gitelman syndrome, Bartter syndrome, nephrogenic diabetes insipidus, and familial magnesium wasting syndrome [6].

In developing non-invasive approaches to identify disease processes internal to the kidney, it is ideal to use an easily collected biological sample, such as urine. The urine proteome is very complex and therefore it is important to focus on an isolated sub-proteome. This sub-proteome is found in isolated urinary exosomes, which constitutes 3% of the total excreted in the urine [7, 8]. Urinary exosomes arise as internal vesicles of multivesicular bodies (MVBs) that are delivered to the urinary space by fusion of the outer membrane of MVBs with the apical plasma membrane of renal tubule epithelial cells [6]. Exosomes derive from every epithelial cell type facing the urinary space, glomerular podocytes, and cells lining the renal pelvis, ureters and bladder.

The overall objective of this research is to develop methods and knowledge infrastructure for urinary proteomics. The central focus of the proposed work is on urinary exosome analysis by protein mass spectrometry. The analysis of urinary exosomes can be the basis of future disease biomarker discovery studies. Conventional

biomarker studies consist of three phases: discovery, validation and implementation [9].

The specific objectives proposed for this research are:

1. Large-scale proteomic and phosphoproteomic analysis of human urinary exosomes
2. Application of human urinary exosome analysis using polyclonal antibodies
3. Normal variability and normalization of human urinary exosomes: Gender Proteomics

2 Background

2.1 Human urinary exosomes

2.1.1 Introduction

Exosomes originate as the internal vesicles of multivesicular bodies (MVBs) in cells (Figure 2.1). They were first described as products of circulating blood cells [11], such as erythrocytes and lymphocytes, and have been shown to be secreted by most cell types throughout the body. These include B cells, T cells, dendritic cells, reticulocytes, mastocytes, enterocytes and platelets [9]. In the kidney, exosomes are released to the urine by fusion of the outer membrane of the MVBs with the apical plasma membrane (Figure 2.2).

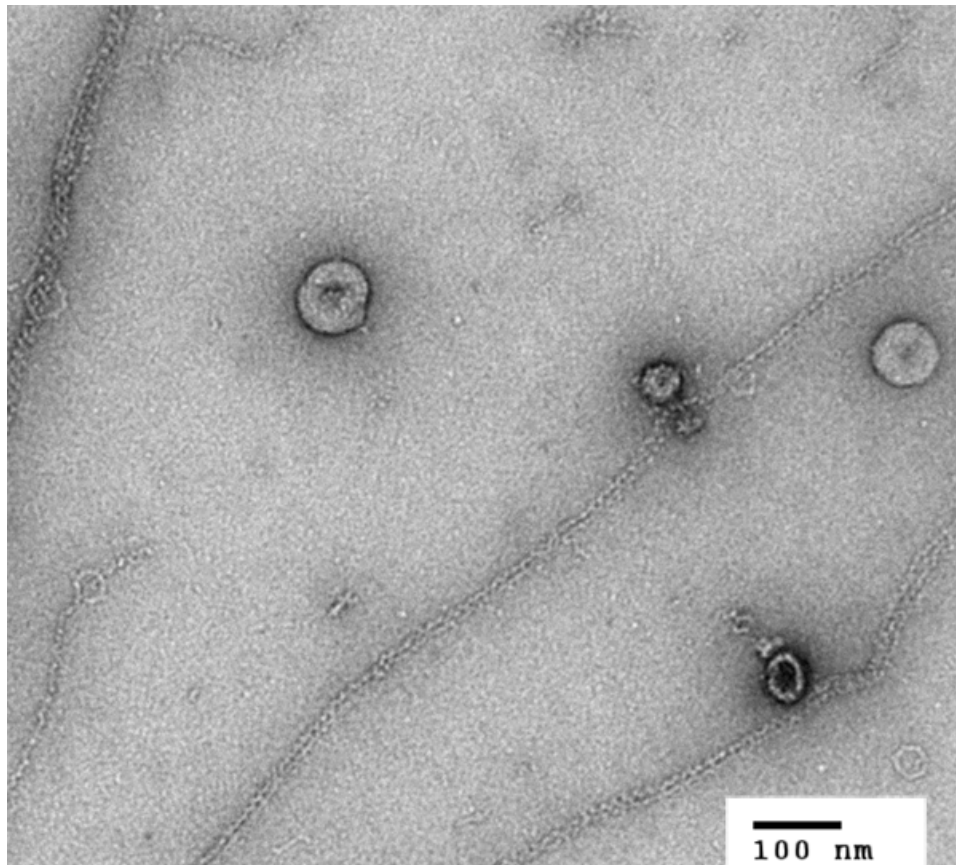


Figure 2.1 Electron micrograph (EM) of human urinary exosomes. Exosomes were isolated from normal human urine and attached to EM grids. The grids were subjected to negative staining (Appendix A).

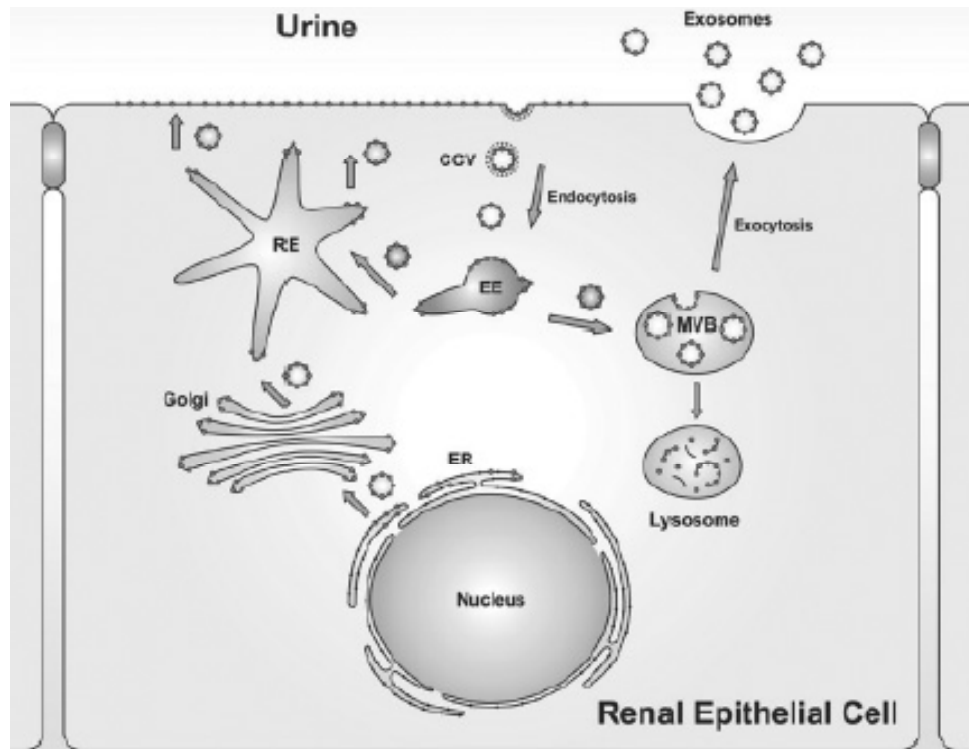


Figure 2.2 Diagram of membrane trafficking process [31]. The mRNA of a transporter (dots) is translated in the rough endoplasmic reticulum (ER) and resulting protein is processed in the Golgi apparatus. Vesicles budding from the cis-Golgi are believed to be delivered to the apical plasma membrane via the endosomal system, by vesicle-mediated translocation to recycling endosomes (RE). The protein in the apical plasma membrane undergoes regulated endocytosis via clathrin-coated vesicles, which deliver it to early endosomes and subsequently to late endosomes (LE) or multivesicular bodies (MBV). The transporter protein accumulates inside MVBs as a component of the internal MVB vesicles. The MVBs can become lysosomes or the membrane of MVBs can fuse with the apical plasma membrane delivering the internal vesicles to the urine.

Multivesicular bodies (MVBs) or late endosomes are part of the endocytic pathway.

The endocytic pathway works through the process known as endocytosis. Endocytosis is a cellular process by which eukaryotic cells internalize fluid, molecules and particles through an internal membrane system [12]. These molecules include proteins, carbohydrates and lipids; they are delivered to the cell exterior via the biosynthetic-secretory pathway which involves the process known as exocytosis.

The endocytic pathway is composed of endocytic vesicles, early endosomes, late endosomes (MVBs) and lysosomes [13]. Molecules are endocytosed into the cell by the invagination of clathrin-coated pits on the plasma membrane. These clathrin-coated pits pinch off into clathrin-coated vesicles inside the cell. The clathrin coat is shed and the endocytic vesicles fuse with early endosomes. The early endosomes mature into multivesicular bodies by accumulating internal vesicles [14, 15]. These internal vesicles contain protein and lipids [13]. The multivesicular bodies have three distinct fates: (1) to target protein and lipids for degradation by fusing with lysosomes, (2) to serve as temporal storage compartments and (3) to release their internal vesicles as exosomes through exocytosis [13].

2.1.2 Urinary exosome biogenesis

Urinary exosomes are delivered to the urine from all renal epithelial cell types. These small extracellular vesicles have a size of 40 to 100nm and contain membrane and cytosolic proteins. The biogenesis of urinary exosomes is coupled to the formation of multivesicular bodies. The biogenesis and sorting of cargo protein of multivesicular body (MVB) depends on the endosomal sorting complexes required for transport (ESCRT) [15, 16, 17]. The mammalian ESCRT machinery is composed of four complexes which are ESCRT-0, ESCRT-I, ESCRT-II, and ESCRT-III [16, 17].

The biogenesis of exosomes starts when plasma membrane proteins are monoubiquitinated to be sorted as cargo by the ESCRT machinery [17]. Ubiquitin is a protein that can be covalently conjugated to cargo proteins by their lysine residues [16, 17]. Initial recognition of monoubiquitinated proteins occurs by ESCRT-0 which is composed by hepatocyte growth factor-regulated Tyr-kinase substrate (HRS) and two

signal transducing adaptor molecules (STAM1 and STAM2) [16, 17]. The cargo is transferred to ESCRT-I, composed of TSG101, VPS28, VPS37A, VPS37B, VPS37C and VPS37D [16]. The transfer occurs through the binding of the proline-serine-alanine-proline (PSAP) motif of HRS by TSG101 [17]. The cargo is transferred to the ESCRT-II complex composed of VPS22, VPS25 and VPS36. The TSG101 protein of the ESCRT-I complex is bound by the VPS36 protein of the ESCRT-II complex. There is evidence that the ESCRT-II complex can be bypassed in mammalian cells if the ESCRT-accessory protein, ALIX, connects ESCRT-I and ESCRT-III [16, 17]. Once the cargo reaches the ESCRT-III complex it is deubiquitinated, sorted and concentrated in maturing endosomes. The ESCRT-III complex is composed of seven chromatin modifying proteins (CHMP 1-7) [17]. The cargo is dissociated from the ESCRT complex by the AAA+ATPase proteins and accessory proteins. The VPS4 protein, an AAA+ATPase, disassembles the clathrin coats of the cargo and initiates the formation of the internal vesicles of MVBs [16]. Ubiquitination of a protein is not necessary for sorting into the formation of MVBs [13] as shown by the presence of tetraspanins (CD9, CD63, CD81 and CD82) in exosomes. Tetraspanins inwardly bud into MVBs because they are localized in lipid rafts [13].

2.1.3 Summary

Urinary exosomes are the internal vesicles of multivesicular bodies (MVBs) and are released into the urine from renal epithelial cells, including glomerular podocytes, renal tubule cells and cells lining the urinary drainage system. Biogenesis of multivesicular bodies (MVBs) is important in the generation of urinary exosomes. Multivesicular bodies or late endosomes are part of the endocytic pathway of cells.

2.2 Proteomics

2.2.1 Introduction

Proteomics is the large-scale study of proteins and it can further characterize protein function, protein interactions and protein sub-cellular distribution [18]. There are two aspects of proteomics to be considered in developing new methods and knowledge infrastructure for urinary exosome proteomics. As MacBeath points out these aspects are ‘discovery-oriented proteomics’ and ‘systems-oriented proteomics’. Discovery-oriented proteomics is the “unbiased approach” to discovering unknown proteins in a biological sample. Systems-oriented proteomics is conducted by analyzing a defined subset of proteins in a biological sample [19].

The discovery-oriented proteomics approach is fundamental to identify potential biomarkers in the proteome of human urinary exosomes. Mass spectrometry (MS) is an analytical technique that measures the masses of molecules and atoms [20]. A mass spectrometer measures the mass-to-charge ratio of charged molecular ions.

2.2.2 Tandem Mass Spectrometry

Mass spectrometry analysis starts by converting molecules into gas-phase molecular ions [19]. A mass analyzer separates the gas-phase molecular ions according to the mass-to-charge ratios (m/z). A detector measures the ion current of mass-separated molecular ions. And a data system records these measurements and displays them as a mass spectrum; the mass spectrum shows the relative abundance of ions versus their m/z values [20].

The development of electrospray ionization (ESI) was important because it opened mass spectrometry to the “identification of proteins and peptides [21].” Electrospray ionization is a technique for the generation of gas-phase ions from electrically charged liquid droplets [22, 23]. This technique is particularly useful with complex molecules because it features the formation of a series of multiply charge ions. In order to analyze complex samples, these should be separated by chromatography before ionization. The separation method best suited for complex samples is reverse-phase high-performance liquid chromatography (RP-HPLC) [24]. Separation with RP-HPLC depends on the hydrophobic interaction between the nonpolar matrix of the stationary phase and the hydrophobic group of the sample [24]. The coupling of the separation method, ionization method and mass spectrometry is known as LC-MS.

The LC-MS approach is coupled to another mass spectrometer in the LC-MS/MS approach for protein identification. The LC-MS/MS approach, is known as tandem mass spectrometry and provides “enhanced sensitivity, selectivity and speed of analysis” of complex mixtures [25]. Tandem mass spectrometry involves mass selection of parent peptide ions (MS1), fragmentation of parent peptides and mass analysis of these peptides (MS2). Fragmentation of peptide ions is accomplished by collision-induced dissociation (CID).

Proteomic analysis is accomplished through the digestion of complex protein mixtures using trypsin, an enzyme that cuts peptide bonds following arginine or lysine. These tryptic peptides are separated using RP-HPLC and then are delivered to the electrospray for ionization. The peptide ions are delivered to the first mass spectrometer (MS1), and m/z information is recorded. Selected parent peptide ions are fragmented by

CID that yields complementary b- and y- ions. The m/z of complementary ions are analyzed in the second mass spectrometer (MS2) and recorded. The m/z data from MS1 and MS2 constitute the experimental spectra and peptides are identified by comparison of experimental spectra to a database of theoretical spectra. Various approaches have been developed to analyze large-scale data generated by tandem mass spectrometry studies such as SEQUEST [26, 27], target-decoy [28] and InsPecT [29].

2.2.3 Analysis of tandem mass spectrometry data

SEQUEST is a program based on a “mechanistic prediction” of peptide fragmentation that is quantified to “determine the quality of the match” between the theoretical and the experimental spectrum [27]. The SEQUEST program accomplishes this in four steps: (1) reduction of data, (2) identification of amino acid sequences in peptide database to match experimental mass spectral data, (3) comparison of predicted fragment ions from the peptide database to the experimental mass spectral data and (4) correlation-based analysis of predicted fragment ions from peptide database and experimental mass spectral data [26].

The target-decoy database search strategy provides an estimate of false-discovery rate (FDR) for large data sets. A composite database is created with the normal (target) and reversed (decoy) peptide sequences of the theoretical spectra. The experimental spectra is compared with the composite database and “by definition, 100% of correctly assigned spectra should come from the target database” and “incorrectly assigned spectra should have an equal chance of coming from either the target or decoy database [28]”. The entire data set is filtered by the SEQUEST parameters and the false-discovery rate (FDR) is calculated.

InsPecT uses *de novo* sequencing to reconstruct peptides from experimental mass spectra and theoretical peptide database search [29]. The program generates sequence information in the form of tag filters from the experimental data. These tag filters are used to search the peptide database and identify peptide sequences that match the experimental data. The identifications are scored by calculating the *P*-value by comparing match quality score to the distribution of quality scores for incorrect matches [29].

2.2.4 Summary

Proteomics is the large-scale study of proteins. A mass spectrometer measures the mass-to-charge ratio (m/z) of molecular ions. Tandem mass spectrometry coupled to liquid chromatography is the technique widely used to identify proteins in complex samples. There are many programs used to analyze large data sets generated by tandem mass spectrometry and these use two well-known approaches: theoretical database comparison and *de novo* sequencing.

2.3 Development of urinary clinical proteomics

Clinical proteomics connects scientific and clinical expertise for the purpose of understanding and treating disease [30]. The proteomic analysis of urinary exosomes has the potential of being useful in clinical proteomics. Clinical proteomics will require the completion of three stages to impact the early diagnosis of disease, the classification of disease with regard to therapeutic responses, in assessment of prognosis and in monitoring response to therapy. These three stages are: (1) discovery of potential biomarkers, (2) validation of potential biomarkers and (3) implementation of biomarker assays [9, 30].

The discovery of potential biomarkers is a hypothesis generation step. It should be done on samples from a relatively small number of extremely well-characterized patients and appropriately chosen control subjects. The tandem mass spectrometry approach can be used to profile the entire proteome of the isolated urinary exosomes from these patients. Quantitative mass spectrometry approaches should be implemented to compare patients with controls and determine combination of potential protein biomarkers. The protein biomarkers discovered in the first stage can be tested in the validation stage to determine their specificity and sensitivity. Validation of protein biomarkers should be done in large clinical data sets with proper adjustment for risk factor distribution [31]. The possible assays used for validation are mass spectrometry, and antibody-based quantification [9]. These assays used for validation should also be used for the final stage, implementation. Implementation of biomarkers assays will be completed by adapting these assays for use in the clinical setting.

3 Large Scale Proteomic Analysis of Human Urinary Exosomes *

3.1 Introduction

Urinary exosomes are small extracellular vesicles (<100 nm in diameter) that originate from the internal vesicles of multivesicular bodies (MVB) in renal epithelial cells, including glomerular podocytes, renal tubule cells, and the cells lining the urinary drainage system [6]. Exosomes are released into the urine when the outer membrane of the MVB fuses with the apical plasma membrane of the epithelial cell. Exosomes can be recovered from the urine by differential centrifugation as a low-density membrane fraction. Exosome isolation can result in marked enrichment of low-abundance urinary proteins that have potential pathophysiologic significance. As a consequence, we and others have been working to define optimal conditions for their isolation and purification as a prelude to their use in biomarker discovery studies [6-7, 33].

In addition, we used a neutral loss scanning approach [34] to investigate the phosphoproteome of human urinary exosomes. Protein phosphorylation is one of the most common posttranslational modifications and it plays an important role in cellular processes.

The purpose of this study was to thoroughly expand the known proteome of human urinary exosomes by using a highly sensitive LC-MS/MS system, improved software for identification of peptide ions and a more elaborate data analysis strategy than in our previous study.

* This chapter was published in part as: Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS, Knepper MA: Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol.* 2009 Feb; 20(2): 363-79.

3.2 Methods

3.2.1 Large-scale proteomic analysis

3.2.1.1 Urinary Exosome Isolation

Urine was collected from eight healthy humans: Four men (aged 22 to 33) and four women (aged 24 to 35) (National Institute of Diabetes and Digestive and Kidney Diseases Clinical Research Protocol 00-DK-0107). Fifty milliliters per subject was collected and mixed together. The urinary exosome isolation procedure is shown below in Figure 3.1. Protease inhibitors were added (1.67 ml of 100 mM NaN₃, 2.5 ml of 11.5 mM 4-[2-aminoethyl] benzenesulfonyl fluoride, and 50 ml of 1 mM leupeptin). The mixed sample was centrifuged at 17,000 x *g* for 10 min at 4°C. The 17,000 x *g* supernatant was ultracentrifuged at 200,000 x *g* for 1 h at 25°C. The ultracentrifugation step was repeated 3 additional times, adding new 17,000 x *g* supernatant volume each time to each of the 12 tubes. Each of the 12 pellets was suspended with 50 ml of “isolation solution” (10mM triethanolamine and 250 mM sucrose). The suspensions were pooled together.

The abundant urinary protein uromodulin or Tamm-Horsfall protein forms very high molecular weight complexes through disulfide linkages. These complexes sediment in the 200,000 x *g* spin unless denatured. To denature the zona pellucida domains in the Tamm-Horsfall protein, we mixed the resuspended pellet with 200 mg/ml dithiothreitol (DTT) at 95°C for 2 min. The resuspended pellet was added to an ultracentrifuge tube, and isolation solution was added to increase the volume to 8 ml. The sample was centrifuged at 200,000 x *g* for 1h at 25°C. The pellet was suspended in 50 ml of isolation solution and frozen at -80°C.

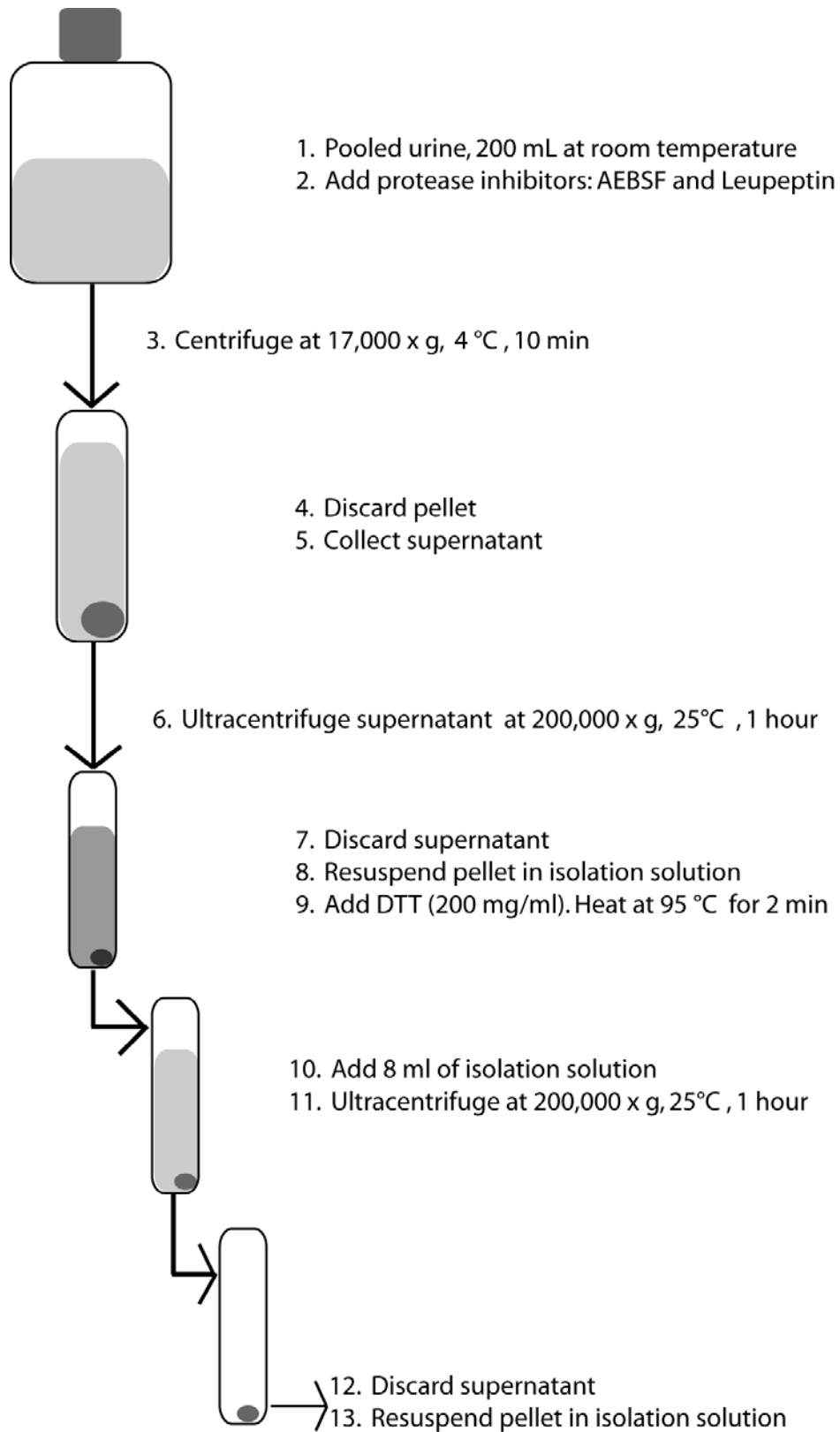


Figure 3.1 Differential centrifugation procedure for the isolation of urinary exosomes from urine.

3.2.1.2 In-Gel Trypsin Digestion

The protein concentration was determined using the Bradford Assay. This sample was solubilized in Laemmli sample buffer (1.5% SDS, 6% glycerol/10 mM Tris HCl, and 60 mg/ml DTT). Proteins in the exosome sample were separated by 1D SDS-PAGE using a Bio-Rad Ready Gel 4 to 15% polyacrylamide gradient gel with 125 μ g distributed among two lanes. The gel was stained with Colloidal Coomassie Blue (GelCode Blue Stain Reagent; Pierce, Rockford IL) for 10 min and destained using ddH₂O (2 x 30 min). The gel was divided from top to bottom into 40 1-mm strips over the entire molecular weight range of the gel. Each strip was diced into small pieces (1 mm³) and placed into labeled centrifuge tubes.

The gels pieces were destained by adding 100 μ l of 25 mM ammonium bicarbonate (NH₄HCO₃)/50% acetonitrile (ACN) for 10 min and were dried using a SpeedVac. The samples were reduced in a solution of 10 mM DTT and 25 mM NH₄HCO₃ at 56°C for 1 h. The samples were alkylated in a solution containing 55 mM iodoacetamide and 25 mM NH₄HCO₃ in the dark at room temperature for 45 min. The gel pieces were washed with 25 mM NH₄HCO₃ and dehydrated in a solution containing 25 mM NH₄HCO₃ and 50% ACN. The samples were dried using the SpeedVac. The samples were rehydrated in a solution containing 12.5 ng/ μ l trypsin (V5113; Promega, Madison, WI) in 25 mM NH₄HCO₃ and digested overnight at 37°C. Peptides were extracted using 50% ACN/0.1% formic acid (FA). The extracted samples were dried using the SpeedVac to remove ACN and then reconstituted with 0.1% FA. All 40 peptide samples were desalted using C₁₈ ZipTips (Millipore, Billerica, MA) before analysis by mass spectrometry.

3.2.1.3 Nanospray LC-MS/MS

A high-sensitivity linear ion trap mass spectrometer, LTQ (Thermo Electron Corp.) equipped with a nanoelectrospray ion source was used to acquire m/z ratios in both precursor ions (MS1) and fragmented ions (MS2) scans. To reduce further the sample complexity before mass analysis, we injected the tryptic peptides extracted from each gel slice using an Agilent 1100 nanoflow system (Agilent Technologies, Palo Alto, CA) into a reversed-phase liquid chromatographic column (PicoFrit, Biobasic C₁₈; New Objective, Woodburn, MA). This LC-MS/MS method allows the acquisition of raw data files that are the MS/MS scans of the five highest intensity peaks after fragmentation with collision-induced dissociation in the LTQ mass analyzer.

3.2.1.4 Analysis of Data

The raw data files were searched against the NCBI Reference Sequences (RefSeq) human protein database by using BIOWORKS software (Thermo Finnigan). BIOWORKS utilizes SEQUEST, which is a program that “finds database candidate sequences whose theoretical spectra are compared with the experimental spectrum [28].” To identify thoroughly peptide sequences, we searched the raw data files using the target-decoy approach and InsPecT.

In addition, we analyzed the data in a two-step process. The first step was to assess and minimize false-discovery peptide identifications using the target-decoy approach, manual inspection of spectra, and InsPecT. The second step was to assess and eliminate ambiguous protein identifications.

3.2.1.5 Target-Decoy

To apply the target-decoy database searching strategy [28], we used the NHLBI

Proteomics Core Facility in-house software to create a composite database containing the forward and reverse sequences of the non-redundant NCBI Reference Sequences (RefSeq) human protein database released on January 26, 2007. We used the forward sequences as the target database and the reversed sequences as the decoy database. We searched the raw data files against this composite database. After the search, we assessed the false discovery rate (FDR) by the number of peptides matched from the reversed sequences. The parameters that determine the stringency of the filtering criteria include XCorr, Sp rank, and delta Cn. These parameters were incrementally adjusted, thereby reducing the false-discovery identifications until a target FDR was achieved. In our case, the data were filtered to a target of 2% FDR, and the actual FDR was 1.91%. The filter settings used were min Xcorr rank 1, min Sp rank 10, min delta Cn 0.08, charge +1 min Xcorr 2.37, charge +2 min Xcorr 2.87, and charge +3 min Xcorr 3.37.

3.2.1.6 InsPecT

We performed an additional analysis of the tandem mass spectrometry data using the InsPecT tool [29]. InsPecT uses *de novo* sequencing to generate sequence information (tag filters) from the experimental data. The tag filters are used to search the human protein database, non-redundant NCBI Reference Sequences (RefSeq) human protein database released on January 26, 2007, and identify peptide sequences that match with the experimental data. The size of the tag filters are three peptides in length on average. As shown in Figure 3.2, the tag filter generated for the protein CHMP1A matches the experimental data accurately. The peptide sequences identified using the tag filters are then scored to estimate that the top match is correct [29]. The score procedure computes the *P* value for each peptide sequence by “comparing the match quality score to

the distribution of quality scores for incorrect matches.” For these data, we accept only peptide matches with $P \leq 0.05$.

R.VYAENAIR.K

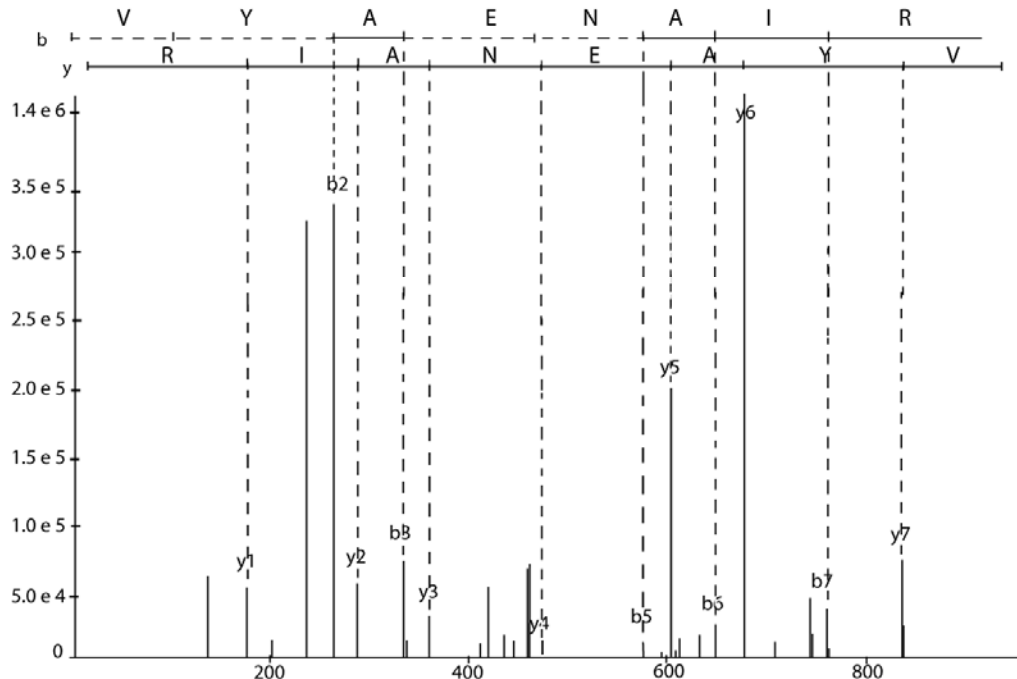


Figure 3.2 Spectrum generated by InsPecT for CHMP1A protein (NP_00279). The peptide sequence is RVYAENAIRK. The tag region for the b-ions and the y-ions are shown by the black solid lines.

3.2.1.7 Minimizing False-Discovery Peptide Identifications

In addition to the target-decoy approach, the InsPecT analysis, we validated the quality of proteins identified by manually checking the spectra of those proteins with one unique peptide. We filtered out the proteins that did not have the expected molecular weight that matched to the corresponding regions in the 1-D SDS PAGE.

3.2.1.8 Elimination of Ambiguous Protein Identifications

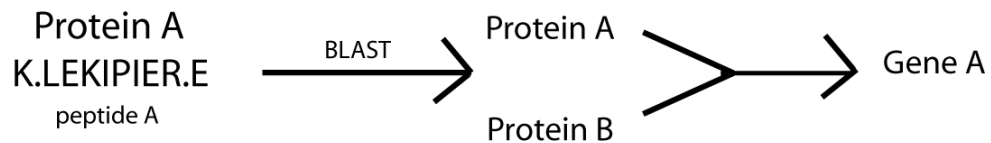
Once proteins were identified using the approaches described, we needed to determine whether all identifications corresponded to unique gene products. An

“ambiguous identification” is defined as an identification for which the peptide sequence that is used to determine the protein identity is found in multiple proteins that are not splice variants of the same gene (Figure 3.3).

A. Unambiguous- Unique Gene Product



B. Unambiguous- Unique Gene Product (multiple splice-variants)



C. Ambiguous Identification- Non-unique Gene Product

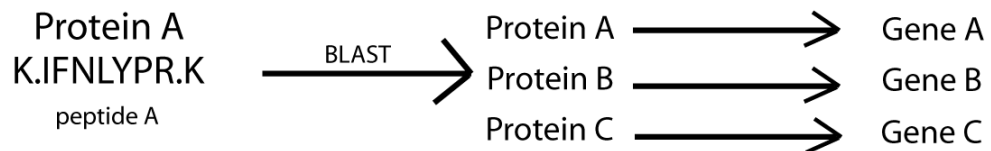


Figure 3.3 Criteria to disambiguate data set. **(A)** An unambiguous identification when the sequence was a 100% match without gaps to one and only one protein. **(B)** An unambiguous identification when the sequence was a 100% match without gaps to more than one protein but these proteins are splice-variant products of one unique gene. **(C)** An ambiguous identification when a peptide sequence was a 100% match without gaps to more than one protein deriving from more than one gene, and the identification was based only on that single peptide.

To disambiguate the data set, we generated software that automates the comparison of each peptide sequence to the protein sequences in the RefSeq Human Protein Database using the BLAST algorithm. An identification was considered unambiguous when the sequence was a 100% match without gaps to one and only one protein (Figure 3.3A). An identification was also considered unambiguous when the sequence was a 100% match

without gaps to more than one protein but these proteins are splice-variant products of one unique gene (Figure 3.3B). An identification was considered ambiguous when a peptide sequence was a 100% match without gaps to more than one protein deriving from more than one gene and the identification was based only on that single peptide (Figure 3.3C). The proteins identified from at least one unambiguous peptide were considered unambiguous proteins. The proteins that contained only ambiguous peptides were considered ambiguous proteins.

3.2.2 Phosphoproteomic Analysis

3.2.2.1 Phosphopeptide Enrichment

We collected urine specimens (200 ml) from six healthy humans, three men and three women. We processed the specimens 400 ml/d for 3 d and pooled them. The exosome isolation was as described previously except that phosphatase inhibitors 10 mM NaF (Sigma, St. Louis, MO), 20 mM β -glycerol phosphate (Fluka, St. Louis, MO), and 1 mM sodium orthovanadate (Sigma) were added. The pellet was resuspended in 6 M guanidine HCl/50 mM NH_4HCO_3 .

The sample was concentrated using a Centricon tube at 13,500 x g, with a starting volume of 420 μl and a final volume of 55 μl . The sample was reduced with 50 mM DTT for 1 h at 56°C. The sample was alkylated by addition of 100 mM iodoacetamide for 1 h (dark) at room temperature and was digested with trypsin overnight at 37°C. The sample was centrifuged at 16,000 x g for 20 min. The supernatant was kept, and 100% FA was added to inactivate the trypsin. The sample was desalted on a 1-ml HLB column (Waters Oasis, Milford, MA) by positive displacement *via* a syringe with a luer adapter. The sample was eluted with two elution buffers. Elution buffer 1 contained 50% ACN

and 0.1% FA, and elution buffer 2 contained 90% ACN and 0.1% FA. The eluents, 50 and 90%, were dried using the SpeedVac.

Phosphopeptides were enriched from the samples using the Pierce Phosphopeptide Isolation Kit (cat. no. 89853) according to the manufacturer's protocol. Phosphopeptide samples were desalted using C18 ZipTips (Millipore) before analysis by mass spectrometry.

Phosphopeptide samples were analyzed on an Agilent 1100 nanoflow system (Agilent Technologies) LC connection to a Finnigan LTQ FT mass spectrometer (Thermo Electron) equipped with a nanoelectrospray ion source as described previously [34]. The five most intense ions were sequentially isolated and fragmented (MS2) in the linear ion trap using collision-induced dissociation. The data-dependent neutral loss algorithm in XCALIBUR software was used to trigger an MS3 scan when a neutral loss of 98.0, 49.0, or 32.7 Da was detected among the two most intense fragment ions in a given MS2 spectrum.

3.2.2.2 Analysis of Phosphopeptide Data Sets

We searched MS raw data files against a composite database containing the forward and reversed peptide sequence of the Human RefSeq Database from January 26, 2007. Putative phosphopeptides were selected and filtered to produce MS2 and MS3 data sets with target FDR of 2% (high stringency) and 20% (low stringency) *via* the PhosphoPIC program [33]. This software was also used to merge MS2 and MS3 data sets into a single file to facilitate subsequent data analysis. Phosphopeptides identified in MS2 spectra were submitted for automated phosphorylation site assignment using the Ascore algorithm [27]. A phosphorylation site with an Ascore ≥ 19 ($\geq 99\%$ confidence)

was considered to be unambiguously assigned. Phosphopeptides present only in MS3 spectra were checked manually. We used Scansite to determine the phosphorylation motif for the identified sites (http://scansite.mit.edu/motifscan_seq.phtml). We searched the PhosphoSite database (<http://www.phosphosite.org>), which is a curated database containing identification of phosphorylation sites from peer-review journals, to determine whether the sites were novel or previously identified.

3.3 Results

3.3.1 Large-Scale Proteomic Profiling of Human Urinary Exosomes

In this study, we carried out proteomic profiling of a low-density membrane fraction from human urine consisting chiefly of exosomes, using a highly sensitive LC-MS/MS system, based on an ion trap mass spectrometer (LTQ; Thermo-Finnigan; Thermo Electron, San Jose, CA). We unambiguously identified 1132 proteins including 205 proteins seen in our previous study and 927 proteins not seen in our previous study of human urinary exosomes [6]. The full list (ambiguous and unambiguous identifications) contains 1412 proteins and can be viewed in Supplemental Table 1 (Appendix B), and the list of proteins that were unambiguously identified in both studies can be viewed at <http://dir.nhlbi.nih.gov/papers/lkem/exosome/>. The expanded list of exosomal proteins includes 177 proteins that are disease related, on the basis of their presence in the OMIM database (Table 3.1).

Predictably, a large number of proteins that were identified were integral membrane proteins involved in solute and water transport (Table 3.2). As seen in our previous study [6], these proteins predominantly represent apical transporters present in every renal tubule segment, including the proximal tubule (sodium-hydrogen exchanger 3, sodium-

glucose co-transporter 1 and 2, and aquaporin-1 [AQP1]), the thick ascending limb (sodium-potassium-chloride co-transporter 2 [NKCC2]), the distal convoluted tubule (thiazide-sensitive Na-Cl co-transporter [NCC]), and connecting tubule/collecting duct (AQP2, rhesus blood group C glycoprotein [RhCG, an ammonia channel], B1 subunit of vacuolar H⁺-ATPase, and pendrin). Note that both polycystin-1 and polycystin-2 were detected in human urinary exosomes.

Exosomes derive from MVB and are delivered to the urine when the outer membranes of MVB fuse with the apical plasma membrane. Interestingly, 22 of the proteins identified in this study are recognized as components of the apparatus responsible for the formation of MVB (Table 3.3). These 22 proteins account for approximately 75% of the proteins that constitute the ESCRT-0, ESCRT-I, ESCRT-II, and ESCRT-III complexes involved in multivesicular body formation [16].

In addition, 17 proteins identified in this study are subunits of the human vacuolar H⁺-ATPase (Table 3.4). Vacuolar H⁺ATPases are ATP-dependent proton pumps for proton transport into intracellular organelles [36]. These proteins also mediate proton transport across the apical plasma membrane of type A intercalated cells and across the basolateral plasma membrane of type B intercalated cells [37]. The B1 subunit is selectively expressed in intercalated cells and its detection in urinary exosomes establish that intercalated cells secrete exosomes as do other types of epithelial cells lining the renal tubule. These proteins constitute 78% of the subunits of the V₀ and V₁ domains of the vacuolar H⁺-ATPase [38].

3.3.2 Phosphoproteomic Analysis of Human Urinary Exosomes

Protein phosphorylation is a key element of most cell regulatory processes.

Recently, technical approaches that allow phosphoproteomic profiling on a large scale have been introduced [26, 34, 35, 39]. We used neutral loss scanning with high-stringency target-decoy analysis to identify phosphorylation sites present in exosomal proteins from human urine samples.

Nineteen phosphorylation sites corresponding to 14 phosphoproteins were identified (Table 3.5). These included both newly identified phosphorylation sites and sites that had been previously identified. Two orphan G-protein-coupled receptors are included in the former group, *viz.* GPRC5B and GPRC5C. In GPRC5B, we identified one new phosphorylation site, T389, and, in GPRC5C, we identified three new phosphorylation sites, T435, S395, and Y426. These proteins are also known as retinoic acid-induced gene 2 (GPRC5B) and retinoic acid-induced gene 3 (GPRC5C).

A new phosphorylation site was also identified in the COOH-terminal tail of the thiazide-sensitive co-transporter (NCC) at S811 (Figure 3.4). This site is distinct from the N-terminal site previously identified [40] and may play a role in regulation of transport. This amino acid is conserved in humans, chimpanzees, rhesus monkeys, and horses but not in mice and rats. Simon *et al* [41] showed that in rat, the amino acid sequence surrounding this site is absent owing to a difference in exon splicing.

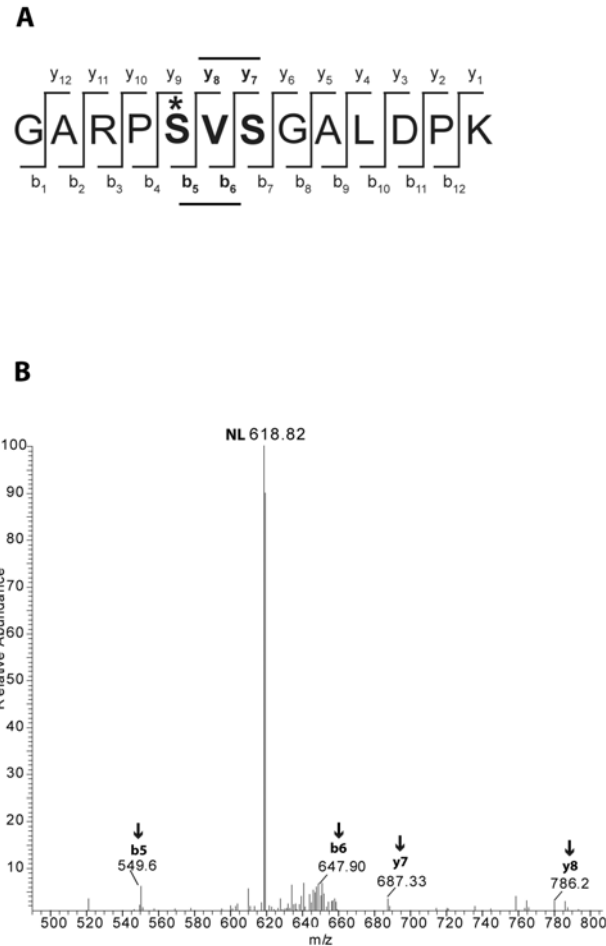


Figure 3.4 A novel phosphorylation site in the NCC protein. The serine-811 on the NCC protein is phosphorylated. (A) The phosphorylation site on the peptide is denoted by an asterisk (*). (B) The neutral loss peak (NL) from the +2 mass spectrum and the site-determining ions, b₅, b₆, y₇ and y₈.

Novel phosphorylation sites were also identified in RRAS2 (TC21), VPS4B (an ESCRT component), cytochrome b reductase, proteasome-3 subunit, and mucin 1. This study also revealed previously identified phosphorylation sites in AQP2 (S256) [42], NKCC2 (T118 and S120) [43], CHMP2B (S199) [44], HSP90AB1 (S255) [45], and SPP1 (S192, S197, and S207) [46]. Phosphorylation of AQP2 at S256 was confirmed by immunoblotting human urinary exosomes samples with a phosphospecific antibody for this site (Figure 3.5).

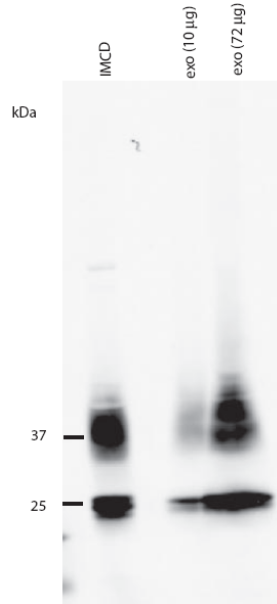


Figure 3.5 Detection of AQP2-S256 phosphorylation in urinary exosomes. IMCD, rat inner medullary collecting duct treated with dDAVP (V2R-selective vasopressin analog) for 30 min; exo (10 µg) is human urinary exosomes, 10 µg; exo (72 µg), human urinary exosomes, 72 µg.

3.4 Discussion

3.4.1 Large-Scale Proteomic Profiling of Human Urinary Exosomes

One of the objectives of this study was to expand the existing human urinary exosome database by using a higher sensitivity LC-MS/MS mass spectrometer and improved computational tools for matching spectra to proteins in the human proteome. The LTQ mass analyzer has an increased trapping efficiency, ion capacity, and ion ejection rate compared with the LCQ mass analyzer [6] used in our previous study. We identified the peptide sequences using the SEQUEST program and analyzed them using the target-decoy database search strategy and the InsPecT tool. The target-decoy database search strategy allows adjustment of SEQUEST search parameters to ensure a given false-discovery rate (FDR) [28]. The InsPecT tool uses *de novo* sequencing to generate tag filters, which are then used to search the database to “look for any peptide that matches

the tag [29].” The data have been made available to the general public and can be downloaded from our laboratory’s web-site (<http://dir.nhlbi.nih.gov/papers/lkem/exosome/>). In addition, the database can be searched using the BLAST algorithm.

3.4.2 Relevance to Renal Biology

Several of the proteins newly identified in urinary exosomes in this study may have considerable relevance to renal biology and the mechanism of renal disease. Our previous study [6] identified proteins that were characteristic of most of the cell types facing the urinary space from podocytes through transitional epithelial cells of the urinary drainage system. In this study, we identified markers of two additional cell types, type A and B intercalated cells. Specifically, the B1 subunit of the H^+ -ATPase is apically located in type A intercalated cells [47], and the anion transporter pendrin is present in type B intercalated cells [48]. Previously, we showed that urinary exosomes are derived from the apical endosomal pathway so that, although the B1 subunit of the H^+ -ATPase is also expressed in type B intercalated cells, its basolateral location probably precludes delivery to urinary exosomes. Overall, we identified 17 different vacuolar H^+ -ATPase subunits in urinary exosomes, 78% of the whole V_0-V_1 complex [36].

We identified all of the subunits of the four ESCRT complexes (ESCRT-0 through ESCRT-III) in urinary exosomes in this study. The ESCRT complexes play a central role in the formation of MVB and the secretion of exosomes [38].

Four different orphan G-protein-coupled receptors were identified in urinary exosomes in this study, namely GPR98, GPRC5A, GPRC5B, and GPRC5C. These receptors are presumably apically located in one or more renal tubule cells. GPR98 (also

known as very large G-protein–coupled receptor 1 or Neurepin) has more than 6000 amino acids. The three GPRC5 proteins are members of the metabotropic glutamate family, but their natural ligands are unknown. It will be of interest in future studies to discover the role of these proteins in renal development and regulation.

3.4.3 Phosphoproteomic Analysis of Human Urinary Exosomes

Posttranslational modifications (PTM) of proteins play an important role in protein function. Among the most important phosphorylation sites identified included six that were previously identified in peer-review journals and eight that had not been previously identified. Among the novel sites was serine-811 in the NCC protein. This amino acid is conserved in humans, chimpanzees, rhesus monkeys, and horses but not in mice and rats. The amino acid sequence surrounding this site is absent in rodents owing to a difference in exon splicing [41]. Finally, AQP2 phosphorylated at serine-256 was readily detectable in urinary exosomes. Because this phosphorylation event is increased by vasopressin-stimulated activation of adenylyl cyclase [42], measurements of the amount of serine 256–phosphorylated AQP2 in urine may provide an improved means of assessing the state of vasopressin activation using phospho-specific antibodies.

3.5 Conclusion

We used a highly sensitive LC-MS/MS system, based on an ion trap mass spectrometer (LTQ; Thermo-Finnigan; Thermo Electron, San Jose, CA) to analyze the human urinary exosomes proteome. We unambiguously identified 1132 proteins including 205 proteins seen in our previous study and 927 proteins not seen in our previous study of human urinary exosomes [6]. The data is available to the general public and can be downloaded from the web-site

(<http://dir.nhlbi.nih.gov/papers/lkem/exosome/>). In addition, the database can be searched using the BLAST algorithm. The expanded list of exosomal proteins includes 177 proteins that are disease related and are potential targets for further biomarkers studies.

Protein phosphorylation regulates cellular signaling processes and may determine protein structure, function, and sub-cellular localization [39]. The ability to detect PTM, such as phosphorylation, in urinary exosomes may provide an additional level of information that could aid in diagnosis and treatment of a variety of renal disorders. Furthermore, discovery of PTM in urinary exosomes can provide clues about physiologic and pathophysiologic mechanism. In this study, we identified 14 phosphoproteins. The specific phosphorylation sites identified included six that were previously identified and eight that had not been previously identified. Among the novel sites was serine-811 in the NCC protein. This amino acid is conserved in humans, chimpanzees, rhesus monkeys, and horses but not in mice and rats. The amino acid sequence surrounding this site is absent in rodents owing to a difference in exon splicing [39]. In addition, AQP2 phosphorylated at serine-256 was readily detectable in urinary exosomes.

Table 3.1. Disease Related Proteins in Human Urinary Exosomes

Information for each protein include "Gene" name, "Protein Name", "Pep" refers to the number of unique peptides identified in LC-MS/MS, "ID" refers to the number of spectra and "Related to Disease [OMIM]" refers to the disease with which the protein is related according to OMIM. The 34 proteins associated with kidney diseases are presented in italics.

Gene	Protein Name	Pep	ID	Related to Disease (OMIM)
ABCB1	ATP-binding cassette sub-family B member 1	23	61	Colchicine Resistance [MIM : 120080] Crohn Disease [MIM:266600]
ABCC9	ATP-binding cassette, sub-family C, member 9 isoform SUR2A-delta-14	1	2	Cardiomyopathy [MIM: 608569]
ABCB11	ATP-binding cassette, sub-family B (MDR/TAP), member 11	1	3	Cholestasis, Progressive Familial Intrahepatic 2 [MIM:601847] Cholestasis, Benign Recurrent Intrahepatic 2 [MIM:605479]
ACAT1	acetyl-Coenzyme A acetyltransferase 1 precursor	1	2	Alpha-methylacetoacetic Aciduria [MIM:203750]
<i>ACE</i>	<i>angiotensin I converting enzyme isoform 1 precursor</i>	23	96	<i>Hypertension [MIM:106180]</i>
<i>ACE</i>	<i>angiotensin I converting enzyme isoform 2 precursor</i>	12	61	<i>Renal Tubular Dysgenesis [267430]</i>
<i>ACE2</i>	<i>angiotensin I converting enzyme 2 precursor</i>	8	17	<i>Hypertension [MIM:300335]</i>
ACOT7	acyl-CoA thioesterase 7 isoform hBACHd	1	1	Mesial temporal lobe epilepsy [MIM:608096]
ACSL4	acyl-CoA synthetase long-chain family member 4 isoform 2	1	2	Mental Retardation, X-linked 63, MRX 63 [MIM:300387]
ACY1	aminoacylase 1	15	43	Aminoacylase 1 Deficiency [MIM:609924]
AHCY	S-adenosylhomocysteine hydrolase	10	28	Hypermethioninemia [MIM:180960]
AK1	adenylate kinase 1	4	4	Hemolytic Anemia due to AK1 Deficiency [MIM:103000]
ALAD	delta-aminolevulinic acid dehydratase isoform a	1	1	Acute Hepatic Porphyria [MIM:125270]
ALB	albumin precursor	36	139	Dysalbuminemic Hyperthyroxinemia Hyperthyroxinemia, dysalbuminemic Analbuminemia Bisalbuminemia [MIM:103600]
ALDOA	aldolase A	7	14	Aldolase Deficiency of Red Cells Myopathy and Hemolytic Anemia [MIM:103850]
ALPL	tissue non-specific alkaline phosphatase precursor	3	4	Hypophostasia [MIM:241500]
AMN	amionless protein precursor	1	1	Megaloblastic Anemia 1 [MIM: 261100]
<i>ANPEP</i>	<i>membrane alanine aminopeptidase precursor</i>	69	412	<i>Hypertension [MIM:151530]</i>
APOA1	apolipoprotein A-I preproprotein	6	17	Primary Hypoalphalipoproteinemia [MIM:604091]
APOA2	apolipoprotein A-II preproprotein	1	1	Apolipoprotein A-II Deficiency, Familial Hypercholesterolemia, Familial [MIM:143890]
<i>APRT</i>	<i>adenine phosphoribosyltransferase isoform a</i>	2	2	<i>2,8-Dihydroxyadenine urolithiasis [MIM:102600]</i>
<i>APRT</i>	<i>adenine phosphoribosyltransferase isoform b</i>	3	10	<i>2,8-Dihydroxyadenine urolithiasis [MIM:102600]</i>
AQP1	aquaporin 1	3	35	Aquaporin 1 Deficiency, Colton-Null [MIM:110450]
<i>AQP2</i>	<i>aquaporin 2</i>	7	36	<i>Autosomal recessive nephrogenic diabetes insipidus, type 1 [MIM:222000]</i> <i>Autosomal dominant nephrogenic diabetes insipidus, type 1 [MIM:125800]</i>
ARL6	ADP-ribosylation factor-like 6	4	7	Bardet-Biedl Syndrome 3 [MIM:209900]
ARSE	arylsulfatase E precursor	1	2	Chondrodysplasia Punctata 1, X-linked Recessive [MIM:302950]
ASAH1	N-acylsphingosine amidohydrolase (acid ceramidase) 1 preproprotein isoform a	7	16	Farber Disease [MIM:228000]
ASAH1	N-acylsphingosine amidohydrolase (acid ceramidase) 1 isoform b	9	34	Farber Disease [MIM:228000]
ASL	argininosuccinate lyase isoform 3	1	1	Argoninosuccinic Aciduria [MIM:207900]
ASS1	argininosuccinate synthetase 1	20	59	Citrullinemia [MIM:215700]
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	1	1	Aica-ribosiduria due to Atic Deficiency [MIM:608688]
<i>ATP6V0A4</i>	<i>ATPase, H+ transporting, lysosomal V0 subunit a4</i>	1	2	<i>Renal Tubular Acidosis, Distal, Autosomal Recessive [MIM:602722]</i>
<i>ATP6V1B1</i>	<i>ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B1</i>	8	17	<i>Renal Tubular Acidosis, Distal, with Progressive Deafness [MIM:267300]</i>
B2M	beta-2-microglobulin precursor	1	1	Hypercatabolic Hypoproteinemia [MIM:241600]

Table 3.1 Continued

Gene	Protein Name	Pep	ID	Related to Disease (OMIM)
B4GALT1	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase 1, membrane-bound form	1	1	Congenital Disorder of Glycosylation Type Iid [MIM:607091]
CA2	<i>carbonic anhydrase II</i>	9	25	<i>Autosomal recessive syndrome of osteopetrosis with renal tubular acidosis</i> [MIM:259730]
CA4	<i>carbonic anhydrase IV precursor</i>	2	2	<i>Proximal renal tubular acidosis</i> [MIM:114760]
CC2D1A	coiled-coil and C2 domain containing 1A	6	6	Mental Retardation Autosomal Recessive 3 [MIM:608443]
CD2AP	CD2-associated protein	14	21	Focal Segmental Glomerulosclerosis FSGS3 [MIM:607832]
CETP	cholesteryl ester transfer protein, plasma precursor	7	14	Cholesterol Ester Transfer Protein Deficiency [MIM:607322]
CFH	complement factor H isoform b precursor	1	1	Hemolytic Uremic Syndrome, Atypical [MIM: 235400]
CFI	complement factor I	1	1	Complement Factor I Deficiency [MIM: 610984]
CHMP2B	chromatin modifying protein 2B	4	15	Frontotemporal Dementia, Chromosome 3-linked [MIM:6000795]
CLTC	<i>clathrin heavy chain 1</i>	12	24	<i>Renal Cell Carcinoma</i> [MIM:118955]
COL18A1	alpha 1 type XVIII collagen isoform 1 precursor	1	1	Knobloch Syndrome [MIM:267750]
COL6A1	collagen, type VI, alpha 1 precursor	6	21	Bethlem Myopathy [MIM:158810] Ullrich Congenital Muscular Dystrophy, Autosomal Dominant [MIM:254090]
COL6A3	alpha 3 type VI collagen isoform 5 precursor	1	2	Ullrich Congenital Muscular Dystrophy [MIM: 254090]
CP	ceruloplasmin precursor	6	15	Aceruloplasminemia [MIM:604290]
CRYAB	crystallin, alpha B	7	12	Alpha-B Crystallinopathy [MIM:608810]
CRYM	crystallin, mu isoform 1	1	3	Autosomal Dominant Nonsyndromic Deafness [MIM:123740]
CST3	cystatin C precursor	1	3	Icelandic-type cerebroarterial amyloidosis [MIM:105150]
CSTB	cystatin B	2	10	Myoclonic Epilepsy of Unverricht and Lundborg [MIM:254800]
CTSC	cathepsin C isoform b precursor	1	1	Papillo-LeFevre Syndrome [MIM: 245000]
CTH	cystathionase isoform 2	1	1	Cystathionuria [MIM:219500]
CTSA	cathepsin A precursor	3	15	Galactosialidosis [MIM:256540]
CTSC	cathepsin C isoform a preproprotein	1	2	Papillon-Lefevre Syndrome [MIM:245000]
CTSD	cathepsin D preproprotein	1	2	Neuronal Ceroid Lipofuscinosis [MIM:610127]
CTNS	<i>cystinosis, nephropathic isoform 1</i>	1	3	<i>Nephropathic Cystinosis</i> [MIM: 219800]
CUBN	cubilin	104	672	Megaloblastic Anemia 1, Finnish Type [MIM:261100]
CUL4B	cullin 4B	1	1	Cabezas Syndrome [MIM:300354] Mental Retardation-Hypotonic Facies Syndrome [MIM:300639]
DDAH1	<i>dimethylarginine dimethylaminohydrolase 1</i>	1	5	<i>Hypertension</i> [MIM: 604743]
DDC	dopa decarboxylase (aromatic L-amino acid decarboxylase)	11	31	Aromatic L-Amino Acid Decarboxylase Deficiency [MIM:608643]
DNM2	dynamin 2 isoform 3	1	1	Charcot-Marie-Tooth Disease, Dominant Intermediate B [MIM:606482]
DNM2	dynamin 2 isoform 4	1	1	Charcot-Marie-Tooth Neuropathy, Dominant Intermediate B [MIM: 606482]
DYSF	dysferlin	1	1	Miyoshi Myopathy [MIM: 254130]
DPYS	dihydropyrimidinase	5	6	Dihydropyrimidinuria [MIM:222748]
DSC2	desmocollin 2 isoform Dsc2b preproprotein	1	3	Arrhythmogenic right ventricular dysplasia-11 [MIM:610476]
DSP	desmoplakin isoform II	10	16	Keratosis Palmoplantaris Striata II Dilated Cardiomyopathy With Woolly Hair and Keratoderma [MIM:605676]
ECE1	endothelin converting enzyme 1	1	1	Hirschsprung Disease [MIM: 142623]

Table 3.1 Continued

Gene	Protein Name	Pep	ID	Related to Disease (OMIM)
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1 precursor	1	3	Doyne Honeycomb Retinal Dystrophy [MIM: 126600]
ELA2	elastase 2, neutrophil preproprotein	1	1	Cyclic Hematopoiesis [MIM:162800]
ENPEP	<i>glutamyl aminopeptidase (aminopeptidase A)</i>	25	100	Hypertension [MIM:138297]
FAH	fumarylacetoacetate hydrolase (fumarylacetoacetase)	2	2	Tyrosinemia Type I [MIM:276700]
FLNB	filamin B, beta (actin binding protein 278)	1	1	Spondylocapotalarsal Synostosis Syndrome [MIM: 272460]
FBP1	fructose-1,6-bisphosphatase 1	7	16	Fructose-1,6-Bisphosphatase Deficiency [MIM:229700]
FGA	fibrinogen, alpha polypeptide isoform alpha-E preproprotein	5	17	Renal Amyloidosis [MIM:105200] Dysfibrinogenemia [MIM:134820]
FGG	fibrinogen, gamma chain isoform gamma-A precursor	1	1	Dysfibrinogenemia[MIM:134850]
FTCD	formiminotransferase cyclodeaminase	4	7	Glutamate Formiminotransferase Deficiency [MIM: 229100]
FTH1	ferritin, heavy polypeptide 1	1	7	Iron Overload, Autosomal Dominant [MIM:134770]
FTL	ferritin, light polypeptide	5	14	Hyperferritinemia-Cataract Syndrome [MIM:600886]
FUCA1	fucosidase, alpha-L-1, tissue	1	1	Fucosidosis [MIM:230000]
FXYD2	<i>FXYD domain-containing ion transport regulator 2 isoform 1</i>	1	7	Hypomagnesemia 2, Renal [MIM:154020]
G6PD	glucose-6-phosphate dehydrogenase isoform a	1	1	Nonspherocytic Hemolytic Anemia due to G6PD Deficiency [MIM:305900]
GAA	acid alpha-glucosidase preproprotein	4	8	Infantile-onset Glycogen Storage Disease Type II [MIM:232300]
GALK1	galactokinase 1	1	1	Galactokinase Deficiency [MIM: 230200]
GBE1	glucan (1,4-alpha-), branching enzyme 1	1	2	Type IV Glycogen Storage Disease [MIM:232500]
GCS1	mannosyl-oligosaccharide glucosidase	1	1	Congenital Disorder of Glycosylation [MIM: 606056]
GK	glycerol kinase isoform a	1	1	Glycerol Kinase Deficiency [MIM: 307030]
GLB1	galactosidase, beta 1 isoform a	16	70	Gangliosidosis GM1 [MIM:230500]
GLUL	glutamine synthetase	2	2	Congenital Glutamine Deficiency [MIM:610015]
GM2A	GM2 ganglioside activator precursor	2	3	Gangliosidosis GM2 AB Variant Tay-Sachs Disease [MIM:272750]
GPI	glucose phosphate isomerase	9	19	Chronic Hemolytic Anemia Duet to GPI Deficiency [MIM:172400]
GPR98	G protein-coupled receptor 98 precursor	1	1	Familial Febrile Seizures [MIM:604352] Usher Syndrome Type IIC [MIM:605472]
GSN	gelsolin isoform b	10	21	Finnish Type Familial Amyloidosis [MIM:105120]
GSS	glutathione synthetase	1	3	Glutathione Synthetase Deficiency [MIM:266130]
HNMT	histamine N-methyltransferase isoform 1	1	1	Susceptibility to Asthma [MIM:600807]
HPD	4-hydroxyphenylpyruvate dioxygenase	1	1	Tyrosinemia Type III [MIM:276710]
HPGD	<i>hydroxyprostaglandin dehydrogenase 15-(NAD)</i>	6	21	Hypertension [MIM:601688]
HSPB1	heat shock 27kDa protein 1	8	30	Charcot-Marie-Tooth Disease, Type 2F [MIM:606595] Distal Hereditary Motor Neuropathy IIB [MIM:608634]
ICAM1	intercellular adhesion molecule 1 precursor	1	1	Graves Disease [MIM: 275000]
IL1RN	interleukin 1 receptor antagonist isoform 1 precursor	2	2	Gastic Cancer Risk [MIM:137215]
IRF6	interferon regulatory factor 6	1	1	Van der Woude Syndrome [MIM:119300] Popliteal Pterygium Syndrom [MIM:119500]
ITM2B	integral membrane protein 2B	5	21	Familial Dementia [MIM:176500]
JUP	junction plakoglobin	9	15	Naxos Disease [MIM:601214]
KALRN	kalirin, RhoGEF kinase isoform 3	1	1	Coronary Heart Disease [MIM:608901]
KHK	ketohexokinase isoform a	2	4	Essential Fructosuria [MIM:229800]
KL	klotho	1	1	Hyperphosphatemic Tumoral Calcinosis [MIM: 211900]
KLK1	kallikrein 1 preproprotein	1	1	Decreased Urinary Activity of Kallikrein [MIM:147910]
LGALS3	galectin 3	1	1	Lymphocyte Function-Associated Antigen 1 [MIM: 116920]
LAMP2	lysosomal-associated membrane protein 2 precursor	4	26	Danon Disease [MIM:300257]
LRRK2	leucine-rich repeat kinase 2	4	5	Parkinson Disease [MIM:607060]
LYZ	lysozyme precursor	1	3	Familal Visceral Amyloidosis [MIM:105200]

Table 3.1 Continued

Gene	Protein Name	Pep	ID	Related to Disease (OMIM)
MIF	macrophage migration inhibitory factor (glycosylation-inhibiting factor)	1	6	Rheumatoid Arthritis [MIM: 604302]
MME	membrane metallo-endopeptidase Neprilysin	48	311	Hypertension Important cell surface marker in the diagnostic of human acute lymphocytic leukemia [MIM:120520]
MPO	myeloperoxidase	7	32	Myeloperoxidase Deficiency [MIM:254600]
MTHFD1	methylenetetrahydrofolate dehydrogenase 1	4	5	Spina Bifida [MIM:601634]
MYH14	myosin, heavy chain 14 isoform 1	1	1	Autosomal Dominant Nonsyndromic Sensorineural Deafness [MIM: 600652]
MYH14	myosin, heavy chain 14 isoform 2	8	17	Autosomal Dominant nonsyndromic sensorineural hearing loss [MIM:600652]
MYH3	myosin, heavy chain 3, skeletal muscle, embryonic	1	1	Freeman-Sheldon Syndrome [MIM:193700]
MYH9	myosin, heavy polypeptide 9, non-muscle	19	51	Fechtner syndrome [MIM:153640] Epstein syndrome [MIM:153650]
MYO15A	myosin XV	1	3	Recessive Congenital Deafness [MIM:600316]
MYO6	myosin VI	7	21	Autosomal Recessive Congenital Sensorineural Deafness [MIM:607821] Autosomal Dominant Nonsyndromic Sensorineural Deafness [MIM:606346]
NAGLU	alpha-N-acetylglucosaminidase precursor	21	63	Mucopolysaccharidosis Type IIIB [MIM:252920]
NDRG1	N-myc downstream regulated gene 1	2	5	Charcot-Marie-Tooth Disease Type 4D [MIM:601455]
NEB	nebulin	2	4	Nemaline Myopathy [MIM:256030]
NPHS2	podocin	6	9	Autosomal recessive steroid-resistant nephrotic syndrome [MIM:600995]
PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit (45kD)	1	1	Miller-Dieker Lissencephaly Syndrome [MIM:607432]
PARK7	DJ-1 protein	1	1	Parkinson Disease 7, Autosomal Recessive [MIM: 606324]
PCBD1	pterin-4 alpha-carbinolamine dehydratase precursor	1	1	Hyperphenylalaninemic [MIM:264070]
PDCD10	programmed cell death 10	2	3	Cerebral Cavernous Malformations [MIM:603285]
PHGDH	phosphoglycerate dehydrogenase	2	2	Phosphoglycerate Dehydrogenase Deficiency [MIM:601815]
PKD1	polycystin 1	1	1	Polycystic Kidney Disease, adult, type I [MIM:601313]
PKD2	polycystin 2	1	2	Polycystic kidney disease, adult, type II [MIM:173910]
PKHD1	polyductin isoform 2	6	9	Autosomal Recessive Polycystic Kidney Disease [MIM:263200]
PKLR	pyruvate kinase, liver and RBC isoform 1	1	1	Pyruvate Kinase Deficiency [MIM:266200]
PLOD1	lysyl hydroxylase precursor	1	1	Ehlers-Danlos Syndrome, Type VIA [MIM:225400]
PRKCH	protein kinase C, eta	1	2	Cerebral Infarction [MIM:601367]
PROM1	prominin 1	23	174	Autosomal Recessive Retinal Degeneration [MIM:604365]
PRNP	prion protein preproprotein	1	1	Creutzfeldt-Jakob Disease [MIM: 123400]
PSAP	prosaposin isoform a preproprotein	3	6	Metachromatic Leukodystrophy Due to SAP1 Deficiency [MIM:249900] Gaucher Disease, Atypical Due to SAP2 Deficiency [MIM:610539]
PSAP	prosaposin isoform c preproprotein	1	4	Metachromatic Leukodystrophy [MIM: 249900]
PSAT1	phosphoserine aminotransferase isoform 1	2	4	Phosphoserine Aminotransferase Deficiency [MIM: 610992]
PTPRJ	protein tyrosine phosphatase, receptor type, J precursor	1	1	Somatic Colon Cancer [MIM:114500]
RAB3GAP1	RAB3 GTPase-activating protein	1	1	Warburg Micro Syndrome #600118
RBP4	retinol-binding protein 4, plasma precursor	2	3	Retinol-Binding Protein Deficiency [MIM:180250]
RDX	radixin	16	23	Autosomal Recessive Deafness 24 [MIM:611022]
ROBO2	roundabout, axon guidance receptor, homolog 2	1	1	Vesicoureteral Reflux 2 [MIM:610878]
RP2	XRP2 protein	3	5	X-Linked Retinitis Pigmentosa 2 [MIM:312600]

Table 3.1 Continued

RYR1	skeletal muscle ryanodine receptor isoform 1	1	1	Malignant Hyperthermia [MIM:145600] Central Core Disease [MIM:117000] Minicore Myopathy with External Ophthalmoplegia [MIM:255320]
Gene	Protein Name	Pep	ID	Related to Disease (OMIM)
SERPING1	complement component 1 inhibitor precursor	7	14	Hereditary Angioedema Type I [MIM:106100]
SLC3A1	<i>solute carrier family 3, member 1</i>	14	25	<i>Cystinuria</i> [MIM:220100]
SLC4A1	<i>solute carrier family 4, anion exchanger, member 1 [kAE1]</i>	2	2	<i>Defective kidney acid secretion leading to distal renal tubular acidosis</i> [MIM:179800]
SLC4A4	<i>solute carrier family 4, sodium bicarbonate cotransporter, member 4 [NBC1]</i>	2	3	<i>Renal Tubular Acidosis, Proximal, With Ocular Abnormalities</i> [MIM:604278]
SLC5A1	<i>solute carrier family 5 (sodium/glucose cotransporter), member 1 [SGLT1]</i>	2	3	<i>Glucose/Galactose Malabsorption</i> [MIM:606824]
SLC5A2	<i>solute carrier family 5 (sodium/glucose cotransporter), member 2 [SGLT2]</i>	4	9	<i>Renal Glucosuria</i> [MIM:233100]
SLC6A19	<i>solute carrier family 6, member 19</i>	4	8	<i>Hartnup Disorder</i> [MIM:234500]
SLC12A1	<i>sodium potassium chloride cotransporter 2 [NKCC2]</i>	25	94	<i>Bartter Syndrome, Antenatal, Type 1</i> [MIM:601678]
SLC12A3	<i>solute carrier family 12 (sodium/chloride transporters), member 3 [NCC]</i>	28	102	<i>Gitelman Syndrome</i> [MIM:263800]
SLC22A12	<i>urate anion exchanger 1 isoform a [URAT1]</i>	1	2	<i>Renal Hypouricemia</i> [MIM:220150]
SLC25A3	solute carrier family 25 member 3 isoform b precursor	1	5	Mitochondrial Phosphate Carrier Deficiency [MIM:610773]
SLC26A4	<i>pendrin</i>	2	4	<i>Pendred Syndrome</i> [MIM:274600] <i>Deafness, Autosomal Recessive 4</i> [MIM:600791]
SLC44A4	NG22 protein isoform 1	6	59	Sialidosis 1 [MIM:606107]
SPR	sepiapterin reductase (7,8-dihydrobiopterin:NADP+ oxidoreductase)	1	2	Dystonia, Dopa-Responsive, Due To Sepiapterin Reductase Deficiency [MIM:251120]
SQSTM1	sequestosome 1	1	2	Paget Disease of Bone [MIM:602080]
SUCLA2	succinate-CoA ligase, ADP-forming, beta subunit	1	1	Mitochondrial DNA Depletion Syndrome [MIM: 609560]
TECTA	tectorin alpha precursor	1	1	Autosomal Dominant nonsyndromic sensorineural hearing loss [MIM:601842]
TF	transferrin	12	20	Alzheimer Disease [MIM:104300]
TPP1	tripeptidyl-peptidase I preproprotein	8	42	Ceroid Lipofuscinosis Neuronal 2 [MIM:204500]
TSG101	tumor susceptibility gene 101	17	66	Breast Cancer [MIM:176960]
TTN	titin isoform novex-1	4	5	Cardiomyopathy [MIM:188840]
UMOD	<i>uromodulin precursor</i>	35	1278	<i>Medullary cystic kidney disease-2 (MCKD2)</i> [MIM:603860] <i>Familial juvenile hyperuricemic nephropathy (FJHN)</i> [MIM:16200]
VCP	valosin-containing protein	2	2	Inclusion Body Myopathy with Early Onset Paget Disease and Frontotemporal Dementia [MIM:167320]
VAMP7	vesicle-associated membrane protein 7	1	1	Beta-Ureidopropionase Deficiency [MIM: 606673]
VCL	vinculin isoform meta-VCL	3	5	Cardiomyopathy, Dilated [MIM: 611407]
VWF	von Willebrand factor preproprotein	1	4	Von Willebrand Disease [MIM: 193400]
ZMPSTE24	zinc metalloproteinase STE24	1	1	Madibuloacral Dysplasia [MIM: 608612]

Table 3.2. Solute and water transporters

Table contains all of the proteins that are solute and water transporters.

Ref Seq	Gene	Protein Name	Pep	ID
NP_000918	ABCB1	ATP-binding cassette sub-family B member 1	23	61
NP_003733	ABCB11	ATP-binding cassette, sub-family B (MDR/TAP), member 11	1	3
NP_005680	ABCB6	ATP-binding cassette, sub-family B, member 6	1	1
NP_064694	ABCC9	ATP-binding cassette, sub-family C, member 9 isoform SUR2A-delta-14	1	2
NP_149163	ABCC11	ATP-binding cassette, sub-family C, member 11 isoform a	1	1
NP_932766	AQP1	aquaporin 1	3	35
NP_000477	AQP2	aquaporin 2	7	36
NP_000692	ATP1A1	Na ⁺ /K ⁺ -ATPase alpha 1 subunit isoform a proprotein	19	57
NP_001001787	ATP1B1	Na ⁺ /K ⁺ -ATPase beta 1 subunit isoform b	1	1
NP_001001937	ATP5A1	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, alpha subunit precursor	3	5
NP_001677	ATP5B	ATP synthase, H ⁺ transporting, mitochondrial F1 complex, beta subunit precursor	4	5
NP_001174	ATP6AP1	ATPase, H ⁺ transporting, lysosomal accessory protein 1 precursor	1	1
NP_001685	ATP6V0C	ATPase, H ⁺ transporting, lysosomal, V0 subunit c	1	9
NP_005168	ATP6V0A1	ATPase, H ⁺ transporting, lysosomal V0 subunit a isoform 1	1	1
NP_065683	ATP6V0A4	ATPase, H ⁺ transporting, lysosomal V0 subunit a4	1	2
NP_004682	ATP6V0D1	ATPase, H ⁺ transporting, lysosomal, V0 subunit d1	1	1
NP_689778	ATP6V0D2	ATPase, H ⁺ transporting, lysosomal 38kDa, V0 subunit D2	2	4
NP_001681	ATP6V1A	ATPase, H ⁺ transporting, lysosomal 70kD, V1 subunit A, isoform 1	22	49
NP_001683	ATP6V1B1	ATPase, H ⁺ transporting, lysosomal 56/58kDa, V1 subunit B1	8	17
NP_001684	ATP6V1B2	vacuolar H ⁺ ATPase B2	12	27
NP_001686	ATP6V1C1	ATPase, H ⁺ transporting, lysosomal 42kDa, V1 subunit C1 isoform A	1	1
NP_001034451	ATP6V1C2	vacuolar H ⁺ ATPase C2 isoform a	1	1
NP_057078	ATP6V1D	H ⁽⁺⁾ -transporting two-sector ATPase	2	3
NP_001687	ATP6V1E1	vacuolar H ⁺ ATPase E1 isoform a	2	2
NP_001034456	ATP6V1E1	vacuolar H ⁺ ATPase E1 isoform c	1	1
NP_001034455	ATP6V1E1	vacuolar H ⁺ ATPase E1 isoform b	1	1
NP_004222	ATP6V1F	ATPase, H ⁺ transporting, lysosomal 14kD, V1 subunit F	1	1
NP_004879	ATP6V1G1	vacuolar H ⁺ ATPase G1	1	2
NP_998784	ATP6V1H	ATPase, H ⁺ transporting, lysosomal 50/57kDa, V1 subunit H isoform 1	8	34
NP_036415	KCNG2	potassium voltage-gated channel, subfamily G, member 2	1	2
NP_853514	PKD1L3	polycystin 1-like 3	1	1
NP_001009944	PKD1	polycystin 1 isoform 1 precursor	1	1
NP_000288	PKD2	polycystin 2	1	2
NP_057405	RHCG	Rhesus blood group, C glycoprotein	5	8
NP_000531	RYR1	skeletal muscle ryanodine receptor isoform 1	1	1
NP_006505	SCN10A	sodium channel, voltage-gated, type X, alpha	1	11
NP_054858	SCN11A	sodium channel, voltage-gated, type XI, alpha	1	1
NP_000329	SLC12A1	sodium potassium chloride cotransporter 2	25	94
NP_000330	SLC12A3	solute carrier family 12 (sodium/chloride transporters), member 3	28	102
NP_064631	SLC12A9	solute carrier family 12 (potassium/chloride transporters), member 9	1	1
NP_003975	SLC13A2	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	5	10
NP_073740	SLC13A3	solute carrier family 13 member 3 isoform a	2	2
NP_001011554	SLC13A3	solute carrier family 13 member 3 isoform b	1	3
NP_004161	SLC1A1	solute carrier family 1, member 1	3	6
NP_066568	SLC15A2	solute carrier family 15 (H ⁺ /peptide transporter), member 2	1	1
NP_060954	SLC22A11	solute carrier family 22 member 11	2	9
NP_653186	SLC22A12	urate anion exchanger 1 isoform a	1	2
NP_003049	SLC22A2	solute carrier family 22 member 2 isoform a	2	3
NP_003051	SLC22A5	solute carrier family 22 member 5	1	1
NP_004781	SLC22A6	solute carrier family 22 member 6 isoform a	1	3
NP_695010	SLC22A6	solute carrier family 22 member 6 isoform d	1	1

Table 3.2 Continued

Ref Seq	Gene	Protein Name	Pep	ID
NP_004245	SLC22A8	solute carrier family 22 member 8	1	2
NP_005838	SLC23A1	solute carrier family 23 (nucleobase transporters), member 1 isoform a	4	6
NP_689898	SLC23A1	solute carrier family 23 (nucleobase transporters), member 1 isoform b	6	10
NP_005975	SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	1	1
NP_998776	SLC25A3	solute carrier family 25 member 3 isoform b precursor	1	5
NP_775897	SLC26A11	solute carrier family 26, member 11	1	1
NP_000432	SLC26A4	pendrin	2	4
NP_003030	SLC2A5	solute carrier family 2 (facilitated glucose/fructose transporter), member 5	12	22
NP_775867	SLC39A5	solute carrier family 39 (metal ion transporter), member 5	1	1
NP_000332	SLC3A1	solute carrier family 3, member 1	14	25
NP_001012679	SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a	10	18
NP_002385	SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform c	3	5
NP_536856	SLC44A1	CDW92 antigen	1	1
NP_065161	SLC44A2	CTL2 protein	18	70
NP_000333	SLC4A1	solute carrier family 4, anion exchanger, member 1	2	2
NP_003750	SLC4A4	solute carrier family 4, sodium bicarbonate cotransporter, member 4	2	3
NP_000334	SLC5A1	solute carrier family 5 (sodium/glucose cotransporter), member 1	2	3
NP_689564	SLC5A10	solute carrier family 5 (sodium/glucose cotransporter), member 10 isoform 1	2	2
NP_001035915	SLC5A10	solute carrier family 5 (sodium/glucose cotransporter), member 10 isoform 2	1	3
NP_848593	SLC5A12	solute carrier family 5 (sodium/glucose cotransporter), member 12 isoform 2	2	4
NP_003032	SLC5A2	solute carrier family 5 (sodium/glucose cotransporter), member 2	4	9
NP_666018	SLC5A8	solute carrier family 5 (iodide transporter), member 8	1	1
NP_001011547	SLC5A9	solute carrier family 5 (sodium/glucose cotransporter), member 9	1	3
NP_057699	SLC6A13	solute carrier family 6 (neurotransmitter transporter, GABA), member 13	1	1
NP_001003841	SLC6A19	solute carrier family 6, member 19	4	8
NP_004165	SLC9A3	solute carrier family 9 (sodium/hydrogen exchanger), isoform 3	2	3
NP_004776	SLC9A3R2	solute carrier family 9 isoform 3 regulator 2	1	1
NP_851322	SLCO4C1	solute carrier organic anion transporter family, member 4C1	2	2
NP_003365	VDAC1	voltage-dependent anion channel 1	6	43
NP_005653	VDAC3	voltage-dependent anion channel 3	1	1

Table 3.3. Proteins of the ESCRT Complex

Table contains all of the proteins that are members of the ESCRT Complex:

ESCRT-0, ESCRT-I, ESCRT-II, ESCRT-III, ATPase complex and accessory.

Gene	Protein	Pep	ID	Ref Seq	ESCRT
HGS	Hepatocyte growth factor-regulated tyrosine kinase substrate	1	1	NP_004703	ESCRT-0
TSG101	Tumor susceptibility gene 101	15	47	NP_006283	ESCRT-I
VPS28	Vacuolar protein sorting 28 isoform 1	5	8	NP_057292	ESCRT-I
VPS28	Vacuolar protein sorting 28 isoform 2	2	8	NP_898880	ESCRT-I
VPS37B	Vacuolar protein sorting 37B	4	10	NP_078943	ESCRT-I
VPS37C	Vacuolar protein sorting 37C	1	1	NP_060436	ESCRT-I
VPS25	EAP25	4	15	NP_115729	ESCRT-II
VPS36	EAP45	2	3	NP_057159	ESCRT-II
SNF8	EAP30	1	1	NP_009172	ESCRT-II
CHMP2A	CHMP2A	6	40	NP_055268	ESCRT-III
CHMP2B	CHMP2B	3	11	NP_054762	ESCRT-III
VPS24	CHMP3	1	3	NP_057163	ESCRT-III
VPS24	CHMP3	1	4	NP_001005753	ESCRT-III
CHMP4B	CHMP4B	2	6	NP_789782	ESCRT-III
CHMP5	CHMP5	2	7	NP_057494	ESCRT-III
CHMP1A	CHMP1A	1	3	NP_002759	ESCRT-III
CHMP1B	CHMP1B	1	2	NP_065145	ESCRT-III
CHMP6	CHMP6	2	3	NP_078867	ESCRT-III
VPS4A	Vacuolar protein sorting factor 4A	11	25	NP_037377	ATPase Complex
VPS4B	Vacuolar protein sorting factor 4B	11	32	NP_004860	ATPase Complex
PDCD6IP	ALIX	27	104	NP_037506	Accessory
C1orf58	Hypothetical protein LOC148362	11	34	NP_653296	Accessory

Table 3.4. Vacuolar H-ATPase subunits in human urinary exosomes

Table contains proteins that are found in human urinary exosomes and are subunits of the human vacuolar H⁺-ATPase.

Ref Seq	Gene	Protein Name	Pep	ID
NP_005168	ATP6V0A1	ATPase, H+ transporting, lysosomal V0 subunit a isoform 1	1	1
NP_065683	ATP6V0A4	ATPase, H+ transporting, lysosomal V0 subunit a4	1	2
NP_001685	ATP6V0C	ATPase, H+ transporting, lysosomal, V0 subunit c	1	1
NP_004682	ATP6V0D1	ATPase, H+ transporting, lysosomal, V0 subunit d1	1	1
NP_689778	ATP6V0D2	ATPase, H+ transporting, lysosomal 38kDa, V0 subunit D2	2	4
NP_001681	ATP6V1A	ATPase, H+ transporting, lysosomal 70kD, V1 subunit A, isoform 1	22	49
NP_001683	ATP6V1B1	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B1	8	17
NP_001684	ATP6V1B2	vacuolar H+ATPase B2	12	27
NP_001686	ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C1 isoform A	1	1
NP_001034451	ATP6V1C2	vacuolar H+ ATPase C2 isoform a	1	1
NP_057078	ATP6V1D	H(+)-transporting two-sector ATPase	2	3
NP_001687	ATP6V1E1	vacuolar H+ ATPase E1 isoform a	2	2
NP_001034455	ATP6V1E1	vacuolar H+ ATPase E1 isoform b	1	1
NP_001034456	ATP6V1E1	vacuolar H+ ATPase E1 isoform c	1	1
NP_004222	ATP6V1F	ATPase, H+ transporting, lysosomal 14kD, V1 subunit F	1	1
NP_004879	ATP6V1G1	vacuolar H+ ATPase G1	1	2
NP_998784	ATP6V1H	ATPase, H+ transporting, lysosomal 50/57kDa, V1 subunit H isoform 1	8	34

Table 3.5. Human urinary exosomes phosphopeptides

Table contains phosphopeptides found in urinary exosomes. MSn refers to spectra for phosphorylation site identification.

Motif refers to phosphorylation motif site and GO Function.

Ref Seq	Protein Name, Sequence	Site	Gene	Novel Site	MSn	Motif	GO Function
NP_061123	G protein-coupled receptor family C, group 5, member C isoform b		GPRC5C				metabotropic glutamate, GABA-B-like receptor activity
	R.AEDMYSAQSHQAA(T*)PPKDGK.N	T435		Yes	MS2, MS3	proline-directed	protein binding
	K.VP(S*)EGAYDILPR.A	S395		Yes	MS2, MS3	acidophilic	receptor activity
	R.AEDM(Y*)SAQSHQAATPPKDGK.N	Y426		Yes	MS2	tyrosine kinase	
NP_001035149	Secreted phosphoprotein 1 isoform c		SPP1				cytokine activity
	K.AIPVAQDLNAPSDWD(S*)R.G	S192		No	MS2	Miscellaneous	growth factor activity
	R.GKD(S*)YETSQLDDQSAETHSHK.Q	S197			MS2, MS3	Miscellaneous	integrin binding
	R.GKDSYETSQLDDQ(S*)AETHSHK.Q	S207			MS2, MS3	acidophilic	protein binding
NP_057319	G protein-coupled receptor, family C, group 5, member B precursor		GPRC5B				metabotropic glutamate, GABA-B-like receptor activity
	R.SNVYQPTEMA VVLNGG(T*)IPTAPPSHTGR.H	T389		Yes	MS2	basophilic	receptor activity sevenless binding
NP_000477	Aquaporin 2		AQP2				transporter activity
	R.RQ(S*)VELHSPQSLPR.G	S256		No	MS2,MS3		water channel activity
NP_004860	Vacuolar protein sorting factor 4B		VPS4B				
	K.EGQSPADEKGNDS(S*)DGEGESDDPEKKK.L	S102		Yes	MS2	acidophilic	
NP_054762	Chromatin modifying protein 2B		CHMP2B				not classified
	K.ATI(S*)DEEIER.Q	S199		No	MS2, MS3 (unfiltered)	basophilic	
NP_687033	Proteasome alpha 3 subunit isoform 2		PSAM3				protein binding
	K.ESLKEEDE(S*)DDDNM	S243		No	MS2, MS3	acidophilic	threonine endopeptidase activity
NP_036382	Related RAS viral (r-ras) oncogene homolog 2		RRAS2				GTP binding
	R.KFQEQCPEP(S*)PEPTRK.E	S186		Yes	MS2, MS3	DNA damage kinase	nucleotide binding protein binding
NP_031381	Heat shock 90kDa protein 1, beta		HSP90AB1				ATP binding
	K.IEDVG(S*)DEEDDSGKDKK.K	S255		No	MS2	basophilic	nitric-oxide synthase regulator activity nucleotide binding TPR domain binding unfolded protein binding

Ref Seq	Protein Name, Sequence	Site	Gene	Novel Site	MSn	Motif	GO Function
NP_612433	Kinesin family member 12 R.VTTRPQAPK(S*)PVAK.Q	S236	KIF12	Yes	MS2, MS3	proline-directed	ATP binding microtubule motor activity nucleotide binding
NP_079119	Cytochrome b reductase 1 R.NLALDEAGQRS(T*)M.	T285	CYBRD1	Yes	MS2	acidophilic	ferric-chelate reductase activity
NP_001037857	Mucin 1 isoform 7 precursor R.DTYHPMSEYPTYH(T*)HGR.Y	T118	MUC1	Yes	MS2, MS3 (unfiltered)	N/A	NF-kappaB binding protein heterodimerization activity protein homodimerization activity RNA binding Tat protein binding unfolded protein binding
NP_000330	Solute carrier family 12 (sodium/chloride transporters), member 3 R.GARP(S*)VSGALDPK.A	S811	SLC12A3	Yes	MS2	basophilic	sodium ion binding sodium:chloride symporter activity symporter activity transporter activity
NP_000329	Sodium potassium chloride cotransporter 2 K.IEYYRN(T*)GSISGPK.V K.IEYYRNTG(S*)ISGPK.V	T118 S120	SLC12A1	No No	MS2, MS3 MS2, MS3	N/A N/A	potassium ion binding sodium ion binding sodium:potassium:chloride symporter activity symporter activity transporter activity

4 Application of urinary exosomes analysis using polyclonal antibodies

4.1 Introduction

4.1.1 Bartter syndrome type I *

The potential clinical usefulness of urinary exosomes was demonstrated using the well-defined renal tubulopathy, Bartter syndrome type I, as an example. Patients with Bartter syndrome type I, have mutations in the SLC12A1 gene, which encodes for the NKCC2 sodium potassium chloride cotransporter 2 protein. The NKCC2 protein is expressed on the apical membrane of the thick ascending limb of Henle (TAL) [48]. It is the main cotransporter of sodium and chloride in the TAL [49].

Patients with Bartter syndrome type I become apparent already in utero or later in infancy [50]. Clinical features of Bartter syndrome type I include dehydration with low blood pressure, seizures, tetany (involuntary contraction of muscles), muscular weakness, paresthesias (feeling of numbness of the hands and feet) and joint pain due to chondrocalcinosis (calcification within tissues) [51]. Physiologic features include hypercalciuria, which results in nephrocalcinosis, deposition of calcium in the kidney, and renal failure [50, 51]. The mutation in the NKCC2 cotransporter leads to loss of function, resulting in salt wasting in the thick ascending limb and volume contraction which also affects the absorption of calcium and magnesium, resulting in hypermagnesiuria and hypercalciuria [50].

Patients with Bartter syndrome type I should not receive a treatment with thiazides

* This portion of the chapter was published as: Gonzales PA, Pisitkun T, Hoffert JD, Tchapyjnikov D, Star RA, Kleta R, Wang NS, Knepper MA. Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol.* 2009 Feb; 20(2): 363-79.

because it would block the salt reabsorption in the distal convoluted tubule (DCT), which will deprive the kidney of the compensatory mechanism that works through the thiazide-sensitive cotransporter (NCC) protein [50]. The NCC protein is functionally expressed on the apical membrane of the distal convoluted tubule of the kidney [49, 52]. It reabsorbs sodium and chloride and it is “essential for sodium chloride conservation in the kidney” [52]. The purpose of this part of the research is to use the anti-NKCC2 and anti-NCC polyclonal antibodies to qualitatively analyze samples obtained from patients clinically diagnosed with Bartter syndrome type I.

4.1.2 Mutation of ABCC4 gene

In addition, we generated a polyclonal antibody by designing a synthetic peptide corresponding to the sequence of the multidrug resistance protein 4, MRP4. The gene coding the MRP4 protein, ABCC4, is located in chromosome 13 and its product has a length of 1,325 amino acids. The multidrug resistance protein 4 is a member of the C subfamily ATP-binding cassette (ABC) transporters. The main characteristics of the ABCC transporters are the ability to bind and to hydrolyze ATP to drive translocation of their substrates across membranes against a diffusion gradient [53]. The protein contains two membrane-spanning domains and each consists of six transmembrane helices with two cytosolic ATP-binding domains [53]. It is an organic anion transporter localized in the apical proximal tubule of the kidney.

The ABCC4 gene contains 31 exons and it is highly polymorphic with at least 25 non-synonymous SNPs [53]. In general, a person’s genome contains unique SNPs that are not responsible for a disease state but have the potential to be disease biomarkers. When a SNP causes a disease, it is important to determine its effect on the gene and the

gene product. The specific non-synonymous polymorphism on the ABCC4 gene has the potential to affect the protein expression. One type of mutation in the gene is due to the single nucleotide polymorphism (SNP) of thymine (T) to cytosine (C) located in exon 2, IVS2-5T>C. This SNP causes the deletion of 121 base pairs and results in a short gene product (67 amino acids) and a normal gene product (1,325 amino acids). Thus, the patients that have this homozygous splice site mutation in the ABCC4 gene abnormally express a short-form of protein in addition to a normal protein. The purpose of the collaboration was to develop an affinity-purified antibody that would recognize the short and normal gene products and to qualitatively analyze the isolated urinary exosomes from these patients.

4.1.3 Designing anti-human polyclonal antibodies

The implementation of proteomic studies in urinary exosomes requires the use of suitable antibodies. These antibodies can be monoclonal and polyclonal. Polyclonal antibodies are produced against synthetic peptides (20-25 amino acids in length) corresponding to the protein of interest which are called immunizing peptides. The characteristics used to select candidates for immunizing peptides are immunogenicity, specificity and the lack of post-translational modifications [54].

The immunogenicity of a peptide sequence in a protein is determined by the content of charged amino acids, which are aspartic acid (D), arginine, glutamic acid (R), lysine (K) and histidine (H) and polar amino acids which are serine (S), threonine (T), tyrosine (Y), asparagines (N), and glutamine (Q). In general it is preferable to have the least number of non-polar amino acids; alanine (A), valine (V), leucine (L), isoleucine (I), methionine (M), phenylalanine (F), tryptophan (W), cysteine (C), glycine (G). A

proline (P) moiety is helpful because the peptide sequence is less likely to have secondary structures, α -helices and β -sheets, that would be obstructive. We used programs such as Emboss Lite (<http://helixweb.nih.gov/emboss-lite/protein.html>) to aid in the design of these synthetic peptides.

The Kyte-Doolittle hydrophathy index measures the hydrophilicity and hydrophobicity of the amino acids in the protein sequence. Hydrophilic amino acids, charged or polar amino acids, are most likely to be found on the exterior of the protein and most likely to be immunogenic. Hydrophobic amino acids, which are non-polar, are most likely to be found in the interior of the proteins or in the transmembrane region of the cell and less likely to be immunogenic.

The specificity of the antibody is determined by the level of recognition of the intended target protein. A BLAST analysis (National Center for Biotechnology Information, www.ncbi.nlm.nih.gov/BLAST/) is used to determine if the peptide sequence is specific for the intended target protein. The peptide sequence of the protein is searched for motifs that would indicate possible post-translational modifications near or on the site of the epitope.

We proposed to qualitatively assess patients' samples by generating a polyclonal antibody to bind to an epitope in the amino-terminal of the human MRP4 protein; and to use previously generated polyclonal antibodies for NKCC2 and NCC. The purpose of producing polyclonal antibodies to recognize epitopes in human proteins is for the future use in proteomic studies involving immunoblotting. Our hypothesis is that the affinity-purified polyclonal antibody will recognize the truncated and normal MRP4 protein in the

urinary exosomes samples obtained from the patients; and that the Bartter syndrome type I patient samples will show the absence of the NKCC2 protein band.

4.2 Methods and Materials

4.2.1 Patients with Bartter Syndrome Type I

We collected spot urine samples from two patients with clinically diagnosed Bartter syndrome type I. The patients were enrolled in the institutional review board–approved protocol 76-HG-0238. We obtained written informed consent from the parents and/or patient. We collected urine samples from healthy humans and used them as controls. We processed all samples using the differential centrifugation method to isolate human urinary exosomes described already. Each sample was prepared for immunoblotting by solubilizing in Laemmli buffer (1.5% SDS, 6% glycerol, 10 mM Tris HCl, and 60 mg/ml DTT). The samples, patient 1 and patient 2, and the control samples, control 1 and control 2, were loaded onto a 1-D SDS-PAGE gel on the basis of time as measured by creatinine excretion. The proteins were transferred to Immobilon-P (Millipore) membranes, blocked, and probed with antigen-specific NKCC2 [55] and NCC primary antibodies. We incubated the blots with species-specific fluorescence secondary antibodies (Alexa 688) and visualized them using the Odyssey Infrared Imaging System (LiCor, Lincoln, NE).

4.2.2 Patients with mutation in ABCC4 gene

We collected spot urine samples (100 to 180 ml) from two patients with a splice site variation in the ABCC4 gene. We obtained written informed consent from the patients. We collected urine samples from healthy humans and used them as controls. We processed all samples using the differential centrifugation method to isolate human

urinary exosomes described already. Each sample was prepared for immunoblotting by solubilizing in Laemmli buffer (1.5% SDS, 6% glycerol, 10 mM Tris HCl, and 60 mg/ml DTT).

The samples, patient 3 and the control, were loaded onto a 1-D SDS-PAGE gel and the samples, patient 4, and the control, were loaded onto another 1-D SDS-PAGE gel on the basis of time as measured by creatinine excretion. The proteins were transferred to Immobilon-P (Millipore) membranes, blocked, and probed with a commercially available antigen-specific MRP4 primary antibody (Abcam). We incubated the blots with species-specific fluorescence secondary antibodies (Alexa 688) and visualized them using the Odyssey Infrared Imaging System (LiCor, Lincoln, NE). The commercially available MRP4 antibody is a rat monoclonal antibody corresponding to amino acids 372-431 of human MRP4.

The samples, patient 3, the control, patient 4, and the control were loaded onto a 1-D SDS-PAGE gel on the basis of time as measured by creatinine excretion. The proteins were transferred to nitrocellulose membranes, blocked and probed with the antigen-specific polyclonal MRP4 primary antibody. We incubated the blot with species-specific fluorescence secondary antibody (Alexa 688) and visualized it using the Odyssey Infrared Imaging System (LiCor, Lincoln, NE).

4.2.3 Polyclonal antibodies

4.2.3.1 Multidrug resistant protein 4

The MRP-4 protein contains 1,325 amino acids, and we designed two synthetic peptides with epitopes in the first 67 amino acids of the protein. The synthetic peptide 1 starts at position 2 and ends at position 19. The sequence is

LPVYQEYKPNPLQDANLC. The synthetic peptide 2 starts at position 32 and ends at position 53. The sequence is KIGHKRRLEEDDMYSVLPEDRS.

The online software program Emboss Lite (<http://helixweb.nih.gov/emboss-lite>, previously known as GCG-Lite) was used to determine the hydropathy (hydrophilicity and hydrophobicity) parameters. We used the ‘pepPlot’ program to output the hydropathy parameters. We used the ‘Motifs’ program to determine the most common post-translational motifs found in the sequence of the protein. We used the protein BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) analysis to determine whether these peptide sequences are unique to the MRP4 protein.

The peptides were synthesized and purified by high performance liquid chromatography at Lofstrand Labs Limited (Gaithersburg, MD). A cysteine residue was added to the carboxyl terminal which enabled the peptide to be covalently attached to the resin for affinity purification of the antibody. The cysteine residue enables the peptide to be conjugated to the carrier protein, keyhole limpet haemocyanin (KLH), for the immunizing injection. The immunization of the rabbits was done at Antibodies Incorporated (Davis, CA). The rabbits were immunized using Freund’s complete adjuvant and the serum was collected as bleeds for affinity-purification of the polyclonal antibody.

4.2.3.2 Thiazide-sensitive co-transporter

The thiazide-sensitive co-transporter, NCC, protein is located in the distal convoluted tubule of the nephron in the kidney. The protein expressed in humans contains 1,030 amino acids; the synthetic peptide was designed to bind to an epitope

found in the human protein. The sequence starts at position 76 and ends at position 97, and it is PGEPRKVRPTLADLHSFLKQEG.

4.2.3.3 Sodium potassium chloride co-transporter 2

The sodium potassium chloride co-transporter 2, NKCC2, protein is located in the thick ascending limb of the nephron in the kidney. The protein expressed in humans contains 1,099 amino acids. The synthetic peptide for the polyclonal antibody, L320, was designed against the NKCC2 protein expressed in rats which contains 1,095 amino acids [52]. The sequence starts at position 33 and it ends at position 55. The sequence is SDSTDPPHYEETSFGDEAQNRLK.

4.2.4 Affinity purification of polyclonal antibodies

The synthetic peptides which contain terminal cysteine residues are immobilized permanently on a column resin prior to purification of polyclonal antibody from serum. The resin is a custom-made affinity resin for purification of antibodies. The resin is porous, crosslinked and it contains 6% beaded agarose activated with iodoacetyl groups. These iodoacetyl groups react with the exposed sulfhydryls (-SH) on the cysteine residues of the synthetic peptides to form covalent and irreversible thioether bonds.

We prepared the peptide columns using 2 mg of the synthetic peptide and followed the protocol of the SulfoLink Kit® (Pierce). Rabbit 316 was immunized with synthetic peptide 1 and rabbit 319 was immunized with synthetic peptide 2. The rabbit serum bleed was diluted in an equal volume of PBS, and were added to the peptide column in 2 ml aliquots. The sample was left in the column for at least 15 minutes during which the antibody binds to the antigen. This procedure is repeated until the

entire sample is added to the column, then the column is washed to get rid of unbound sample. The purified antibody is eluted from the column in 1mL fractions, with an acidic elution buffer, 0.1M glycine-HCl at pH 2.5-3.0. The eluted sample is collected in tubes containing 200 μ l Tris buffer (pH 8, 1M) to neutralize elution. We collected seven fractions and monitored by measuring the optical density at 280 nm. We expect the optical density to be the highest for fractions 3, 4 and 5. Once the optical density was measured and we added 15 μ l of 100 mM sodium azide to preserve the purified antibody.

4.2.5 Characterization of polyclonal antibody

The MRP4 antibody was characterized by immunoblotting where we expect to see a predominant band in the tissue in which it is shown to be expected. We obtained tissue sample from the human kidney cortex and outer medulla, and a predominant band is expected at the glycosylated molecular weight of 170 kD. The molecular weight without glycosylation is 146 kD. The predominant band is expected in the membrane fractions and not in cytosol of the human kidney cortex and outer medulla. The predominant band is not expected in the tissue sample when the antibody was pre-absorbed with the immunizing peptide.

4.3 Results

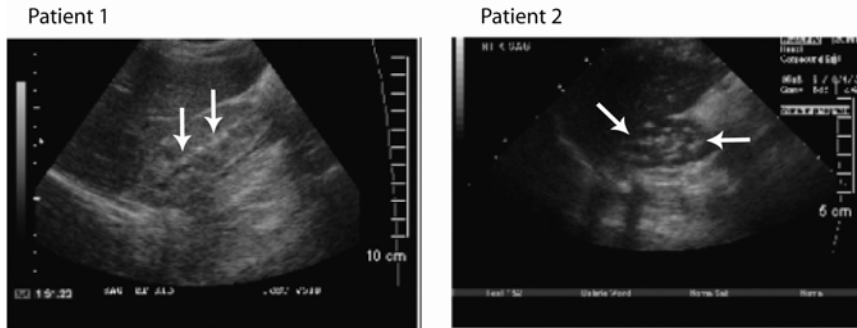
4.3.1 Bartter syndrome type I

An example of the utility of exosome analysis is shown in Figure 4.1, describing immunoblotting in patients with Bartter syndrome type I, associated with mutations in the SLC12A1 gene, which encodes for the NKCC2 sodium-potassium-chloride cotransporter protein [51]. Urine samples were obtained from two patients (patients 1 and 2) with clinical phenotypes consistent with Bartter syndrome type I (Figure 4.1A) [50].

A. Clinical Phenotypes

Patient	Unique Phenotype	Premature Birth	Hypokalemia	Renin	Aldosterone	Urinary PGE2	Gender	Age
1	nephrocalcinosis	yes	yes	elevated	elevated	elevated	Female	20y10m
2	nephrocalcinosis	yes	yes	elevated	elevated	elevated	Female	3y7m

B. Ultrasound Images



C. Immunoblots

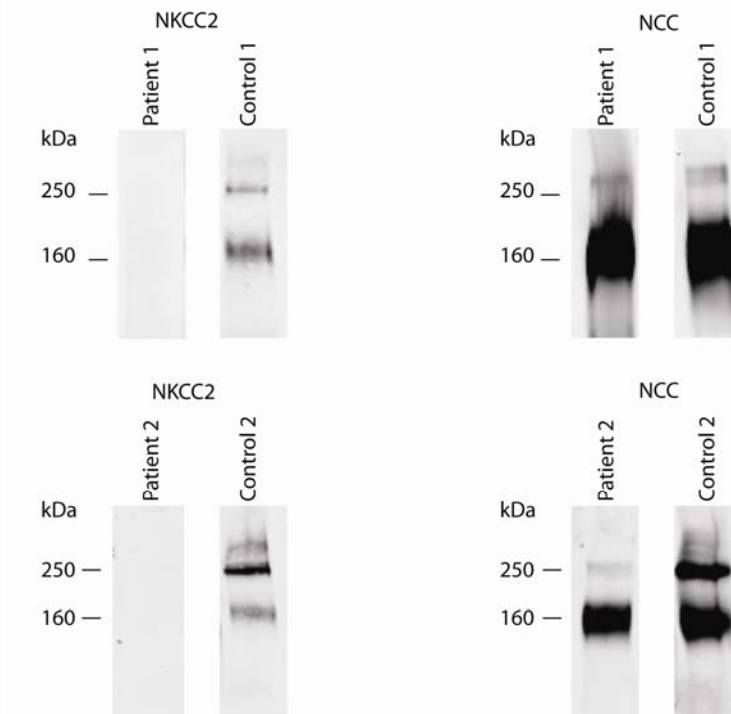


Figure 4.1 Disease-related protein NKCC2 and Bartter syndrome Type 1. **(A)** Details of clinical phenotype for patients with Bartter syndrome type I, patient 1 and patient 2. **(B)** Ultrasound images showing calcium deposits (white arrowheads) in the kidneys of patients 1 and 2. **(C)** Immunoblot of urinary exosomes samples from patient 1, patient 2, control 1 and control 2 using polyclonal rabbit anti-NKCC2 and NCC.

The clinical diagnosis for the patients with Bartter syndrome type I was confirmed

by the ultrasound images showing deposits of calcium in the kidney also known as nephrocalcinosis [50, 51] (Figure 4.1B) and other typical laboratory findings. The two urinary exosome samples obtained from the patients with Bartter syndrome type I were analyzed by immunoblotting for the presence of the NKCC2 protein (Figure 4.1C). Compared with the respective control samples, patients 1 and 2 showed an absence of the NKCC2 protein bands, expected at 160 kD for monomeric NKCC2 and 320kD for dimeric NKCC2. In addition, the samples (patients 1 and 2) were probed for the thiazide-sensitive co-transporter (NCC) protein to ensure that urinary exosomes were successfully isolated and loaded properly. Strong NCC bands were obtained in samples from both patients with Bartter syndrome type I and control samples.

The polyclonal antibodies used for the immunoblotting of the patient samples were generated against synthetic peptide sequences in the NKCC2 and NCC proteins. The anti-NKCC2 polyclonal antibody was generated against a 23-amino acid synthetic peptide at the amino-terminal of the protein in the rat proteome. This antibody was previously characterized in kidney tissue samples obtained from rats [55]. The synthetic peptide sequence was analyzed using the BLAST analysis tool to determine if the polyclonal antibody would recognize the epitope in the human protein. The peptide sequence matches 19 amino acids out of 21 amino acids found in the human NKCC2 protein. The anti-NCC polyclonal antibody was generated against a 22-amino acid synthetic peptide of the protein in the human proteome. The sequence matches the NCC protein in human, horse, rat, cow, chimpanzee, rhesus monkey and dogs. The anti-NCC polyclonal antibody was previously characterized (data not shown).

4.3.2 Mutation of ABCC4 gene

Urinary exosomes samples from two patients with a mutation of the ABCC4 gene, which produces a short-form and normal MRP4 protein, were analyzed by immunoblotting. A commercially-available monoclonal antibody against MRP4 was used to detect expression changes in MRP4 protein (Figure 4.2). The commercial monoclonal antibody, anti-MRP4, was generated to recognize an epitope in the carboxyl-terminal of the protein, amino acids 372-431. We generated a polyclonal antibody to bind to an epitope in the amino-terminal of the MRP4 protein. Synthetic peptide 1 targets amino acids 2 to 19 and synthetic peptide 2 targets amino acids 32 to 53 of the MRP4 protein.

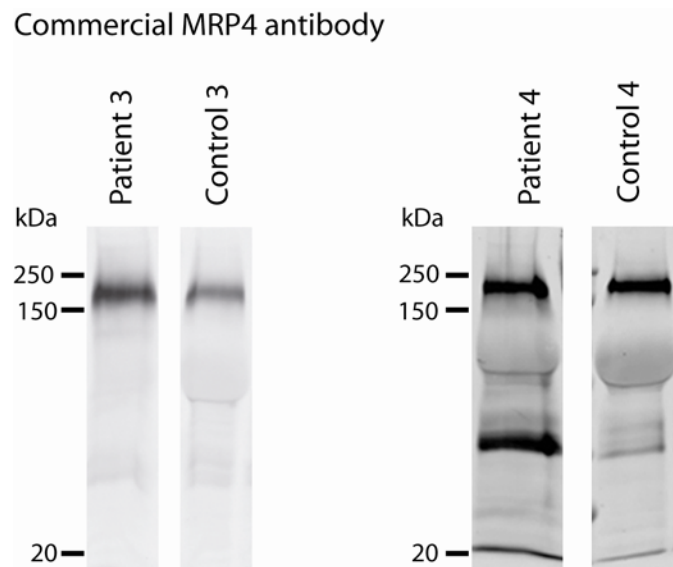


Figure 4.2 Genetic mutation of ABCC4 results in short-form and normal MRP4 protein. Patient samples (patient 3 and patient 4) and control samples (control 3 and control 4) were probed with commercial monoclonal anti-MRP4 antibody. Expected molecular weight for normal protein is 170 kD and for the short protein is between 15 and 20 kD.

The synthetic peptide 1 has 6 amino acids that are either charged or polar. Also, it contains three prolines, which are helpful in preventing free rotation and there is less likelihood for secondary structures. The synthetic peptide 2 has 14 amino acids that are

either charged or polar. It contains one proline and there are only four non-polar amino acids. The software program, Emboss Lite, uses the Kyte-Doolittle method for calculating the hydropathy index, hydrophilic and hydrophobic tendencies, evaluated systematically for each amino acid in the sequence [56]. The hydropathy index was calculated for the first 60 amino acids which includes the sequence for the epitopes, amino acids 2 to 19 (synthetic peptide 1) and amino acids 32-53 (synthetic peptide 2). Both sequences target epitopes that are hydrophilic and likely to be immunogenic (Table 4.1).

Both peptide sequences were analyzed using the BLAST tool to determine their specificity for the MRP4 protein. The results from this analysis show that synthetic peptide 1 highly matches the MRP4 protein in humans, chimpanzee, and rhesus monkey as defined by the parameters used in the BLAST search. Synthetic peptide 2 highly matches the MRP4 protein in humans, cows, chimpanzee, rhesus monkey and dogs. In addition, the MRP4 protein sequence was analyzed for post-translational modifications that could potentially mask the epitope site using the Emboss Lite program. A tyrosine phosphorylation site is predicted between amino acids 37 to 45, which are found in the epitope for synthetic peptide 2.

The antibody fraction purified from rabbit serum 316 has a concentration of $\mu\text{g}/\mu\text{l}$ and the antibody fraction purified from rabbit serum 319 has a concentration of $\mu\text{g}/\mu\text{l}$. The antibodies generated with synthetic peptide 1 and 2, were characterized in kidney tissue samples where the MRP4 protein is expected to be expressed. The MRP4 protein is expected in the kidney cortex and outer medulla, which contain the proximal tubule. The expected molecular weight with glycosylation is 170 kD, without glycosylation the

expected molecular weight is 146 kD. As shown in Figure 4.3, antibody 316 fractions 4 and 5 do not show the expected band; antibody 316 fraction 3 shows the expected band very faintly at 170 kD. Antibody 319 fraction 5 (319-5) shows a strong band at the expected molecular weight (Figure 4.4A).

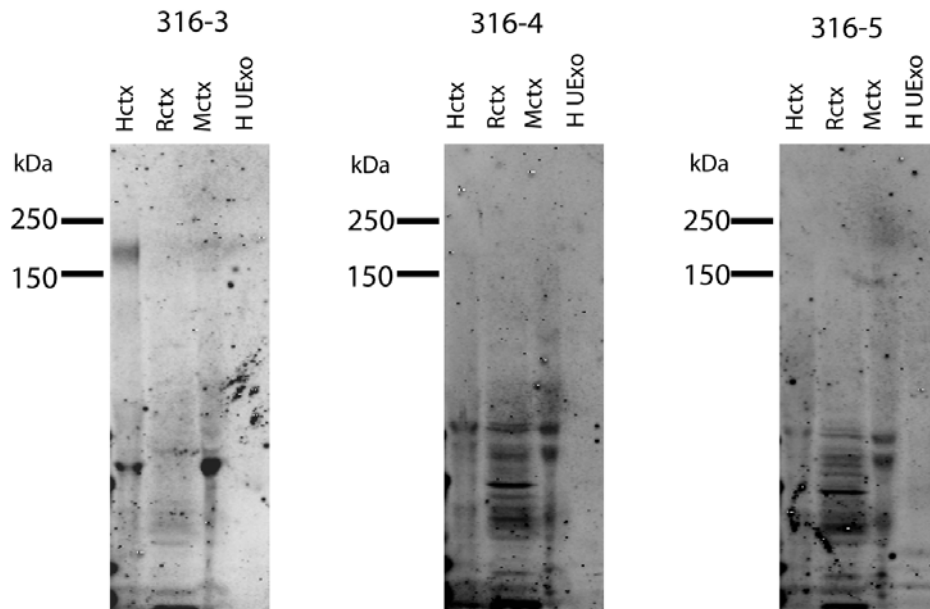
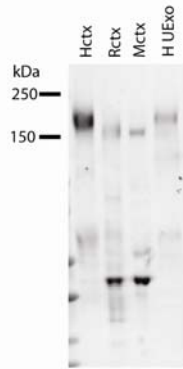


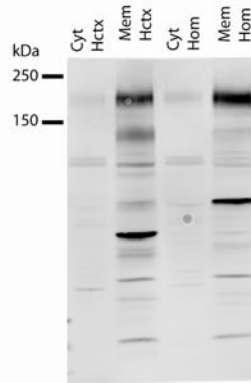
Figure 4.3 Characterization of rabbit polyclonal anti-MRP4 antibody (316) generated with synthetic peptide 1 (LPVYQEVKPNPLQDANLC). Three fractions (316-3, -4, and-5) were used to probe kidney tissue samples from human cortex (Hctx), rat cortex (Rctx), mouse cortex (Mctx) and human urinary exosomes (H UExo).

Antibody 319 fraction 5 was characterized in the membrane and cytosol fractions of human cortex and medulla. The membrane fractions of the tissue sample show strong bands at the expected molecular weight and the cytosol fractions do not (Figure 4.4B). The antibody 319-5 was pre-absorbed with synthetic peptide 2; the membrane and cytosol fractions were incubated with this pre-absorbed antibody and the expected bands were ablated (Figure 4.4C).

A. Characterization of rabbit polyclonal anti-MRP4 (319-4) in kidney tissue samples and human urinary exosome sample



B. Characterization of rabbit polyclonal anti-MRP4 (319-4) in membrane and cytosol kidney tissue samples



C. Peptide absorption of rabbit polyclonal anti-MRP4 (319-4)

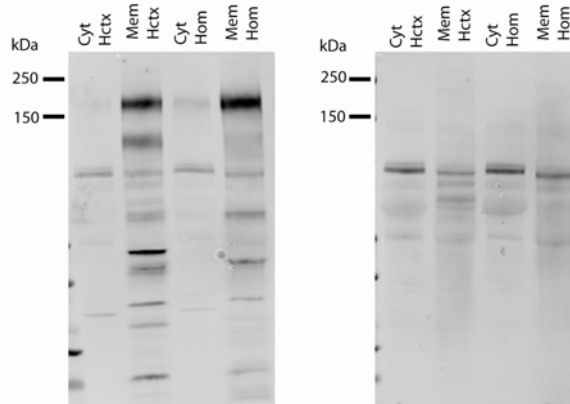


Figure 4.4 Characterization of rabbit polyclonal anti-MRP4 antibody (319) generated with synthetic peptide 2 (KIGHKRRLEEDDMYSVLPEDRS). (A) Fraction 4 was used to probe kidney tissue samples from human cortex (Hctx), rat cortex (Rctx), mouse cortex (Mctx) and human urinary exosomes (H UExo). Expected molecular weight is 170 kD. (B) Fraction 4 was used to probe kidney tissue membrane (Mem Hctx) and cytosolic (Cyt Hctx) samples. (C) Fraction 4 was incubated with synthetic peptide 2 and membrane and cytosolic kidney tissue samples were probed.

The samples from patient 4, patient 5, control 4 and control 5 were incubated with polyclonal antibody 319-5. Patient 4 and patient 5 have a mutation in the ABCC4 gene which causes the expression of a normal protein (1,325 amino acids) and a short-form protein (67 amino acids). The immunoblotting (Figure 4.5) shows the expected band for the MRP4 protein at 170 kD in the control samples and the patient samples. In addition, there are bands between 15 and 20 kD (with glycosylation) present in patient 4 and patient 5 but missing in control 4 and control 5.

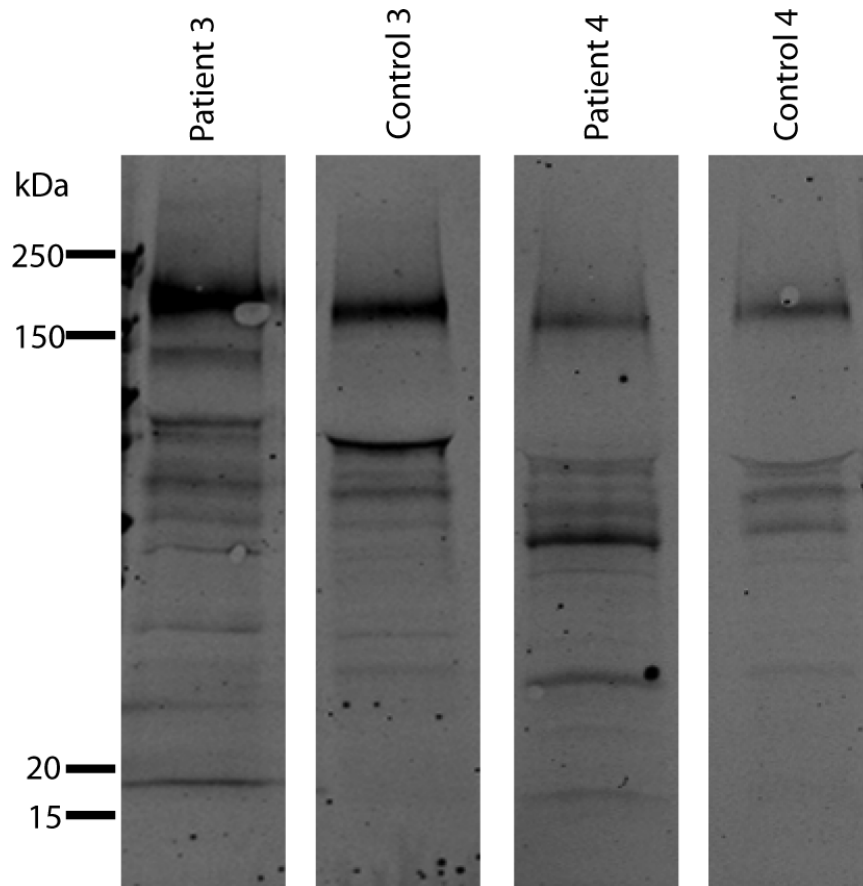


Figure 4.5 Genetic mutation of ABCC4 results in production of short-form and normal MRP4 protein. Patient samples (patient 3 and patient 4) and control samples (control 3 and control 4) were probed with rabbit polyclonal anti-MRP4 antibody (319 fraction 4). Expected molecular weight for the normal protein is 170 kD and for the short-form protein is between 15 and 20 kD.

4.4 Discussion

4.4.1 Bartter syndrome type I

As illustrated in Figure 4.1, analysis of human urinary exosomes by mass spectrometry (Chapter 3) and immunoblotting can provide information with regard to genetic diseases involving apical membrane proteins as shown by the qualitative assessment of urinary exosome samples from patients with Bartter syndrome type I. Mutations in the *SLC12A1* gene cause Bartter syndrome type I [49]. Many such mutations presumably result in misfolding of the NKCC2 protein, preventing the apical delivery of the protein and therefore preventing incorporation into urinary exosomes. The anti-NKCC2 and anti-NCC polyclonal antibodies generated against synthetic peptides were used in the immunoblotting technique to complete the qualitative assessment. The anti-NKCC2 polyclonal antibody was generated against a 23-amino acid synthetic peptide that matches the protein in the human proteome [53]. The anti-NCC polyclonal antibody was generated against a 22-amino acid synthetic peptide that matches the protein in the human proteome. Immunoblotting analysis of the urinary exosome samples from Bartter syndrome type I patients showed a complete absence of the NKCC2 protein. As a loading control, the patient samples showed the presence of the NCC protein, found in the distal convoluted tubule of the kidney and was identified in the proteome of the urinary exosomes.

4.4.2 Mutation in *ABCC4* gene

The urinary exosomes samples from two genotyped patients with a mutation of the *ABCC4* gene, producing a short and normal MRP4 protein, were analyzed by immunoblotting. The MRP4 protein is an organic anion found in the apical membrane of

the renal proximal tubule [57]. There are 98 single nucleotide polymorphisms (SNPs) in the ABCC4 gene and there is little evidence of a complete loss of function of the MRP4 protein [58]. The single nucleotide polymorphism in the intron region between exon 2 and 3, IVS2-5T>C, results in the production of a truncated protein, 67 amino acids in length, and a normal protein, 1,325 amino acids in length. This mutation of the ABCC4 gene could have an effect on the organic anion transport in and it might be involved in drug resistance. We wanted to conduct an initial qualitative analysis of urinary exosomes samples from patients with this mutation.

The immunoblotting analysis was done with a commercially-available monoclonal antibody against MRP4 (Figure 4.2). The commercial monoclonal antibody, anti-MRP4, recognizes an epitope in the carboxyl-terminal of the protein, amino acids 372-431. We generated a polyclonal antibody targeted to an epitope in the amino-terminal of the MRP4 protein. The polyclonal antibody generated was raised against a synthetic peptide that was designed to be immunogenic, unique and lacks-post translational modifications. The synthetic peptide was determined to be highly hydrophilic and as result expected to be immunogenic. The immunogenicity was determined using the hydropathy index using the Kyte-Doolittle method for calculating hydrophobicity and hydrophilicity of peptide regions in a protein. A hydrophobic region is more likely to span across the membrane and less likely to be immunogenic. A hydrophilic region is more likely to be exposed on the surface of the protein and more likely to be immunogenic. A hydrophobic amino acid has a positive index value and a hydrophilic amino acid has a negative index value.

The polyclonal antibody (319-5) was characterized in kidney tissue samples (Figure 4.4). The patient samples were analyzed by immunoblotting using this polyclonal antibody. The patient samples (Figure 4.5) showed a band at the expected molecular weight of the normal MRP4 protein, 170 kD and a fainter band at the expected molecular weight for the short-form protein, between 15 and 20 kD. The band for the short-form protein was missing from the control samples.

4.5 Conclusion

Urinary exosome analysis has the potential of usefulness in genetic mutation analysis. This potential application was demonstrated to be feasible using immunoassays such as immunoblotting with highly specific antibodies.

As examples of the usefulness of the urinary exosomes analysis in genetic mutation analysis, we qualitatively analyzed samples from patients clinically diagnosed with the well-defined renal tubulopathy, Bartter syndrome type I, and samples from patients with a single nucleotide polymorphism (SNP) in the ABCC4 gene. The anti-NKCC2 polyclonal antibody was used to probe the urinary exosomes patient samples (patient 1 and patient 2) and these showed an absence of the NKCC2 protein band. The anti-MRP4 polyclonal antibody was generated to probe urinary exosomes patient samples (patient 3 and patient 4) and showed a strong band for the normal MRP4 protein and faint band for the short-form MRP4 protein. Genetic mutation analysis of well-characterized patient samples is the first step in application of urinary exosomes analysis in large-scale biomarker discovery.

Table 4.1. Kyte-Doolittle Hydropathy Index for MRP4 protein

The hydropathy values are shown for amino acids 1 to 60 (out of 1,325)

Position	Residue	Result
1	M	0.000
2	L	0.000
3	P	0.000
4	V	0.000
5	Y	0.033
6	Q	-0.356
7	E	-1.167
8	V	-1.167
9	K	-1.211
10	P	-1.456
11	N	-1.456
12	P	-0.867
13	L	-1.722
14	Q	-0.867
15	D	-0.411
16	A	-0.111
17	N	-0.433
18	L	-0.389
19	C	0.311
20	S	1.011
21	R	0.711
22	V	1.000
23	F	1.000
24	F	0.333
25	W	0.244
26	W	1.167
27	L	1.011
28	N	0.267
29	P	0.456
30	L	0.511

Position	Residue	Result
31	F	0.256
32	K	-0.600
33	I	-0.711
34	G	-1.033
35	H	-1.033
36	K	-1.733
37	R	-1.689
38	R	-2.578
39	L	-2.922
40	E	-2.356
41	E	-2.067
42	D	-1.656
43	D	-0.689
44	M	-0.689
45	Y	-0.478
46	S	-0.478
47	V	-0.478
48	L	-0.589
49	P	-0.889
50	E	-1.133
51	D	-1.400
52	R	-1.444
53	S	-1.911
54	Q	-2.122
55	H	-2.122
56	L	-1.311
57	G	-1.200
58	E	-1.156
59	E	-0.456
60	L	-0.200

5 Normal variability and normalization of urinary exosomes: Gender Proteomics

5.1 Introduction

Urinary exosomes analysis provides a non-invasive window to cell processes in the kidney and a potential means of enrichment of disease biomarkers. Future biomarker studies using tandem mass spectrometry (LC-MS/MS) and immunoblotting require knowledge of normal variability among human populations and development of normalization techniques. The qualitative approach used to identify proteins in a sample as shown in Chapter 3 is not sufficient when developing strategies for biomarker studies. For this we propose to use a quantitative approach in addition to a qualitative approach. The quantification of expressed proteins using tandem mass spectrometry in a sample can be accomplished by a label-free approach [21, 60]. The quantification of protein-fold change using the immunoblotting technique can be accomplished by normalizing the loading of samples based on equal time.

The two types of approaches commonly used for quantification of protein expression using tandem mass spectrometry are label-based and label-free. Label-based approaches generally use chemical labels “to discriminate between a peptide in one sample and the same peptide in another sample” [21]. The difference in the masses is detected by the mass analyzer providing quantitative data. There are currently many methods that are considered label-base approaches such ICAT, IVICAT [61], SILAC and iTRAQ [21].

Label-free approaches use direct comparison of peptide peak intensities. An important factor in this approach is the mass accuracy of the analyzer in the first stage of mass spectrometry. The major advantages of this approach in future biomarker studies are that there is no added cost for use of labeling reagents and its application is not limited by the number of samples as some label-based approaches.

We proposed to study the normal variability between proteomes of female and male urinary exosomes. In the process, we developed valuable knowledge and tools to quantitatively analyze proteomic data. The quantification of the tandem mass spectrometry data is a relative quantification and was done using a label-free approach.

The relative quantification of protein abundance between these two samples is accomplished by plotting the intensity of the signal as the peptide is eluted from the chromatographic column [18]. The area under the curve of this plot is the extracted ion current (XIC) as defined by Ong and Mann. The XIC of the peptide is linearly related to the amount of the peptide and can be compared when the same peptide is identified in two different chromatographic runs. In our case the two chromatographic runs are the two samples of isolated urinary exosomes from males and females.

Our hypothesis for the quantitative study of tandem mass spectrometry data is that there should be a high concordance between proteomes of urinary exosomes isolated from males and females. We define high concordance as a change of less than 5% of the total number of proteins between the proteomes of males and females. In addition, any proteins that have at least a 2-fold increase or decrease will be further studied by the immunoblotting technique.

In collaboration with Dr. Jeffrey Kopp from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), we extended our study to analyze the proteins that have a 2-fold increase/decrease by the immunoblotting technique. The immunoblotting technique will be normalized by loading isolated urinary exosomes samples based on timed urine collection. In this study we plan to enroll three female and three male volunteers.

Implementing a normalization method to analyze isolated urinary exosomes samples using the immunoblotting assay is an important quantitative approach. In our research we propose to use timed urine collection. We hypothesize that time will linearly correlate with urinary exosomes excretion which will be quantified by signal intensity of multivesicular body (MVB) markers, TSG-101 and Alix [7].

5.2 Methods and Materials

5.2.1 Tandem mass spectrometry study

Urine was collected from eight healthy humans: Four men (aged 22 to 33) and four women (aged 24 to 35) (National Institute of Diabetes and Digestive and Kidney Diseases Clinical Research Protocol 00-DK-0107). Fifty milliliters per subject was collected. Protease inhibitors were added (1.67 ml of 100 mM NaN₃, 2.5 ml of 11.5 mM 4-[2-aminoethyl] benzenesulfonyl fluoride, and 50 ml of 1 mM leupeptin). The samples from men were mixed and the samples from women were mixed. The two mixed samples were centrifuged at 17,000 x g for 10 min at 4°C. The 17,000 x g supernatant was ultracentrifuged at 200,000 x g for 1 h at 25°C. The ultracentrifugation step was repeated 3 additional times, adding new 17,000 x g supernatant volume each time to each of the 12 tubes (6 tubes for male mixed sample, 6 tubes for female mixed sample). Each

of the 12 pellets was suspended with 50 μ l of “isolation solution” (10mM triethanolamine and 250 mM sucrose). The suspensions were pooled together.

To denature the zona pellucida domains in the Tamm-Horsfall protein, we mixed the re-suspended pellets with 200 mg/ml dithiothreitol (DTT) at 95°C for 2 min. The re-suspended pellets were added to ultracentrifuge tubes, and isolation solution was added to increase the volume to 8 ml. The samples were ultracentrifuged at 200,000 x g for 1h at 25°C. The pellets were suspended in 50 μ l of isolation solution and frozen at -80°C.

5.2.1.1 In-Gel Trypsin Digestion

The protein concentration was determined using the Bradford Assay. The male sample had a concentration of 0.76 mg/ml and the female sample had a concentration of 0.85 mg/ml. The male mixed sample and the female mixed sample were solubilized in Laemmli sample buffer (1.5% SDS, 6% glycerol/10 mM Tris HCl, and 60 mg/ml DTT). Proteins in the exosome samples were separated by 1D SDS-PAGE using a Bio-Rad Ready Gel 4 to 15% polyacrylamide gradient gel with 36 μ g in each lane. The gel was stained with Colloidal Coomassie Blue (GelCode Blue Stain Reagent; Pierce, Rockford IL) for 10 min and destained using ddH₂O (2 x 30 min). The gel was divided from top to bottom into 20 1-mm strips over the entire molecular weight range of the gel. Each strip was diced into small pieces (1 mm³) and placed into labeled centrifuge tubes.

The gels pieces were destained by adding 100 μ l of 25 mM ammonium bicarbonate (NH₄HCO₃)/50% acetonitrile (ACN) for 10 min and were dried using a SpeedVac. The samples were reduced in a solution of 10 mM DTT and 25 mM NH₄HCO₃ at 56°C for 1 h. The samples were alkylated in a solution containing 55 mM iodoacetamide and 25 mM NH₄HCO₃ in the dark at room temperature for 45 min. The

gel pieces were washed with 25 mM NH_4HCO_3 and dehydrated in a solution containing 25 mM NH_4HCO_3 and 50% ACN. The samples were dried using the SpeedVac. The samples were rehydrated in a solution containing 12.5 ng/ μl trypsin (V5113; Promega, Madison, WI) in 25 mM NH_4HCO_3 and digested overnight at 37°C. Peptides were extracted using 50% ACN/0.1% formic acid (FA). The extracted samples were dried using the SpeedVac to remove ACN and then reconstituted with 0.1% FA. All 40 peptide samples (20 peptide sample from male and 20 peptide samples from female) were desalted using C_{18} Zip Tips (Millipore, Billerica, MA) before analysis by mass spectrometry.

5.2.1.2 Nanospray LC-MS/MS

A high-accuracy Fourier transform mass spectrometer, FT-ICR (Thermo Electron Corp.) equipped with a nano-electrospray ion source was used to acquire m/z ratios in precursor ions (MS1). And high-sensitivity linear ion trap mass analyzer, LTQ (Thermo Electron Corp.) was used to acquire m/z ratios in fragmented ions (MS2) scans. To reduce further the sample complexity before mass analysis, we injected the tryptic peptides extracted from each gel slice using an Agilent 1100 nanoflow system (Agilent Technologies, Palo Alto, CA) into a reversed-phase liquid chromatographic column (PicoFrit, Biobasic C_{18} ; New Objective, Woodburn, MA). This LC-MS/MS method allows the acquisition of raw data files that are the MS/MS scans of the top three highest intensity peaks after fragmentation with collision-induced dissociation in the FT-ICR mass analyzer.

5.2.2 Analysis of Data

The raw data files were searched against the NCBI Reference Sequences (RefSeq)

human protein database by using the BIOWORKS software (Thermo Finnigan). BIOWORKS utilizes SEQUEST, and to identify all peptide sequences, we searched the raw data files using the target-decoy approach.

To apply the target-decoy database searching strategy, we used the NHLBI Proteomics Core Facility in-house software to create a composite database containing the forward and reverse sequences of the non-redundant NCBI Reference Sequences (RefSeq) human protein database released on January 26, 2007. We searched separately, the raw data files for the female and male samples against this composite database. After the search, we assessed the False-Discovery Rate (FDR) by the number of peptides matched from the reversed sequences.

The parameters that determine the stringency of the filtering criteria include XCorr, Sp rank, and delta Cn. These parameters were incrementally adjusted, thereby reducing the false-discovery identifications until a FDR target was achieved. In our case, the data were filtered to a target of 2% FDR, and the actual FDR was 1.59% for the female sample, and 1.71% for the male sample. The filter settings used for the male sample were max Xcorr rank 1, min Sp rank 1, min delta Cn 0.1, charge +1 min Xcorr 1.91, charge +2 min Xcorr 0, and charge +3 min Xcorr 0.

5.2.3 Label-free quantification

The quantification of the male and female samples was carried out using the custom computer program, QUOIL (Quantification without Isotope Labeling), developed by the NHLBI Proteomic Core Facility [60]. The program requires that one of the samples be chosen as the reference sample, this choice has no effect on the final outcome of the quantification. We chose the male sample as the reference. The raw data files

containing the chromatographic data were read by the QUOIL program and ion chromatograms were reconstructed. Intensities are counted when these meet pre-defined mass tolerance and threshold as well as signal-to-noise ratios. We pre-defined the precursor mass tolerance (m/z) to 0.05 daltons and the minimum signal-to-noise ratio to 1.5. These reconstructed ion chromatograms contain peak areas of peptides identified in both proteomes [60]. The peak area is calculated by integration using the trapezoidal approximation within a defined window around the peak center as shown in Figure 6.x. The defined window is determined by the retention time and the peak center is defined by the highest peak within that window [60].

The ratio of peak areas is calculated (female/male) for all peptides identified per protein and were manually inspected. The ratio of peak areas were log-transformed (log base 2) and the median of the log-transformed ratios was determined. The criteria to consider proteins as potential candidates of variability between these two proteomes were: (1) Proteins with two or more peptides with a peak area ratio less than 1 or greater than 1; (2) Reconstructed chromatograms of peak areas pass manual inspection; (3) The median of the log-transformed peak area ratios for the protein's peptides was 2 or greater.

5.2.3 Immunoblotting study

5.2.3.1 Urine collection

Six volunteers were enrolled in the study, three males and three females, for six days (National Institute of Diabetes and Digestive and Kidney Diseases Clinical Research Protocol). The first three days were the diet adaptation phase (days 1-3) [62] and the last three days were the urine collection phase (days 4-6). Urine is collected at room temperature every 12-hour during the last three days. Protease inhibitors were added to

the collection container (1.67 ml of 100 mM NaN_3 , 2.5 ml of 11.5 mM 4-[2-aminoethyl] benzenesulfonyl fluoride, and 50 ml of 1 mM leupeptin). During the six day admission, volunteers had a normal diet and a water intake appropriate for their body weight. Table 6.1 shows the volume of urine collected per 12-hour period per volunteer for the three days of urine collection.

5.2.3.2 Urinary exosome isolation

The urinary exosomes were isolated from all six 12-hour samples collected from each volunteer using the procedure described previously (Chapter 3). Briefly, a fraction of the total volume of 12-hour urine sample is centrifuged at $17,000 \times g$ for 10 min at 37°C . The $17,000 \times g$ pellet is re-suspended in 1mL of isolation solution, 200 mg of DTT is added and the sample is heated to 37°C for 5 minutes. The pellet is vortexed for at least 5 min to dissolve cross-linked Tamm-Horsfall protein. Isolation solution is added to make a total volume of 8 mL and the sample is centrifuged at $17,000 \times g$ for 10 minutes at 37°C . The supernatant is added to the previous $17,000 \times g$ supernatant and ultracentrifuged at $200,000 \times g$ for 1 h at 37°C . Pellets are re-suspended in 0.1% SDS.

5.2.3.3 Immunoblotting

Protein concentrations for all samples were determined using the Bradford Assay. The samples were solubilized in Laemmli sample buffer (1.5% SDS, 6% glycerol/10 mM Tris HCl). The samples were loaded onto a 1-D SDS-PAGE gel on the basis of timed urine collection, the time for loading is 0.07 hours. The samples were loaded as follows: Gel 1 has samples from male volunteer M2 and female volunteer F1; Gel 2 has samples from male volunteer M3 and female volunteer F2; Gel 3 has samples from male volunteer M4 and female volunteer F3. The proteins were transferred to nitrocellulose (Millipore)

membranes, blocked, and probed with antigen-specific TSG-101 and Alix primary antibodies. We incubated the blots with species-specific fluorescence secondary antibodies (Alexa 688) and visualized them using the Odyssey Infrared Imaging System (LiCor, Lincoln, NE).

5.3 Results

In this study, we used tandem mass spectrometry to determine the normal variability among male and female proteomes isolated from urinary exosomes. We used a high accuracy FT-ICR mass spectrometer to analyze these samples. The protein expression in these samples was quantified using a label-free approach with the NHLBI Proteomic Core program, QUOIL.

A total of 388 proteins were identified from both proteomes in male and female samples after target/decoy analysis predicting a false positive identification rate of less than 2% for both samples as shown in Table 5.1. Out of the total 388 proteins identified from both proteomes, only 27 proteins were not previously identified in the large-scale proteomic analysis of urinary exosomes (Chapter 3). And among all proteins, only 19 were identified that differ by at least 2-fold as shown in Table 5.2.

The immunoblots in Figure 5.1 were incubated with Alix and TSG-101 which are markers for multivesicular bodies (MVB). The Alix protein has a 100 kDa molecular weight and it is the upper band on the immunoblots. The TSG-101 protein has a 47 kDa molecular weight and it is the lower band on the immunoblots. As shown in Figure 5.1A all six samples from male volunteer M-2 have bands for Alix and TSG-101. Only three samples from female volunteer F-1 have bands for Alix and TSG-101.

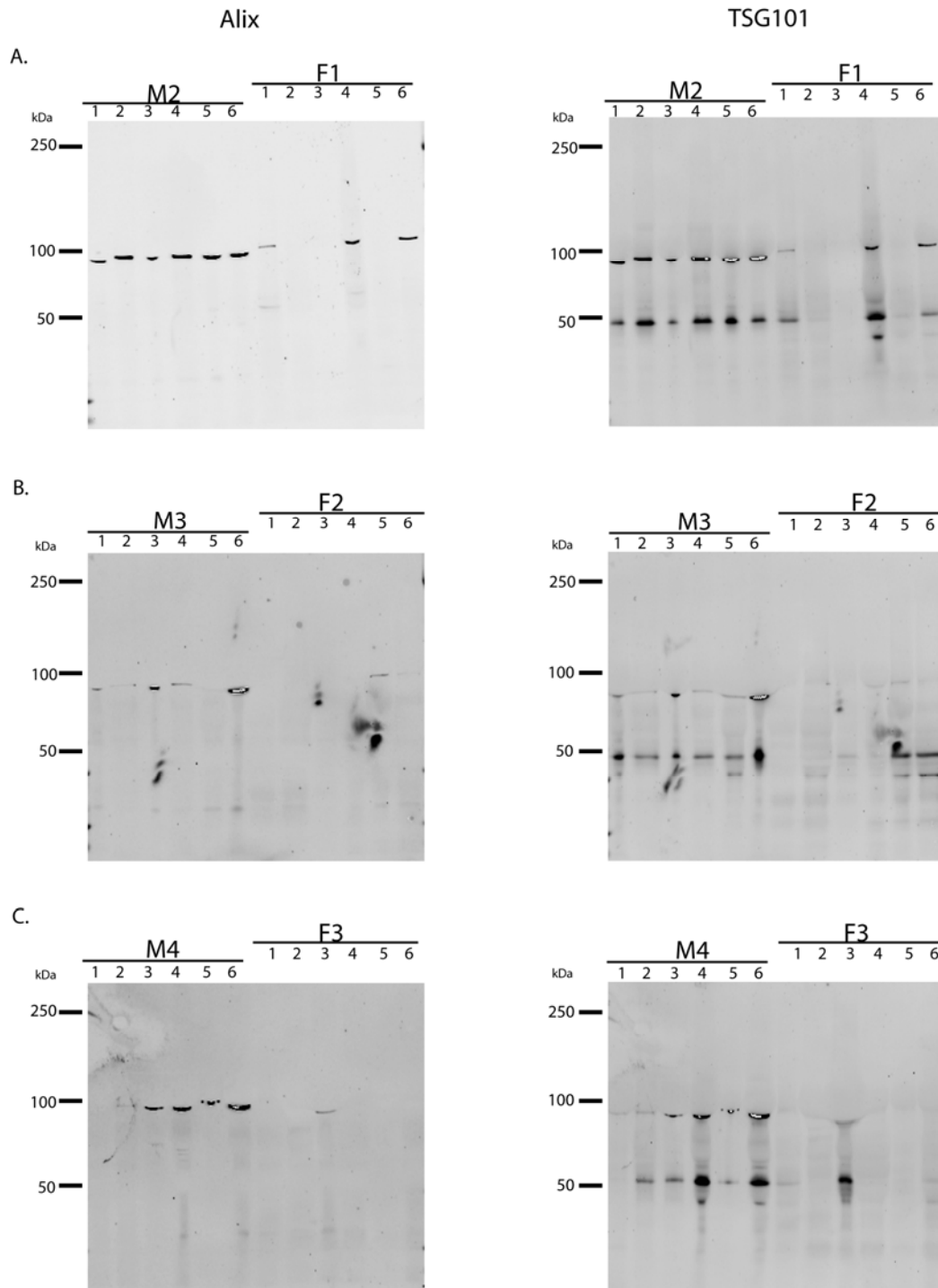


Figure 5.1 Immunoblots of urinary exosomes samples from three male healthy volunteers (M2, M3, M4) and three female healthy volunteers (F1, F2, F3). Six samples were collected per volunteer corresponding to 12-hour urine collections. Samples were loaded based on time, $t=0.07$ hours. Anti-TSG101 and anti-Alix antibodies were used to probe samples. Expected molecular weight of TSG101 is 100 kD and Alix is 47 kD.

Figure 5.1B shows that all six samples from male volunteer M-3 have bands for Alix and TSG-101. The last two samples for female volunteer F-2 have bands for TSG-101. Figure 5.1C shows that only five samples from male volunteer M-4 show bands for TSG-101 and Alix. And only one sample from female volunteer F-3 shows bands for TSG-101 and Alix.

5.4 Discussion

One of the objectives of this study was to acquire knowledge about the normal variability among human populations for future biomarker studies using urinary exosomes isolation. We compared exosomes isolated from pooled samples of urine from four normal human females and four human males. The samples were solubilized in SDS and proteins were separated by SDS-PAGE followed by in-gel trypsin digestion of contiguous gel slices. The digests were analyzed by tandem mass spectrometry using the FT-ICR mass spectrometer (Thermo-Finnegan LTQ) and we compared the two samples quantitatively using a label-free approach with the integration software, QUOIL, developed by the NHLBI Proteomic Core. The relative quantification of protein abundance using the QUOIL program was accomplished by plotting the intensity of the identified peptide signal as the peptide is eluted from the chromatographic column. These reconstructed ion chromatograms contain peak areas of peptides identified in both proteomes. The peak areas were calculated by integration and the ratio of peak areas was calculated. This ratio is log-transformed and the median is used to determine the fold-change.

There were 388 proteins identified in both male and female proteomes from isolated urinary exosomes. Only 19 proteins had a 2-fold change and overall represent

4.9% of the total number of proteins identified in both proteomes. This result shows that there is high concordance between proteomes of urinary exosomes isolated from males and females.

The second objective of the proposed research was to extend our study to analyze the proteins that have a 2-fold change by the immunoblotting technique. In order to quantify protein expression with the immunoblotting technique we proposed to normalize samples by loading based on timed urine collection. We enrolled six healthy volunteers, three male and three female for a six-day study. We collected 12-hour urine samples for the last three days of the study. The urinary exosomes samples were isolated by centrifugation and loaded onto SDS-PAGE gels. The proteins were transferred to nitrocellulose membranes and probed with antibodies against MVB markers. We expected a linear correlation of excreted urinary exosomes with timed urine collection, as shown by the signal intensity of MVB markers, TSG-101 and ALIX. The results of the normalization did not show a linear correlation of timed urine collection and signal intensity of MVB markers.

5.5 Conclusion

A total of 388 proteins were identified in the two urine samples after the application of high stringency acceptance criteria and target/decoy analysis predicting a false positive identification rate of 1%. Among all proteins, only 19 proteins were identified that differ by at least 2-fold between female and male urinary exosomes. These include 15 that were lower in female than male samples and 4 that were higher in female than male samples. Therefore, we conclude that there is high concordance between proteomes of urinary exosomes proteins derived from males and females.

Finally, the analysis of proteins that differ by at least a 2-fold between female and male urinary exosomes depended on implementing a normalization method. The proposed normalization method was to load samples based on timed urine collections. The samples were probed with antibodies against MVB markers, TSG101 and Alix. The results did not show correlation between signal intensity of TSG101 and Alix and timed urine collection. A possible explanation is that the excretion of Tamm-Horsfall protein, prevent the efficient isolation of urinary exosomes. Tamm-Horsfall tends to form vast aggregates in urine that can entrain the exosomes.

Table 5.1 Proteins identified in male and female urinary exosomes proteomes

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_000026	aldolase B, fructose-bisphosphate	ALDOB	10	55	11	67
NP_000032	apolipoprotein E precursor	APOE	7	23	5	5
NP_000038	arylsulfatase E precursor	ARSE	1	1	1	1
NP_000041	argininosuccinate synthetase 1	ASS1	5	13	4	8
NP_000058	carbonic anhydrase II	CA2	6	16	6	12
NP_000090	cystatin C precursor	CST3	1	4	1	1
NP_000261	purine nucleoside phosphorylase	NP	3	4	3	5
NP_000276	Xaa-Pro dipeptidase	PEPD	2	6	1	1
NP_000282	phosphoglycerate kinase 1	PGK1	8	19	5	11
NP_000299	cathepsin A precursor	CTSA	1	2	1	3
NP_000329	sodium potassium chloride cotransporter 2	SLC12A1	18	99	13	51
NP_000330	solute carrier family 12 (sodium/chloride transporters), member 3	SLC12A3	15	55	5	13
NP_000332	solute carrier family 3, member 1	SLC3A1	6	15	6	13
NP_000334	solute carrier family 5 (sodium/glucose cotransporter), member 1	SLC5A1	2	3	2	4
NP_000356	triosephosphate isomerase 1	TPI1	2	4	2	6
NP_000382	tripeptidyl-peptidase I preproprotein	TPP1	2	8	2	6
NP_000395	galactosidase, beta 1	GLB1	10	36	10	52
NP_000434	ATP-binding cassette, subfamily B, member 4 isoform A	ABCB4	2	3	3	4
NP_000468	albumin preproprotein	ALB	16	42	9	15
NP_000469	tissue non-specific alkaline phosphatase precursor	ALPL	3	4	1	4
NP_000477	aquaporin 2	AQP2	2	18	2	12
NP_000498	fructose-1,6-bisphosphatase 1	FBP1	3	4	3	6
NP_000509	beta globin	HBB	1	2	1	1
NP_000552	glutathione S-transferase mu 1 isoform 1	GSTM1	3	6	1	2
NP_000615	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5	SERPINA5	9	28	6	36
NP_000628	glutathione reductase	GSR	1	2	1	1
NP_000657	aminoacylase 1	ACY1	11	19	8	28
NP_000678	S-adenosylhomocysteine hydrolase	AHCY	2	4	4	6
NP_000691	annexin I	ANXA1	4	9	10	52
NP_000692	Na ⁺ /K ⁺ -ATPase alpha 1 subunit isoform a proprotein	ATP1A1	21	70	12	26
NP_000704	biliverdin reductase B (flavin reductase (NADPH))	BLVRB	2	4	1	1
NP_000780	angiotensin I converting enzyme isoform 1 precursor	ACE	15	51	13	60
NP_000837	glutathione S-transferase A2	GSTA2	3	4	6	14
NP_000840	glutathione S-transferase mu 3	GSTM3	3	4	3	6
NP_000843	glutathione transferase	GSTP1	4	12	6	14
NP_000851	hydroxyprostaglandin dehydrogenase 15-(NAD)	HPGD	4	8	5	8
NP_000884	kininogen 1 isoform 2	KNG1	6	13	6	22
NP_000918	ATP-binding cassette, subfamily B, member 1	ABCB1	9	37	10	33
NP_001002857	annexin A2 isoform 2	ANXA2	11	31	12	44
NP_001003841	solute carrier family 6, member 19	SLC6A19	1	2	1	8
NP_036344	NADP-dependent leukotriene B4 12-hydroxydehydrogenase	PTGR1	2	6	4	8
NP_001005502	carboxypeptidase M precursor	CPM	2	5	1	1
NP_001007070	syntenin isoform 3	SDCBP	3	14	3	6
NP_036361	5'-nucleotidase, cytosolic II	NT5C2	1	2	1	1
NP_001008741	peptidylprolyl isomerase A-like	LOC388817	1	2	1	1
NP_001009881	zinc finger, CCHC domain containing 11 isoform a	ZCCHC11	1	25	1	26
NP_001012679	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a	SLC3A2	5	20	1	7
NP_001014390	heat shock protein HSP 90-beta	HSP90AB2P	4	6	2	4
NP_001014436	drebrin-like isoform b	DBNL	1	7	1	3
NP_001017421	actin-like protein	FKSG30	2	137	2	126
NP_001018017	mucin 1 isoform 3 precursor	MUC1	4	30	4	19
NP_001019770	histone cluster 2, H2bf	HIST2H2BF	2	4	3	8
NP_003370	ezrin	EZR	19	127	10	86
NP_001026888	glycerophosphodiester phosphodiesterase domain containing 3 isoform 1	GDPD3	1	1	1	3
NP_001027467	serine/threonine kinase 24 (STE20 homolog, yeast) isoform b	STK24	2	6	2	3

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_001027536	gamma-glutamyltransferase 1 precursor	GGT1	3	5	1	2
NP_001055	transketolase isoform 1	TKT	1	1	3	4
NP_001060	tubulin, beta 2	TUBB2A	4	7	3	9
NP_060712	solute carrier family 47, member 1	SLC47A1	2	3	3	6
NP_001072	cubilin	CUBN	52	238	40	318
NP_074035	smooth muscle myosin heavy chain 11 isoform SM2A	MYH11	2	8	1	9
NP_001141	membrane alanine aminopeptidase precursor	ANPEP	38	514	31	492
NP_001144	annexin IV	ANXA4	12	22	18	40
NP_001145	annexin 5	ANXA5	7	15	3	5
NP_001146	annexin VI isoform 1	ANXA6	14	30	7	18
NP_001166	Rho GDP dissociation inhibitor (GDI) beta	ARHGDI3	2	3	2	6
NP_003050	solute carrier family 22 member 4	SLC22A4	1	2	1	2
NP_001227	carbonyl reductase 3	CBR3	1	2	1	3
NP_001279	chloride intracellular channel 1	CLIC1	2	4	1	3
NP_001419	enolase 1	ENO1	17	43	8	48
NP_001485	GDP dissociation inhibitor 2 isoform 1	GDI2	7	17	5	10
NP_001531	heat shock protein beta-1	HSPB1	3	6	6	11
NP_001601	acid phosphatase 2, lysosomal isoform 1 precursor	ACP2	4	14	2	7
NP_001604	alpha 2 actin	ACTA2	11	133	9	84
NP_001605	actin, gamma 1 propeptide	ACTG1	7	117	5	120
NP_001619	aldose reductase	AKR1B1	2	11	2	4
NP_001630	serum amyloid P component precursor	APCS	4	8	3	6
NP_001638	apolipoprotein D precursor	APOD	2	11	8	27
NP_001650	ADP-ribosylation factor 3	ARF3	1	2	1	5
NP_001653	ADP-ribosylation factor 5	ARF5	3	6	3	11
NP_001681	ATPase, H+ transporting, lysosomal V1 subunit A	ATP6V1A	12	25	7	13
NP_001683	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B1	ATP6V1B1	4	9	2	3
NP_001684	vacuolar H+ATPase B2	ATP6V1B2	7	22	4	9
NP_001687	vacuolar H+ ATPase E1 isoform a	ATP6V1E1	1	2	2	5
NP_001704	betaine-homocysteine methyltransferase	BHMT	4	8	2	4
NP_001738	gelsolin-like capping protein	CAPG	7	31	5	15
NP_001748	carbonyl reductase 1	CBR1	3	5	3	8
NP_001760	CD9 antigen	CD9	3	19	3	9
NP_001771	CD63 antigen isoform A	CD63	3	28	2	30
NP_001782	cell division cycle 42 isoform 1	CDC42	1	2	1	2
NP_001814	brain creatine kinase	CKB	6	21	2	3
NP_001879	crystallin, mu isoform 1	CRYM	5	9	4	6
NP_001880	crystallin, zeta isoform a	CRYZ	3	3	3	3
NP_071396	kinesin family member 13A	KIF13A	1	3	1	2
NP_001926	dipeptidylpeptidase IV	DPP4	22	104	20	95
NP_001954	epidermal growth factor (beta-urogastrone)	EGF	13	65	11	92
NP_001958	eukaryotic translation initiation factor 4A2	EIF4A2	1	2	1	2
NP_001968	glutamyl aminopeptidase (aminopeptidase A)	ENPEP	10	20	10	22
NP_002022	fyn-related kinase	FRK	4	6	1	2
NP_002023	ferritin, heavy polypeptide 1	FTH1	1	3	1	1
NP_002037	glyceraldehyde-3-phosphate dehydrogenase	GAPDH	5	19	4	10
NP_002049	gamma-glutamyltransferase 2	GGT2	8	53	9	109
NP_002058	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	GNA11	4	9	2	4
NP_002060	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	GNAI1	2	5	1	2
NP_002061	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2	GNAI2	5	16	3	8
NP_002063	guanine nucleotide binding protein (G protein), q polypeptide	GNAQ	3	11	3	6
NP_002070	aspartate aminotransferase 1	GOT1	6	9	1	2
NP_002077	growth factor receptor-bound protein 2 isoform 1	GRB2	1	1	1	1
NP_002146	heat shock 70kDa protein 6 (HSP70B')	HSPA6	2	13	2	6
NP_002188	aconitase 1	ACO1	4	11	3	8
NP_002291	L-lactate dehydrogenase B	LDHB	8	33	6	34
NP_002347	myristoylated alanine-rich protein kinase C substrate	MARCKS	3	8	2	3
NP_002435	moesin	MSN	13	27	11	24

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_002515	neuroblastoma RAS viral (v-ras) oncogene homolog	NRAS	1	4	1	1
NP_002551	purinergic receptor P2X4	P2RX4	1	6	1	3
NP_002558	prostatic binding protein	PEBP1	2	8	2	7
NP_002582	cytosolic phosphoenolpyruvate carboxykinase 1	PCK1	2	3	5	8
NP_002605	PDZ domain containing 1	PDZK1	10	21	5	13
NP_002624	phosphoglucomutase 1	PGM1	6	10	4	6
NP_002635	polymeric immunoglobulin receptor	PIGR	6	11	10	20
NP_002645	pyruvate kinase, muscle isoform M2	PKM2	1	3	1	5
NP_002700	protein phosphatase 1, catalytic subunit, beta isoform 1	PPP1CB	1	2	1	2
NP_002785	proteasome beta 2 subunit	PSMB2	1	1	1	1
NP_002796	proteasome 26S ATPase subunit 5	PSMC5	1	2	1	5
NP_002859	RAB5B, member RAS oncogene family	RAB5B	2	3	1	1
NP_002872	v-ral simian leukemia viral oncogene homolog B	RALB	3	7	5	8
NP_002875	RAP1A, member of RAS oncogene family	RAP1A	2	10	1	6
NP_002897	radixin	RDX	7	11	2	2
NP_003030	solute carrier family 2 (facilitated glucose/fructose transporter), member 5 isoform 1	SLC2A5	4	16	5	10
NP_003032	solute carrier family 5 (sodium/glucose cotransporter), member 2	SLC5A2	4	8	6	11
NP_001883	casein kinase 1, alpha 1 isoform 2	CSNK1A1	1	1	1	2
NP_003051	solute carrier family 22 member 5	SLC22A5	1	2	1	3
NP_003078	synuclein, gamma (breast cancer-specific protein 1)	SNCG	2	7	2	6
NP_003095	sorbitol dehydrogenase	SORD	2	11	2	4
NP_003121	sorcin isoform a	SRI	2	3	1	2
NP_003304	tissue specific transplantation antigen P35B	TSTA3	1	1	1	1
NP_003320	thioredoxin	TXN	1	6	1	4
NP_003325	ubiquitin-activating enzyme E1	UBA1	3	5	1	1
NP_003339	ubiquitin-conjugating enzyme E2N	UBE2N	1	6	1	2
NP_003352	uromodulin precursor	UMOD	19	1720	19	1921
NP_003364	vinculin isoform VCL	VCL	2	3	2	2
NP_003365	voltage-dependent anion channel 1	VDAC1	2	10	3	11
NP_001003962	calpain, small subunit 1	CAPNS1	2	4	1	4
NP_003390	X-prolyl aminopeptidase 2, membrane-bound	XPNPEP2	21	61	16	53
NP_003397	tyrosine 3/tryptophan 5 -monooxygenase activation protein, zeta polypeptide	YWHAZ	4	14	5	16
NP_003536	histone cluster 1, H4e	HIST1H4E	2	3	4	10
NP_001020527	anoctamin 6	ANO6	1	2	1	5
NP_003680	aldo-keto reductase family 7, member A2	AKR7A2	2	3	2	3
NP_003730	aldo-keto reductase family 1, member C3	AKR1C3	2	4	2	7
NP_003733	ATP-binding cassette, sub-family B (MDR/TAP), member 11	ABCB11	1	5	1	2
NP_003750	solute carrier family 4, sodium bicarbonate cotransporter, member 4 isoform 2	SLC4A4	4	20	1	4
NP_003818	N-ethylmaleimide-sensitive factor attachment protein, alpha	NAPA	1	2	1	1
NP_003861	IQ motif containing GTPase activating protein 1	IQGAP1	7	12	8	9
NP_003881	Fc fragment of IgG binding protein	FCGBP	5	8	8	11
NP_003900	copine III	CPNE3	8	21	7	24
NP_003920	RAB7, member RAS oncogene family-like 1 isoform 1	RAB7L1	2	4	1	4
NP_003962	sperm associated antigen 9 isoform 3	SPAG9	1	1	2	2
NP_003970	G protein-coupled receptor, family C, group 5, member A	GPRC5A	1	11	1	6
NP_003975	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	SLC13A2	2	3	4	10
NP_003977	gamma-butyrobetaine dioxygenase	BBOX1	1	1	4	8
NP_004024	annexin VI isoform 2	ANXA6	1	1	1	1
NP_004025	annexin VII isoform 2	ANXA7	10	28	6	13
NP_004033	arylsulfatase F precursor	ARSF	5	10	6	21
NP_004046	calpain 5	CAPN5	4	11	1	2
NP_004090	stomatin isoform a	STOM	11	46	8	54
NP_004095	fatty acid synthase	FASN	3	3	1	1
NP_004152	RAB1A, member RAS oncogene family isoform 1	RAB1A	1	3	2	7
NP_004153	RAB5A, member RAS oncogene family	RAB5A	5	8	4	7
NP_004172	ubiquitin carboxyl-terminal esterase L1	UCHL1	1	1	1	1
NP_004209	RAB11B, member RAS oncogene family	RAB11B	5	8	6	12

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_004247	organic cation transporter like 3	SLC22A13	3	5	1	1
NP_004288	G alpha 14	GNA14	2	4	2	7
NP_004300	Rho GDP dissociation inhibitor (GDI) alpha	ARHGDI A	2	7	2	4
NP_004326	bone marrow stromal cell antigen 2	BST2	1	2	1	1
NP_004404	dipeptidase 1	DPEP1	18	112	12	98
NP_004438	epidermal growth factor receptor pathway substrate 8	EPS8	14	24	3	5
NP_004512	kinesin family member 5B	KIF5B	1	2	1	2
NP_004628	RAB7, member RAS oncogene family	RAB7A	4	10	3	10
NP_005991	tubulin, alpha 4a	TUBA4A	3	6	1	2
NP_004659	maltase-glucoamylase	MGAM	20	62	13	29
NP_004682	ATPase, H ⁺ transporting, lysosomal, V0 subunit d1	ATP6V0D1	3	7	1	2
NP_004842	napsin A preproprotein	NAPSA	3	47	3	52
NP_004860	vacuolar protein sorting factor 4B	VPS4B	13	40	5	17
NP_004896	peroxiredoxin 6	PRDX6	5	6	5	7
NP_004915	actinin, alpha 4	ACTN4	9	14	7	10
NP_004920	calbindin 1	CALB1	4	8	3	5
NP_004921	F-actin capping protein beta subunit	CAPZB	3	5	3	4
NP_004990	myosin VI	MYO6	3	11	2	5
NP_005013	profilin 1	PFN1	4	5	3	4
NP_005023	plastin 3	PLS3	4	4	1	1
NP_005100	oxidative-stress responsive 1	OXSR1	2	2	1	1
NP_005106	major vault protein	MVP	2	3	3	7
NP_005130	annexin A3	ANXA3	2	4	6	13
NP_005168	ATPase, H ⁺ transporting, lysosomal V0 subunit a1 isoform c	ATP6V0A1	1	1	1	1
NP_005263	guanine nucleotide binding protein, alpha transducing activity polypeptide 2	GNAT2	1	5	1	3
NP_005264	guanine nucleotide-binding protein, beta-2 subunit	GNB2	5	32	4	16
NP_005265	guanine nucleotide binding protein (G protein), gamma 5	GNG5	1	2	1	1
NP_005336	heat shock 70kDa protein 1A	HSPA1A	10	40	9	33
NP_005361	mel transforming oncogene	RAB8A	4	20	4	16
NP_005388	podocalyxin-like isoform 2 precursor	PODXL	6	50	3	41
NP_005393	ras related v-ral simian leukemia viral oncogene homolog A	RALA	1	2	2	3
NP_005424	viral oncogene yes-1 homolog 1	YES1	4	7	2	3
NP_005498	cofilin 1 (non-muscle)	CFL1	2	3	1	2
NP_005552	lysosomal-associated membrane protein 1	LAMP1	3	12	2	13
NP_005557	L-lactate dehydrogenase A isoform 1	LDHA	4	14	4	12
NP_005558	galectin 3 binding protein	LGALS3BP	14	95	12	91
NP_005663	prostate stem cell antigen preproprotein	PSCA	1	5	1	9
NP_005718	tetraspan 1	TSPAN1	1	5	1	3
NP_005722	actin related protein 2/3 complex subunit 2	ARPC2	3	3	1	2
NP_005755	PDZK1 interacting protein 1	PDZK1IP1	2	18	2	4
NP_005887	isocitrate dehydrogenase 1 (NADP+), soluble	IDH1	9	21	5	7
NP_005908	cytosolic malate dehydrogenase	MDH1	6	12	5	14
NP_006000	tubulin, alpha 1a	TUBA1A	8	25	5	16
NP_004554	mitochondrial phosphoenolpyruvate carboxykinase 2 isoform 1 precursor	PCK2	1	1	1	1
NP_006008	prominin 1	PROM1	15	90	8	30
NP_006026	CDC42-binding protein kinase beta	CDC42BPB	1	1	1	1
NP_006057	aldo-keto reductase family 1, member A1	AKR1A1	6	11	6	15
NP_006077	tubulin, beta, 4	TUBB3	4	14	2	6
NP_006079	tubulin, beta, 2	TUBB2C	1	4	1	5
NP_006127	capping protein (actin filament) muscle Z-line, alpha 2	CAPZA2	2	2	1	1
NP_277050	Rho GTPase activating protein 18	ARHGAP18	1	3	1	1
NP_006216	phospholipase C, delta 1	PLCD1	6	13	2	3
NP_006261	related RAS viral (r-ras) oncogene homolog	RRAS	1	3	2	3
NP_006279	Thy-1 cell surface antigen preproprotein	THY1	2	10	2	11
NP_006283	tumor susceptibility gene 101	TSG101	13	22	5	10
NP_006308	brain abundant, membrane attached signal protein 1	BASP1	6	19	6	10
NP_006364	vesicle amine transport protein 1	VAT1	6	19	6	10
NP_006409	olfactomedin 4 precursor	OLFM4	10	41	5	16
NP_006479	ketohehexokinase isoform b	KHK	1	2	1	3
NP_006487	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	GNAI3	3	8	2	6

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_006563	guanine nucleotide binding protein (G protein), alpha 13	GNA13	7	10	2	2
NP_006588	heat shock 70kDa protein 8 isoform 1	HSPA8	10	43	6	27
NP_006601	mannan-binding lectin serine protease 2 isoform 1 precursor	MASP2	4	142	3	33
NP_775931	downstream of tyrosine kinase 7	DOK7	1	20	1	24
NP_006624	IQ motif containing GTPase activating protein 2	IQGAP2	2	3	2	2
NP_006751	uroplakin 2	UPK2	3	8	3	5
NP_006752	tyrosine 3/tryptophan 5 -monooxygenase activation protein, epsilon polypeptide	YWHAE	8	16	6	23
NP_006786	EH-domain containing 1	EHD1	4	9	4	5
NP_006817	tyrosine 3/tryptophan 5 -monooxygenase activation protein, theta polypeptide	YWHAQ	2	6	2	3
NP_006852	RAB35, member RAS oncogene family	RAB35	3	7	3	6
NP_008880	syntaxin binding protein 2	STXBP2	5	9	9	13
NP_008883	uroplakin 1B	UPK1B	1	2	2	6
NP_008931	uroplakin 1A	UPK1A	4	11	5	19
NP_009057	valosin-containing protein	VCP	2	5	1	3
NP_009111	trehalase (brush-border membrane glycoprotein)	TREH	2	3	2	3
NP_009148	programmed cell death 10	PDCD10	1	3	1	2
NP_009160	protein kinase C and casein kinase substrate in neurons 2	PACSIN2	3	4	1	2
NP_009172	EAP30 subunit of ELL complex	SNF8	1	2	1	2
NP_009219	membrane metallo-endopeptidase	MME	31	199	27	189
NP_036205	chaperonin containing TCPI1, subunit 5 (epsilon)	CCT5	1	1	1	3
NP_036226	peroxiredoxin 5 isoform a precursor	PRDX5	1	4	1	2
NP_036269	dimethylarginine dimethylaminohydrolase 1 isoform 1	DDAH1	2	5	2	4
NP_036322	aldehyde dehydrogenase 1 family, member L1	ALDH1L1	5	7	4	5
NP_036376	solute carrier family 7 (cationic amino acid transporter, y+ system), member 8 isoform a	SLC7A8	2	7	1	1
NP_060784	ubiquitin-conjugating enzyme E2-like isoform b	UEVLD	9	25	5	6
NP_065806	serine incorporator 1	SERINC1	1	3	1	1
NP_036382	related RAS viral (r-ras) oncogene homolog 2 isoform a	RRAS2	1	3	1	3
NP_036524	penta-EF-hand domain containing 1	PEF1	2	5	2	4
NP_036611	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	YWHAG	2	6	3	8
NP_037364	programmed cell death 6	PDCD6	4	5	2	2
NP_037377	vacuolar protein sorting factor 4A	VPS4A	9	20	2	2
NP_037506	programmed cell death 6 interacting protein	PDCD6IP	24	81	17	56
NP_038464	NCK-associated protein 1 isoform 1	NCKAP1	2	4	1	1
NP_039234	chloride intracellular channel 4	CLIC4	3	5	3	7
NP_054701	lysosomal-associated membrane protein 2 isoform B precursor	LAMP2	2	12	1	8
NP_054762	chromatin modifying protein 2B	CHMP2B	3	5	2	6
NP_055111	calpain 7	CAPN7	19	74	15	44
NP_060691	Mob4B protein	MOBK1B	1	1	1	1
NP_055135	heme binding protein 2	HEBP2	1	1	1	1
NP_055179	glyceraldehyde-3-phosphate dehydrogenase, spermatogenic	GAPDHS	2	5	1	4
NP_055268	chromatin modifying protein 2A	CHMP2A	3	16	2	7
NP_055289	acid sphingomyelinase-like phosphodiesterase 3B isoform 1	SMPDL3B	1	4	1	2
NP_055415	EH-domain containing 3	EHD3	5	10	3	6
NP_055440	podocin	NPHS2	5	15	1	2
NP_055535	serine/threonine kinase 2	SLK	5	8	3	3
NP_055576	putative MAPK activating protein PM28	KIAA0174	13	75	11	49
NP_055814	RAB21, member RAS oncogene family	RAB21	1	2	2	4
NP_055956	glycerol-3-phosphate dehydrogenase 1-like	GPD1L	1	4	1	2
NP_056348	dihydroxyacetone kinase 2	DAK	9	27	9	45
NP_056991	leucine aminopeptidase 3	LAP3	2	3	1	1
NP_057058	lambda-crystallin	CRYL1	2	3	2	4
NP_057078	H(+)-transporting two-sector ATPase	ATP6V1D	2	3	1	4
NP_057159	vacuolar protein sorting 36	VPS36	2	6	1	1
NP_057163	vacuolar protein sorting 24 isoform 1	VPS24	1	1	2	3
NP_057215	ras-related GTP-binding protein RAB10	RAB10	1	7	2	11
NP_057292	vacuolar protein sorting 28 isoform 1	VPS28	4	4	5	8

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_057319	G protein-coupled receptor, family C, group 5, member B precursor	GPRC5B	4	21	2	7
NP_057370	dicarbonyl/L-xylulose reductase	DCXR	2	5	3	5
NP_057373	calcium binding protein 39	CAB39	4	6	3	4
NP_057406	GTPase Rab14	RAB14	3	4	3	4
NP_057494	chromatin modifying protein 5	CHMP5	5	17	3	8
NP_057569	Vps20-associated 1 homolog	VTA1	3	8	6	14
NP_057635	transmembrane 7 superfamily member 3	TM7SF3	3	8	2	5
NP_057699	solute carrier family 6 (neurotransmitter transporter, GABA), member 13	SLC6A13	1	2	1	1
NP_059830	WD repeat-containing protein 1 isoform 1	WDR1	5	8	5	10
NP_055113	quinolinate phosphoribosyltransferase	QPRT	1	1	2	3
NP_060187	mu-protocadherin isoform 2	MUCDHL	1	1	1	1
NP_001008274	transgelin 3	TAGLN3	1	2	1	2
NP_060347	hypothetical protein LOC54978	C2orf18	1	4	1	8
NP_060436	vacuolar protein sorting 37C	VPS37C	1	3	1	1
NP_001071089	vacuolar protein sorting 37D	VPS37D	3	10	3	9
NP_060705	CNDP dipeptidase 2	CNDP2	10	30	8	23
NP_001129975	hypothetical protein LOC284422	LOC284422	3	13	3	7
NP_059996	leptin receptor gene-related protein	LEPROT	1	1	1	2
NP_061123	G protein-coupled receptor family C, group 5, member C isoform b	GPRC5C	12	194	10	143
NP_061329	G-protein gamma-12 subunit	GNG12	2	10	1	2
NP_061330	BAI1-associated protein 2-like 1	BAIAP2L1	3	6	1	2
NP_061485	ras-related C3 botulinum toxin substrate 1 isoform Rac1b	RAC1	3	11	3	5
NP_061816	tubulin, alpha 8	TUBA8	1	11	1	4
NP_061907	ras homolog gene family, member F	RHOF	1	1	1	1
NP_062547	sushi domain containing 2	SUSD2	2	5	1	2
NP_062815	transient receptor potential cation channel, subfamily V, member 5	TRPV5	1	2	1	1
NP_064424	tubulin, beta polypeptide 4, member Q	TUBB4Q	2	5	2	8
NP_064524	3-hydroxybutyrate dehydrogenase, type 2	BDH2	3	4	4	7
NP_065145	chromatin modifying protein 1B	CHMP1B	4	9	3	5
NP_065161	solute carrier family 44, member 2	SLC44A2	4	13	3	5
NP_060261	feline leukemia virus subgroup C cellular receptor family, member 2	FLVCR2	1	2	1	4
NP_065842	TAO kinase 1	TAOK1	1	2	1	2
NP_066289	ubiquitin C	UBC	4	83	4	78
NP_066299	myosin, light chain 6, alkali, smooth muscle and non-muscle isoform 1	MYL6	1	1	1	1
NP_066361	RAP2A, member of RAS oncogene family	RAP2A	3	4	2	3
NP_068576	angiotensin I converting enzyme 2 precursor	ACE2	8	16	5	10
NP_068733	cofilin 2	CFL2	1	2	1	3
NP_068814	heat shock 70kDa protein 2	HSPA2	10	28	7	27
NP_068839	integral membrane protein 2B	ITM2B	2	10	2	4
NP_071738	GLI pathogenesis-related 2	GLIPR2	1	1	1	2
NP_071399	fidgetin-like 1	FIGNL1	1	10	1	8
NP_071412	MAWD binding protein isoform a	PBLD	2	3	2	6
NP_071435	transmembrane BAX inhibitor motif containing 1	TMBIM1	1	6	1	4
NP_001173	antiquitin	ALDH7A1	1	2	1	1
NP_073609	epidermal growth factor receptor pathway substrate 8-like protein 2	EPS8L2	5	8	2	2
NP_073740	solute carrier family 13 member 3 isoform a	SLC13A3	2	4	2	4
NP_006139	LIM and SH3 protein 1	LASP1	1	1	1	1
NP_078867	chromatin modifying protein 6	CHMP6	2	7	2	2
NP_078943	vacuolar protein sorting 37B	VPS37B	4	8	4	9
NP_079119	cytochrome b reductase 1	CYBRD1	2	4	1	2
NP_079526	tweety 3	TTYH3	2	2	1	3
NP_079533	NG22 protein isoform 1	SLC44A4	3	38	3	15
NP_109589	phosphotriesterase related	PTER	2	2	2	3
NP_112601	serine carboxypeptidase vitellogenic-like	CPVL	2	6	1	2
NP_115729	vacuolar protein sorting 25	VPS25	1	1	1	1
NP_115788	hypothetical protein LOC84418	C5orf32	1	17	1	9
NP_149124	2',3'-cyclic nucleotide 3' phosphodiesterase	CNP	8	15	2	3
NP_203693	myosin IC	MYO1C	21	54	8	20
NP_003560	syntaxin 7	STX7	1	2	1	1

Table 5.1 Continued

Ref Seq	Protein Name	Gene Symbol	Pep Male	ID Male	Pep Female	ID Female
NP_443102	sorting nexin associated golgi protein 1	SNAG1	2	4	1	2
NP_542416	protein tyrosine phosphatase, non-receptor type 13 isoform 4	PTPN13	1	2	1	2
NP_612410	family with sequence similarity 125, member A	FAM125A	2	6	2	4
NP_612433	kinesin family member 12	KIF12	9	15	4	6
NP_620153	MIT, microtubule interacting and transport, domain containing 1	MITD1	4	13	3	5
NP_620164	carboxymethylenebutenolidase	CMBL	2	3	2	3
NP_631900	crumbs 3 isoform a precursor	CRB3	1	2	1	2
NP_644670	EH-domain containing 4	EHD4	10	20	10	22
NP_647537	attractin isoform 1	ATRN	2	4	2	3
NP_647539	tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, beta polypeptide	YWHAB	6	14	5	14
NP_653186	urate anion exchanger 1 isoform a	SLC22A12	2	4	4	6
NP_653296	hypothetical protein LOC148362	C1orf58	6	30	5	18
NP_653308	prominin 2	PROM2	9	43	6	18
NP_659409	lactate dehydrogenase A-like 6A	LDHAL6A	1	3	1	4
NP_665683	glutathione S-transferase alpha 1	GSTA1	1	1	1	1
NP_665875	annexin A11	ANXA11	23	144	19	73
NP_683718	serine hydroxymethyltransferase 1 (soluble) isoform 2	SHMT1	2	4	5	10
NP_689564	solute carrier family 5 (sodium/glucose cotransporter), member 10	SLC5A10	2	3	1	2
NP_689742	G protein-coupled receptor 155 isoform 10	GPR155	1	7	1	5
NP_689898	solute carrier family 23 (nucleobase transporters), member 1 isoform b	SLC23A1	2	2	4	7
NP_705898	copine VIII	CPNE8	5	16	3	5
NP_006623	solute carrier family 17 (sodium phosphate), member 3	SLC17A3	1	1	1	1
NP_789782	chromatin modifying protein 4B	CHMP4B	4	9	1	2
NP_808592	N-acylsphingosine amidohydrolase (acid ceramidase) 1 preproprotein isoform a	ASAH1	4	8	2	3
NP_821133	tubulin, beta	TUBB	2	6	3	5
NP_857635	peroxiredoxin 5 isoform c precursor	PRDX5	1	1	1	2
NP_859047	peroxiredoxin 1	PRDX1	3	5	4	11
NP_859427	peroxiredoxin 2 isoform b	PRDX2	2	4	3	4
NP_861441	solute carrier family 36 (proton/amino acid symporter), member 2	SLC36A2	7	17	7	18
NP_872270	pyruvate kinase, muscle isoform M1	PKM2	13	35	10	21
NP_872372	vitelline membrane outer layer 1 isoform 1	VMO1	1	2	2	2
NP_874384	synaptophysin-like 1 isoform b	SYPL1	1	8	1	5
NP_908930	fructose-bisphosphate aldolase A	ALDOA	4	10	2	6
NP_932766	aquaporin 1	AQP1	2	46	2	51
NP_937895	gelsolin isoform b	GSN	8	15	3	5
NP_944492	transmembrane 4 superfamily member 8 isoform 2	TSPAN3	1	16	2	11
NP_955376	SFT2 domain containing 2	SFT2D2	1	2	1	3
NP_958842	RAB5C, member RAS oncogene family isoform a	RAB5C	4	9	4	9
NP_974197	regulator of G-protein signalling 19 interacting protein 1 isoform 1	GIPC1	2	4	2	3
NP_976074	CD59 antigen preproprotein	CD59	1	11	1	2
XP_060887	PREDICTED: similar to peptidyl-Pro cis trans isomerase	LOC128192	1	2	1	2

Table 5.2. Proteins that changed at least 2-fold

Ref Seq	Protein Name	Gene
NP_00630	brain abundant, membrane attached signal protein 1	BASP1
NP_05749	chromatin modifying protein 5	CHMP5
NP_00390	copine III	CPNE3
NP_00590	cytosolic malate dehydrogenase	MDH1
NP_00443	epidermal growth factor receptor pathway substrate 8	EPS8
NP_00039	galactosidase, beta 1	GLB1
NP_00660	mannan-binding lectin serine protease 2 isoform 1 precursor	MASP2
NP_00069	Na ⁺ /K ⁺ -ATPase alpha 1 subunit isoform a proprotein	ATP1A1
NP_07953	NG22 protein isoform 1	SLC44A4
NP_85904	peroxiredoxin 1	PRDX1
NP_65330	prominin 2	PROM2
NP_00462	RAB7, member RAS oncogene family	RAB7A
NP_00033	solute carrier family 12 (sodium/chloride transporters), member	SLC12A3
NP_00628	tumor susceptibility gene 101	TSG101
NP_00388	Fc fragment of IgG binding protein	FCGBP
NP_00153	heat shock protein beta-1	HSPB1
NP_62015	MIT, microtubule interacting and transport, domain containing 1	MITD1
NP_00069	annexin I	ANXA1
NP_82113	tubulin, beta	TUBB

6 Summary and Conclusion

Urinary exosomes are small extracellular vesicles that originate from the internal vesicles of multivesicular bodies in renal epithelial cells, including glomerular podocytes, renal tubule cells and the cells lining the renal pelvis, ureters and bladder. Exosomes contain membrane and cytosolic proteins. The mechanism for exosome formation and excretion starts by the endocytosis of apical membrane proteins into the cell via clathrin-coated vesicles. These vesicles are delivered to early endosomes and subsequently fuse with the multivesicular body (MVB). The apical plasma membrane proteins are internalized into the MVB by membrane invagination forming “cytosol-in” vesicles. The outer membrane of the MVB fuses with the apical plasma membrane and releases these internal vesicles through exocytosis in the urinary space. Urinary exosomes are isolated by ultracentrifugation and provide a non-invasive means of acquiring information about the physiological or pathophysiological state of renal cells.

The overall objective of this research was to develop methods and knowledge infrastructure for urinary proteomics. The central focus of the proposed work was on urinary exosome analysis by protein mass spectrometry. The first objective had two components: (1) to expand the proteome of human urinary exosomes by using a highly sensitive LC-MS/MS system, improved software for identification of peptide ions and a more elaborate data analysis strategy, and (2) to investigate the phosphoproteome of human urinary exosomes using neutral loss scanning. We unambiguously identified 1132 proteins including: 177 proteins that are disease related (Table 3.1), 81 proteins involved in solute and water transport (Table 3.2), 22 proteins that are recognized components of the apparatus responsible for the formation of MVB (Table 3.3), and 17 proteins that are

subunits of the human vacuolar H⁺-ATPase (Table 3.4). The phosphoproteomic profiling identified 19 phosphorylation sites corresponding to 14 phosphoproteins. The phosphorylation sites identified include six previously known and eight previously unknown. These results show that the newly identified proteins in urinary exosomes may have considerable relevance to renal biology and the mechanism of renal disease. The detection of phosphorylation, the most important post-translational modification of proteins, provides another level of information about physiological and pathophysiological state of renal cells.

The second objective was to apply urinary exosome analysis using polyclonal antibodies. We proposed to qualitatively assess patients' samples by generating polyclonal antibodies to bind to an epitope in the amino-terminal of the human MRP4 protein, and to use previously generated polyclonal antibodies to bind to epitopes of the NKCC2 and NCC proteins. The potential usefulness of urinary exosome analysis was demonstrated using the well-defined renal tubulopathy, Bartter syndrome type I and using the single nucleotide polymorphism in the ABCC4 gene.

Patients with Bartter syndrome type I have mutations in the SLC12A1 gene which encodes for the NKCC2 protein. NKCC2 is expressed on the apical membrane of the thick ascending limb of Henle. We used a previously generated anti-NKCC2 antibody to immunoblot the urinary exosomes samples from two patients with Bartter syndrome type I. The patient samples showed a complete absence of the NKCC2 protein. In addition, we used a previously generated anti-NCC antibody to ensure that urinary exosomes were successfully isolated. The patient samples showed strong NCC bands.

Patients with a single nucleotide polymorphism of the ABCC4 gene which results in the expression of a short and normal gene product have evidence of potential loss of function of the MRP4 protein. We generated a polyclonal anti-MRP4 antibody to bind to an epitope in the amino-terminal of the protein. The polyclonal antibody was raised against a synthetic peptide that was designed to be highly specific. The synthetic peptide is immunogenic, unique and lacks post-translational modifications. The anti-MRP4 antibody was characterized in kidney tissue samples and shown to be highly specific. Urinary exosome patient samples were probed with our polyclonal antibody and showed a band at the expected molecular weight of the MRP4 protein and a fainter band at the expected molecular weight for the short protein. The results showed that urinary exosome analysis can be useful in gaining insight of genetic mutations that cause disease.

The third objective was to study the normal variability between proteomes of female and male urinary exosomes and to implement a normalization method to analyze urinary exosome samples. We compared urinary exosome samples from pooled samples of urine from four normal human females and four normal human males. The samples were solubilized in SDS and the proteins were separated by SDS-PAGE followed by in-gel trypsin digestion of contiguous gel slices. The digests were analyzed by tandem mass spectrometry. The samples were compared quantitatively using a label-free approach with the integration software, QUOIL. There were 388 proteins identified in both male and female urinary exosome proteomes. Only 19 proteins had a 2-fold change representing 4.9% of the total number of proteins identified. Therefore, there is high concordance between proteomes of urinary exosomes isolated from males and females. In addition, we implemented the normalization method, timed urine collection, to

quantify the excretion of exosomal proteins. We collaborated with Dr. Jeffrey Kopp to enroll six healthy volunteers, three males and three females, and analyzed their urinary exosome samples using immunoblotting. The results showed that the normalization method we implemented did not correlate with the intensity signal of MVB markers, TSG101 and Alix. Further work needs to be completed to implement timed urine collection as a normalization method.

The work completed showed that the analysis of urinary exosomes by protein mass spectrometry can be the basis for future biomarker studies as well as physiological studies. We used a highly sensitive LC-MS/MS system to expand the known proteome of human urinary exosomes. We used neutral loss scanning in LC-MS/MS to investigate the phosphoproteome of human urinary exosomes. We applied the analysis of human urinary exosomes to gain understanding of genetic mutations and we generated highly-specific polyclonal antibodies to be used for immunoassays. And, we studied the normal variability among human populations based on gender.

The analysis of urinary exosomes provides a non-invasive method potentially useful in future biomarker discovery studies. Therefore, in this work we proposed to investigate the possible use of analysis of urinary exosomes in clinical proteomics. The overall objective of this research is to develop methods and knowledge infrastructure for urinary proteomics. The central focus was urinary exosome analysis by protein mass spectrometry. From the results found, it is clear that expanding the known proteome of human urinary exosomes and investigating the phosphoproteome of human urinary exosomes provides a basis for biomarker studies. Human urinary exosome analysis can be applied to understand and analyze genetic mutations. And finally, we showed the

importance of studying normal variability among human populations as well as implementing normalization methods for future biomarker studies.

7 Future Works

The potential of urinary exosome analysis as a route to biomarker discovery in renal diseases still faces many limitations despite our current progress [7, 8, 31]. More work needs to be done to standardize protocols to collect, process, and store urine samples in a clinical setting. Our current isolation approach for urinary exosomes includes an ultracentrifugation step that needs to be replaced for future clinical studies.

7.1 Isolation Method

The ultracentrifugation step requires expensive instrumentation and long processing times would hinder clinical studies. Filtration methods have been proposed to replace the ultracentrifugation step [32] but these tend to retain and concentrate soluble proteins in addition to exosomes. The Tamm-Horsfall protein hinders the concentration of exosomes using filtration methods.

7.2 Tamm-Horsfall

A particular problem is the excretion of Tamm-Horsfall protein, an extraordinary abundant urinary protein that interferes with successful mass spectrometry and immunoblotting [9]. Tamm-Horsfall protein is an 80 kDa glycoprotein produced abundantly by kidney cells [63]. This protein is excreted in varying amounts that hinder the quantification of samples. Current work addressing this problem shows that Tamm-Horsfall can be depolymerized using a reducing agent, DTT, after the 17,000 x g centrifugation step in addition to heat (37 °C).

7.2 Quantification

In the long run, the most important technical challenge may be to develop quantification approaches that allow detection of changes in excretion rates of particular

biomarker candidates. Both labeling and non-labeling methods have been developed to make protein mass spectrometry quantitative [21]. Use of creatinine as a normalizing variable may be inadequate because of high subject-to-subject variability in its rate of excretion but may need to be used in combination with other methods of quantification [8]. An alternative method is to quantify the excretion of exosomal proteins by counting the number of exosomes using negative-stain electron micrograph (EM). The number of exosomes can be correlated to exosomes markers such as TSG-101 and Alix.

Finally, we have shown that even without quantification, urinary exosome analysis can be valuable in situations such as genetic diseases where a protein may be entirely absent from urinary exosomes.

8 Bibliography

1. Pandey A, Mann M: Proteomics to Study Genes and Genomes. *Nature* 405: 837-846, 2000
2. Peng J, Gygi SP: Proteomics: the Move to Mixtures. *J Mass Spectrom* 36: 1083-1091, 2001
3. Knepper MA: Proteomics and the Kidney. *J Am Soc Nephrol*. 13: 1398-1408, 2002
4. Madsen KM, Tisher CC: Anatomy of the Kidney. In: *The Kidney*, 7th ed., edited by Brenner BM, Philadelphia: WB Saunders, 2004
5. Knepper M, Burg M: Organization of Nephron Function. *Am J Phy Renal Phy* 244: F579-F589, 1983
6. Pisitkun T, Shen RF, Knepper MA: Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci U S A* 101: 13368–13373, 2004
7. Zhou H, Yuen PS, Pisitkun T, Gonzales PA, Yasuda H, Dear JW, Gross P, Knepper P, Star RA: Collection, storage, preservation, and normalization of human urinary exosomes for biomarker discovery. *Kidney Int* 69: 1471–1476, 2006
8. Hoorn EJ, Pisitkun T, Zietse R, Gross P, Frokiaer J, Wang NS, Gonzales PA, Star RA, Knepper MA: Prospects for urinary proteomics: Exosomes as a source of urinary biomarkers. *Nephrology (Carlton)* 10: 283–290, 2005
9. Pisitkun T, Johnstone R, Knepper MA: Discovery of urinary biomarkers. *Mol Cell Proteomics* 5: 1760–1771, 2006
10. Nielsen S, Frokier J, Marples D, Kwon TH, Agre P, Knepper MA: Aquaporins in the Kidney: From molecules to Medicine. *Physiol Rev* 82: 202-244, 2002
11. Johnstone RM, Adam M, Hammond JR, Orr L, Turbide C: Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). *J Biol Chem* 262: 9412-9420, 1987
12. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: Intracellular Vesicular Traffic: Introduction. In: *Molecular Biology of the Cell*, 4th ed., 2002, pp 711
13. Stoorvogel W., Kleijmeer M. J., Geuze H.J., Raposo G: The Biogenesis and Functions of Exosomes. Review. *Traffic* 2002; 3: 321-330.
14. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: Intracellular Vesicular Traffic: Transport into the Cell from the Plasma Membrane: Endocytosis. In: *Molecular Biology of the Cell*, 4th ed., 2002, pp 746-757.

15. Simpson RJ, Jensen SS, Lim JWE: Proteomic profiling of exosomes: Current perspectives. *Proteomics* 8: 4083-4099, 2008
16. Williams RL, Urbé S: The emerging shape of the ESCRT machinery. *Nat Rev Mol Cell Biol* 8: 355–368, 2007
17. Tanaka N., Kyuuma M., Sugamura K.: Endosomal sorting complex required for transport proteins in cancer pathogenesis, vesicular transport, and non-endosomal functions. *Cancer Sci* 99 (7): 1293-1303, 2008
18. Ong SE, Mann M: Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology* 1 (5): 252-262, 2005
19. MacBeath G: Protein Microarrays and Proteomics. *Nature Genetics Supplement*. 32: 526-532 (2002)
20. Dass C: Introduction to mass spectrometry. In: *Principles and Practice of Biological Mass Spectrometry*. New York, Wiley-Interscience, 2001, pp 1-10
21. Pisitkun T, Hoffert JD, Yu MJ, Knepper MA: Tandem mass spectrometry in physiology. *Physiology (Bethesda)* 22: 390–400, 2007
22. Dass C: Ionization methods. In: *Principles and Practice of Biological Mass Spectrometry*. New York, Wiley-Interscience, 2001, pp 11-58
23. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246(4926): 64-71, 1989
24. Dass C: Coupling of separation techniques with mass spectrometry. In: *Principles and Practice of Biological Mass Spectrometry*. New York, Wiley-Interscience, 2001, pp 143-174
25. Dass C: Tandem mass spectrometry. In: *Principles and Practice of Biological Mass Spectrometry*. New York, Wiley-Interscience, 2001, pp 95-115
26. Eng J.K., McCormack A. L., Yates III J.R.: An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J Am Soc Mass Spectrom* 5: 976-989, 1994
27. Sadygov RG, Cociorva D, Yates JR 3rd: Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nat Methods* 1: 195–202, 2004
28. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP: A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat*

- Biotechnol* 24: 1285-1292, 2006
29. Tanner S, Shu H, Frank A, Wang LC, Zandi E, Mumby M, Pevzner P, Bafna V: InsPecT: Identification of posttranslationally modified peptides from tandem mass spectra. *Anal Chem* 77: 4626–4639, 2005
 30. Granger CB, Van Eyk JE, MOckrin SC, Anderson NL: National Heart, Lung and Blood Institute Clinical Proteomics Working Group report. *Circulation* 109 (14): 1697-1703, 2004
 31. Mayr M, Zhang J, Greene AS, Gutterman D, Ping P: Proteomics-based development of biomarkers in cardiovascular disease. Mechanistic, clinical and therapeutic insights. Review. *Mol Cell Proteomics* 5(10): 1853-1864, 2006
 32. Gonzales P., Pisitkun T., Knepper M.A.: Urinary exosomes: is there a future?. *Nephrol Dial Transplant* 23: 1799-1801, 2008
 33. Cheruvanky A, Zhou H, Pisitkun T, Kopp JB, Knepper MA, Yuen PS, Star RA: Rapid isolation of urinary exosomal biomarkers using a nanomembrane ultrafiltration concentrator. *Am J Physiol Renal Physiol* 292: F1657–F1661, 2007
 34. Hoffert JD, Wang G, Pisitkun T, Shen RF, Knepper MA: An automated platform for analysis of phosphoproteomic datasets: Application to kidney collecting duct phosphoproteins. *J Proteome Res* 6: 3501– 3509, 2007
 35. Hoffert JD, Pisitkun T, Wang G, Shen RF, Knepper MA: Quantitative phosphoproteomics of vasopressin-sensitive renal cells: Regulation of aquaporin-2 phosphorylation at two sites. *Proc Natl Acad Sci USA* 103: 7159-7164, 2006
 36. Forgac M. Vacuolar ATPases: Rotary proton pumps in physiology and pathophysiology. *Nat Rev Mol Cell Biol* 8: 917-929, 2007
 37. Gluck SL, Underhill DM, Iyori M, Holliday LS, Kostrominova TY, Lee BS: Physiology and biochemistry of the kidney vacuolar H⁺-ATPase. *Annu Rev Physiol* 58: 427– 445, 1996
 38. Wagner CA, Finberg KE, Breton S, Marshansky V, Brown D, Geibel JP: Renal vacuolar H⁺-ATPase. *Physiol Rev* 84: 1263-1314, 2004.
 39. Chen WG, White FM: Proteomic analysis of cellular signaling. *Expert Rev Proteomics* 1: 343-354, 2004.
 40. Moriguchi T, Urushiyama S, Hisamoto N, Lemura S, Uchida S, Natsume T, Matsumoto K, Shibuya H: WNK1 regulates phosphorylation of cation-chloride-coupled cotransporters via the STE20-related kinases, SPAK and OSR1. *J Biol Chem* 280: 42635-42693, 2005

41. Simon DB, Nelson-Williams C, Bia MJ, Ellison D, Karet FE, Molina AM, Vaara I, Iwata F, Cushner HM, Koolen M, Gainza FJ, Gitelman HJ, Lifton RP: Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl cotransporter. *Nat Genet* 12: 24–30, 1996
42. Kuwahara M, Fushimi K, Terada Y, Bai L, Marumo F, Sasaki S: cAMP-dependent phosphorylation stimulates water permeability of aquaporin-collecting duct water channel protein expressed in *Xenopus* oocytes. *J Biol Chem* 270: 10384–10387, 1995
43. Gimenez I, Forbush B: Regulatory phosphorylation sites in the N-terminus of the renal Na-K-Cl cotransporter (NKCC2). *Am J Physiol Renal Physiol* 289: F1341–F1345, 2005
44. Villén J, Beausoleil SA, Gerber SA, Gygi SP: Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A* 104: 1488–1493, 2007
45. Lees-Miller SP, Anderson CW: Two human 90-kDa heat shock proteins are phosphorylated *in vivo* at conserved serines that are phosphorylated *in vitro* by casein kinase II. *J Biol Chem* 264: 2431–2437, 1989
46. Christensen B, Nielsen MS, Haselmann KF, Petersen TE, Sorensen ES: Post-translationally modified residues of native human osteopontin are located in clusters: Identification of 36 phosphorylation and five O-glycosylation sites and their biological implications. *Biochem J* 390: 285–292, 2005
47. Finberg KE, Wagner CA, Bailey MA, Paunescu TG, Breton S, Brown D, Giebisch G, Geibel JP, Lifton RP: The B1-subunit of the H⁺ ATPase is required for maximal urinary acidification. *Proc Natl Acad Sci U S A* 102: 13616–13621, 2005
48. Royaux IE, Wall SM, Karniski LP, Everett LA, Suzuki K, Knepper MA, Green ED: Pendrin, encoded by the Pendred syndrome gene, resides in the apical region of renal intercalated cells and mediates bicarbonate secretion. *Proc Natl Acad Sci U S A* 98: 4221–4226, 2001
49. Gamba G, Miyanoshita A, Lombardi M, Lytton J, Lee WS, Hediger MA, Hebert SC: Molecular cloning, primary structure, and characterization of two members of the mammalian electroneutral sodium-(potassium)-chloride cotransporter family expressed in kidney. *J Biol Chem* 269 (26): 17713–17722, 1994
50. Kleta R, Bockenhauer D: Bartter syndromes and other salt-losing tubulopathies. *Nephron Physiol* 104: 73080, 2006.
51. Simon DB, Karet FE, Hamdan JM, DiPietro A, Sanjad SA, Lifton RP: Bartter's syndrome, hypokalaemic alkalosis with hypercalciuria, is caused by mutations in the Na-K-2Cl cotransporter NKCC2. *Nat Genet* 13: 183–188, 1996

52. Kim G-H, Masilamani S, Turner R, Mitchell C, Wade JB, Knepper MA: The thiazide-sensitive Na-Cl cotransporter is an aldosterone-induced protein. *Proc Natl Acad Sci* 95: 14552-14557, 1998
53. Russel F.G.M, Koenderink J.B., Masereeuw R.: Multidrug resistance protein 4 (MRP4/ABCC4): a versatile efflux transporter for drugs and signaling molecules. *J Mol Biol* 157: 105-132, 1982
54. Knepper M.A., Masilamani S.: Targeted proteomics in the kidney using ensembles of antibodies. *Acta Physiol Scand* 173: 11-21, 2001
55. Kim G-H, Ecelbarger CA, Mitchell C, Packer RK, Wade JB, Knepper MA: Vasopressin increases Na-K-2Cl cotransporter expression in thick ascending limb of Henle's loop. *Am J Physiol Renal Physiol* 276: 96-103, 1999
56. Kyte J., Doolittle R.F.: A Simple Method for Displaying the Hydrophobic Character of a Protein. *J. Mol. Biol* 157: 105-132, 1982
57. Van Aobel R, Smeets P, Peters J, Bindels R, Russel F: The *MRP4/ABCC4* Gene Encodes a Novel Apical Organic Anion Transporter in Human Kidney Proximal Tubules: Putative Efflux Pump for Urinary cAMP and cGMP. *J Am Soc Nephrol* 13: 595-603, 2002
58. Abla N, Chinn LW, Nakamura T, Liu L, Huan CC, Johns SJ, Kawamoto M, Stryke D, Taylor TR, Ferrin TE, Giacomini KM, Kroetz DL: The human multidrug resistance protein (MRP4, ABCC4): function analysis of a highly polymorphic gene. *J Pharmacol Exp Ther* 325 (3): 859-868, 2008
59. Liddell E and Weeks I: Production of Monoclonal Antibodies. In: *Antibody Technology*, edited by Graham JM, Billington D, 1995, pp 25-44.
60. Wang G., Wu W.W., Zeng W., Chou C-L, Shen R-F.: Label-Free Protein Quantification Using LC-Coupled Ion Trap or FT Mass Spectrometry: Reproducibility, Linearity, and Application with Complex Proteomes. *J. Proteome Res* 5 (5): 1214-1223, 2006
61. Simmons BL, Wang G, Shen RF, Knepper MA: In vacuo isotope coded alkylation technique (IVICAT); an N-terminal stable isotopic label for quantitative liquid chromatography/mass spectrometry proteomics. *Rapid Commun Mass Spectrom* 20(16): 2463-2477, 2006
62. Masilamani S, Wang X, Kim GH, Brooks H, Nielsen J, Nielsen S, Nakamura K, Stokes JB, Knepper MA: Time course of renal Na-K-ATPase, NHE3, NKCC2, NCC and ENaC abundance changes with dietary NaCl restriction. *Am J Physiol Renal Physiol* 283 (4): F648-F657, 2002

63. Kobayashi K, Fukuoka S: Conditions for Solubilization of Tamm-Horsfall Protein/Uromodulin in Human Urine and Establishment of a Sensitive and Accurate Enzyme-Linked Immunosorbent Assay (ELISA) Method. *Arch Biochem Biophys* 388 (1): 113-120, 2001

Appendix A

Electron Micrograph Negative Staining Protocol with Uranyl Acetate

NHLBI EM Core Standard Procedure

Materials:

1. Staining Solution: 1% uranyl acetate (w/v) in deionized water. Stir in dark until completely dissolved. Allow any undissolved material to settle. Store in dark. In order to avoid precipitates, do not withdraw solution from bottom of container.

Solution may be made by dilution of 2% uranyl acetate prepared in the same way.

Note: Uranyl acetate is manufactured from depleted uranium but it still gives off a small amount of alpha radiation. Therefore all wet and dry uranyl acetate waste must be disposed of in marked containers as radioactive waste. As it is a heavy metal, uranyl acetate is toxic, and should be handled and disposed of as such.

2. Specimen Grids: Copper EM grids with formvar and carbon film (Electron Microscopy Sciences).

Procedure:

1. On day that staining is to be done, subject the formvar/carbon grids to glow discharge (plasma cleaning) treatment for one minute to make the film hydrophilic.
2. Holding a grid in a forceps, apply 5 μ l of the particulate suspension to the grid. It should spread immediately to cover the grid.
3. After 1 min, remove excess suspension by touching a piece of filter paper to the edge of the grid.

4. Immediately wash and stain by quickly running 4 drops of 1% uranyl acetate over the grid.
5. Immediately remove excess uranyl acetate from the grid with filter paper as in step 3.
6. Allow the grid to air dry at least 5 min before inserting into the electron microscope.
7. The samples were viewed with a JEM 1200EX electron microscope (JEOL USA) at 80kV and a magnification of 20,000X. Images were acquired with an AMT XR60B 6.8 megapixel digital camera (Advanced Microscope Technologies).

Appendix B: Supplementary Table 1. Complete List of Proteins in Human Urinary Exosomes

Amb	Technique	Gi Number	Ref Seq	Gene	Protein Name	Pep	ID	GO Component	GO Process	GO Function	Related to Disease [OMIM]
	2	21071030	NP_570602	A1BG	alpha 1B-glycoprotein	1	1	extracellular region	not classified	not classified	
Yes	2	66932947	NP_000005	A2M	alpha-2-macroglobulin precursor	21	43	extracellular region	intracellular protein transport protein homooligomerization	enzyme binding interleukin-1 binding interleukin-8 binding protein carrier activity serine-type endopeptidase inhibitor activity tumor necrosis factor binding wide-spectrum protease inhibitor activity	
	2	74271845	NP_653271	A2ML1	alpha-2-macroglobulin-like 1	1	1	not classified	not classified	endopeptidase inhibitor activity	
	2	20357568	NP_056238	AASDHPPT	aminoadipate-semialdehyde dehydrogenase	1	3	not classified	fatty acid biosynthesis	magnesium ion binding phosphopantetheinyltransferase activity transferase activity	
	1,2	42741659	NP_000918	ABCB1	ATP-binding cassette sub-family B member 1	23	61	cell surface integral to membrane membrane membrane fraction	response to drug transport	ATP binding ATPase activity hydrolase activity nucleotide binding transporter activity xenobiotic-transporting ATPase activity	Colchicine Resistance [MIM : 120080] Crohn Disease [MIM:266600]
	2	21536378	NP_003733	ABCB11	ATP-binding cassette, sub-family B (MDR/TAP	1	3	integral to plasma membrane membrane membrane fraction	transport	ATP binding ATPase activity bile acid-exporting ATPase activity nucleotide binding sodium-exporting ATPase activity, phosphorylative mechanism transporter activity	Cholestasis, Progressive Familial Intrahepatic 2 [MIM:601847] Cholestasis, Benign Recurrent Intrahepatic 2 [MIM:605479]
Yes	2	4505771	NP_000434	ABCB4	ATP-binding cassette, subfamily B, member	7	20	integral to membrane integral to plasma membrane membrane membrane fraction	lipid metabolism response to drug response to xenobiotic stimulus transport	ATP binding ATPase activity ATPase activity, coupled to transmembrane movement of substances hydrolase activity nucleotide binding xenobiotic-transporting ATPase activity	
	2	9955963	NP_005680	ABCB6	ATP-binding cassette, sub-family B, membe	1	1	ATP-binding cassette (ABC) transporter complex integral to membrane membrane mitochondrial envelope mitochondrion	iron ion homeostasis transport	ATP binding ATPase activity ATPase activity, coupled to transmembrane movement of substances nucleotide binding	
	2	15149474	NP_149163	ABCC11	ATP-binding cassette, sub-family C, membe	1	1	integral to membrane membrane	transport	ATP binding ATPase activity ATPase activity, coupled to transmembrane movement of substances nucleotide binding	
	2	110832839	NP_064694	ABCC9	ATP-binding cassette, sub-family C, member 9	1	2	integral to membrane membrane	potassium ion transport transport	ATP binding ATPase activity ATPase activity, coupled to transmembrane movement of substances nucleotide binding receptor activity sulfonylurea receptor activity transporter activity	Cardiomyopathy [MIM: 608569]
	1,2	14249382	NP_116139	ABHD14B	abhydrolase domain containing 14B	6	20	nucleus	not classified	not classified	

Supplementary Table 1

	2	61743948	NP_001012770	ABI1	abl-interactor 1 isoform d	2	2	cytoskeleton cytosol endoplasmic reticulum nucleus soluble fraction	negative regulation of cell proliferation transmembrane receptor protein tyrosine kinase signaling pathway	cytoskeletal protein binding	
	2	4557237	NP_000010	ACAT1	acetyl-Coenzyme A acetyltransferase 1 precursor	1	2	mitochondrion	not classified	acetyl-CoA C-acetyltransferase activity acyltransferase activity transferase activity	Alpha-methylacetoacetic Aciduria [MIM:203750]
	1,2	4503273	NP_000780	ACE	angiotensin I converting enzyme isoform 1 precursor	23	96	integral to membrane membrane membrane fraction plasma membrane soluble fraction	blood pressure regulation metabolism proteolysis	carboxypeptidase activity chloride ion binding hydrolase activity, acting on glycosyl bonds metal ion binding peptidyl-dipeptidase A activity zinc ion binding	Hypertension [MIM:106180]
	2	23238214	NP_690043	ACE	angiotensin I converting enzyme isoform 2 precursor	12	61	integral to membrane membrane membrane fraction plasma membrane soluble fraction	blood pressure regulation metabolism proteolysis	carboxypeptidase activity chloride ion binding hydrolase activity, acting on glycosyl bonds metal ion binding peptidyl-dipeptidase A activity zinc ion binding	Renal Tubular Dysgenesis [MIM:267430]
	2	11225609	NP_068576	ACE2	angiotensin I converting enzyme 2 precursor	8	17	extracellular region integral to membrane membrane	entry of virus into host cell proteolysis	carboxypeptidase activity chloride ion binding metal ion binding peptidyl-dipeptidase A activity protein binding viral receptor activity zinc ion binding	Hypertension [MIM:300335]
	2	38569423	NP_942127	ACLY	ATP citrate lyase isoform 2	1	1	citrate lyase complex	ATP catabolism citrate metabolism coenzyme A metabolism lipid biosynthesis metabolism	ATP binding ATP citrate synthase activity catalytic activity magnesium ion binding nucleotide binding transferase activity	
	2	8659555	NP_002188	ACO1	aconitase 1	6	9	cytoplasm	metabolism tricarboxylic acid cycle	4 iron, 4 sulfur cluster binding aconitate hydratase activity iron ion binding lyase activity metal ion binding RNA binding	
	2	22165400	NP_671517	ACOT11	thioesterase, adipose associated isoform B	1	1	cytoplasm	fatty acid metabolism intracellular signaling cascade response to temperature stimulus	acyl-CoA thioesterase activity hydrolase activity serine esterase activity	
	2	32528286	NP_863656	ACOT7	acyl-CoA thioesterase 7 isoform hBACHd	1	1	cytoplasm mitochondrion	lipid metabolism	acyl-CoA binding hydrolase activity palmitoyl-CoA hydrolase activity serine esterase activity	Mesial temporal lobe epilepsy [MIM:608096]
	2	6005988	NP_009030	ACP1	acid phosphatase 1 isoform b	1	2	cytoplasm soluble fraction	protein amino acid dephosphorylation	acid phosphatase activity hydrolase activity non-membrane spanning protein tyrosine phosphatase activity protein tyrosine phosphatase activity	
	1,2	4557010	NP_001601	ACP2	lysosomal acid phosphatase 2 precursor	7	19	integral to membrane lysosomal membrane lysosome	not classified	acid phosphatase activity hydrolase activity	

Supplementary Table 1

	1,2	6382064	NP_001090	ACPP	prostatic acid phosphatase precursor	8	13	extracellular region	regulation of progression through cell cycle	acid phosphatase activity hydrolase activity protein tyrosine phosphatase activity	
Yes	2	42794752	NP_004448	ACSL3	acyl-CoA synthetase long-chain family mem	1	2	integral to membrane membrane peroxisome	fatty acid metabolism lipid metabolism metabolism	fatty-acyl-CoA synthase activity ligase activity long-chain-fatty-acid-CoA ligase activity magnesium ion binding	
	2	12669909	NP_075266	ACSL4	acyl-CoA synthetase long-chain family member	1	2	integral to membrane membrane peroxisome	fatty acid metabolism learning and/or memory lipid metabolism metabolism	ligase activity long-chain-fatty-acid-CoA ligase activity magnesium ion binding	Mental Retardation, X-linked 63, MRX 63 [MIM:300387]
	2	115511026	NP_443188	ACSM1	acyl-CoA synthetase medium-chain family m	1	1	mitochondrial matrix	benzoate metabolism butyrate metabolism energy derivation by oxidation of organic compounds fatty acid oxidation metabolism	acyl-CoA ligase activity butyrate-CoA ligase activity catalytic activity	
Yes	2	4501883	NP_001604	ACTA2	alpha 2 actin	15	121	actin filament cytoskeleton	not classified	integral to membrane lysosomal membrane lysosome	
Yes	2	4501885	NP_001092	ACTB	beta actin	26	389	actin filament cytoskeleton TIP60 histone acetyltransferase complex	not classified	ATP binding nucleotide binding protein binding structural constituent of cytoskeleton	
Yes	2	4501887	NP_001605	ACTG1	actin, gamma 1 propeptide	13	128	actin filament cytoskeleton	not classified	ATP binding nucleotide binding protein binding structural constituent of cytoskeleton	
Yes	2	4501889	NP_001606	ACTG2	actin, gamma 2 propeptide	2	23	actin filament cytoskeleton	not classified	ATP binding nucleotide binding protein binding structural constituent of cytoskeleton	
	2	4501891	NP_001093	ACTN1	actinin, alpha 1	1	2	cytoskeleton focal adhesion pseudopodium	focal adhesion formation negative regulation of cell motility regulation of apoptosis	actin binding calcium ion binding integrin binding protein binding vinculin binding	
Yes	2	4501893	NP_001094	ACTN2	actinin, alpha 2	5	9	actin filament cytoskeleton dendritic spine filopodium focal adhesion pseudopodium	cell adhesion focal adhesion formation microspike biogenesis protein homotetramerization regulation of apoptosis	actin binding calcium ion binding FAT2 1 binding identical protein binding integrin binding protein dimerization activity structural constituent of muscle ZASP binding	
Yes	2	12025678	NP_004915	ACTN4	actinin, alpha 4	15	35	cytoplasm nucleus perinuclear region protein complex pseudopodium	positive regulation of cell motility regulation of apoptosis regulation of progression through cell cycle	actin filament binding calcium ion binding integrin binding nucleoside binding protein binding protein homodimerization activity	
	2	5031569	NP_005727	ACTR1A	ARP1 actin-related protein 1 homolog A, cen	1	1	actin filament cytoplasm cytoskeleton	not classified	ATP binding nucleotide binding protein binding structural constituent of cytoskeleton	
	2	53692187	NP_001005386	ACTR2	actin-related protein 2 isoform a	1	1	cytoskeleton	not classified	protein binding structural molecule activity	

Supplementary Table 1

	2	5031573	NP_005712	ACTR3	ARP3 actin-related protein 3 homolog	6	13	Arp2/3 protein complex	cell motility	ATP binding nucleotide binding protein binding structural molecule activity	
Yes	2	92373393	NP_001035225	ACTR3B	actin-related protein 3-beta isoform 2	1	3	cytoskeleton	actin filament polymerization	actin filament binding protein binding	
	1,2	4501901	NP_000657	ACY1	aminoacylase 1	15	43	cytoplasm cytosol	amino acid metabolism proteolysis	aminoacylase activity hydrolase activity metallopeptidase activity protein dimerization activity zinc ion binding	Amimoacylase 1 Deficiency [MIM:609924]
	2	18087825	NP_542389	ACY3	aspartoacylase (aminocyclase) 3	1	1	not classified	metabolism	aspartoacylase activity hydrolase activity, acting on ester bonds	
	2	31083193	NP_066939	ADCY1	brain adenylate cyclase 1	1	1	integral to membrane membrane	cAMP biosynthesis intracellular signaling cascade	calcium- and calmodulin-responsive adenylate cyclase activity calmodulin binding magnesium ion binding	
	2	71565154	NP_000662	ADH5	class III alcohol dehydrogenase 5 chi subun	1	2	not classified	ethanol oxidation	alcohol dehydrogenase activity electron carrier activity fatty acid binding formaldehyde dehydrogenase (glutathione) activity metal ion binding S-(hydroxymethyl)glutathione dehydrogenase activity zinc ion binding	
	2	4501939	NP_000663	ADH6	class V alcohol dehydrogenase 6	1	1	not classified	ethanol oxidation	alcohol dehydrogenase activity electron carrier activity metal ion binding oxidoreductase activity zinc ion binding	
	2	67551265	NP_056261	AHCTF1	transcription factor ELYS	1	3	nucleus	development transcription	DNA binding	
	2	9951915	NP_000678	AHCY	S-adenosylhomocysteine hydrolase	10	28	cytoplasm	one-carbon compound metabolism	adenosylhomocysteine activity hydrolase activity	Hypermethioninemia [MIM:180960]
	2	61743954	NP_001611	AHNAK	AHNAK nucleoprotein isoform 1	1	1	nucleus	nervous system development	protein binding	
	2	6912280	NP_036243	AHSA1	AHA1, activator of heat shock 90kDa protein	1	1	cytoplasm endoplasmic reticulum	protein folding response to stress	ATPase stimulator activity chaperone activator activity chaperone binding	
	2	4502005	NP_001613	AHSG	alpha-2-HS-glycoprotein (Fetuin-A)	1	6	extracellular region extracellular space	acute-phase response negative regulation of bone mineralization negative regulation of insulin receptor signaling pathway pinocytosis positive regulation of phagocytosis regulation of inflammatory response skeletal development	cysteine protease inhibitor activity kinase inhibitor activity	
	2	4502011	NP_000467	AK1	adenylate kinase 1	4	4	cytoplasm cytosol	ATP metabolism nucleobase, nucleoside, nucleotide and nucleic acid metabolism	adenylate kinase activity ATP binding kinase activity nucleotide binding transferase activity	Hemolytic Anemia due to AK1 Deficiency [MIM:103000]

Supplementary Table 1

	2	5174391	NP_006057	AKR1A1	aldo-keto reductase family 1, member A1	7	16	not classified	aldehyde metabolism glucose metabolism	alcohol dehydrogenase (NADP+) activity aldehyde reductase activity electron carrier activity oxidoreductase activity protein binding	
	1,2	4502049	NP_001619	AKR1B1	aldo-keto reductase family 1, member B1	3	17	extracellular space	carbohydrate metabolism	aldehyde reductase activity electron carrier activity oxidoreductase activity	
Yes	2	4503285	NP_001345	AKR1C2	aldo-keto reductase family 1, member C2	1	1	not classified	digestion lipid metabolism prostaglandin metabolism steroid metabolism	3-alpha-hydroxysteroid dehydrogenase (A-specific) activity carboxylic acid binding oxidoreductase activity trans-1,2-dihydrobenzene-1,2-diol dehydrogenase activity	
Yes	2	24497583	NP_003730	AKR1C3	aldo-keto reductase family 1, member C3	2	4	intracellular	electron transport prostaglandin metabolism	3-alpha-hydroxysteroid dehydrogenase (A-specific) activity aldo-keto reductase activity estradiol 17-beta-dehydrogenase activity oxidoreductase activity prostaglandin-F synthase activity trans-1,2-dihydrobenzene-1,2-diol dehydrogenase activity	
Yes	2	24497585	NP_001809	AKR1C4	aldo-keto reductase family 1, member C4	1	1	cytoplasm	androgen metabolism bile acid transport	3-alpha-hydroxysteroid dehydrogenase (B-specific) activity aldo-keto reductase activity bile acid transporter activity chlordecone reductase activity electron carrier activity	
Yes	2	93277124	NP_00103526	AKR1CL2	aldo-keto reductase family 1, member C-like	1	1	not classified	not classified	oxidoreductase activity	
	2	41327764	NP_003680	AKR7A2	aldo-keto reductase family 7, member A2	2	5	not classified	aldehyde metabolism carbohydrate metabolism	aldehyde reductase activity electron carrier activity	
	2	41152114	NP_036199	AKR7A3	aldo-keto reductase family 7, member A3	2	7	cytosol	aldehyde metabolism	aldo-keto reductase activity electron carrier activity	
	2	51558759	NP_001003945	ALAD	delta-aminolevulinic acid dehydratase isoform a	1	1	not classified	heme biosynthesis	lyase activity metal ion binding porphobilinogen synthase activity zinc ion binding	Acute Hepatic Porphyria [MIM:125270]
	1,2	4502027	NP_000468	ALB	albumin precursor	36	139	extracellular region extracellular space protein complex	body fluid osmoregulation cellular response to starvation hemolysis of host red blood cells maintenance of mitochondrion localization negative regulation of apoptosis negative regulation of non-apoptotic programmed cell death transport water homeostasis	antioxidant activity carrier activity copper ion binding DNA binding drug binding fatty acid binding metal ion binding oxygen binding protein binding pyridoxal phosphate binding toxin binding water binding	Dysalbuminemic Hyperthyroxinemia Hyperthyroxinemia, dysalbuminemic Analbuminemia Bisalbuminemia [MIM:103600]
	2	68163411	NP_001618	ALCAM	activated leukocyte cell adhesion molecule	1	1	integral to plasma membrane membrane membrane fraction	antimicrobial humoral response (sensu Vertebrata) cell adhesion signal transduction	receptor binding	

Supplementary Table 1

	2	21361176	NP_000680	ALDH1A1	aldehyde dehydrogenase 1A1	10	15	cytoplasm	aldehyde metabolism metabolism	aldehyde dehydrogenase (NAD) activity androgen binding oxidoreductase activity Ras GTPase activator activity retinal dehydrogenase activity	
	2	21614513	NP_036322	ALDH1L1	aldehyde dehydrogenase 1 family, member	9	10	cytoplasm	10-formyltetrahydrofolate catabolism biosynthesis one-carbon compound metabolism protein biosynthesis	cofactor binding formyltetrahydrofolate dehydrogenase activity hydroxymethyl-, formyl- and related transferase activity methyltransferase activity oxidoreductase activity	
	2	25777732	NP_000681	ALDH2	mitochondrial aldehyde dehydrogenase 2 pr	1	1	mitochondrion	alcohol metabolism carbohydrate metabolism metabolism	aldehyde dehydrogenase (NAD) activity aldehyde dehydrogenase [NAD(P)+] activity electron carrier activity	
	2	4502043	NP_000685	ALDH3B1	aldehyde dehydrogenase 3B1 isoform a	1	1	not classified	alcohol metabolism aldehyde metabolism lipid metabolism metabolism	3-chloroalyl aldehyde dehydrogenase activity aldehyde dehydrogenase [NAD(P)+] activity oxidoreductase activity	
	2	71773324	NP_00102518	ALDH3B1	aldehyde dehydrogenase 3B1 isoform b	1	1	not classified	alcohol metabolism aldehyde metabolism lipid metabolism metabolism	3-chloroalyl aldehyde dehydrogenase activity aldehyde dehydrogenase [NAD(P)+] activity oxidoreductase activity	
Yes	2	73695879	NP_000686	ALDH3B2	aldehyde dehydrogenase 3B2	2	2	not classified	alcohol metabolism aldehyde metabolism lipid metabolism metabolism	3-chloroalyl aldehyde dehydrogenase activity aldehyde dehydrogenase [NAD(P)+] activity oxidoreductase activity	
	2	12007648	NP_072090	ALDH8A1	aldehyde dehydrogenase 8A1 isoform 1	1	1	intracellular	cell surface extracellular matrix (sensu Metazoa) membrane	oxidoreductase activity retinal dehydrogenase activity	
	2	25777739	NP_000687	ALDH9A1	aldehyde dehydrogenase 9A1	5	8	cytoplasm	aldehyde metabolism electron transport hormone metabolism metabolism neurotransmitter biosynthesis	4-trimethylammoniumbutyraldehyde dehydrogenase activity aldehyde dehydrogenase (NAD) activity aminobutyraldehyde dehydrogenase activity oxidoreductase activity	
	2	34577110	NP_908930	ALDOA	aldolase A	7	14	not classified	fructose metabolism glycolysis metabolism striated muscle contraction	fructose-bisphosphate aldolase activity lyase activity	Aldolase Deficiency of Red Cells Myopathy and Hemolytic Anemia [MIM:103850]
Yes	1,2	40354205	NP_000026	ALDOB	aldolase B	14	122	cytoplasm	fructose metabolism glycolysis metabolism	fructose-bisphosphate aldolase activity lyase activity	
	2	4885063	NP_005156	ALDOC	fructose-bisphosphate aldolase C	1	4	not classified	fructose metabolism glycolysis metabolism	fructose-bisphosphate aldolase activity lyase activity	

Supplementary Table 1

	2	29029632	NP_004295	ALK	anaplastic lymphoma kinase Ki-1	1	1	integral to plasma membrane	brain development nervous system development protein amino acid N-linked glycosylation protein amino acid phosphorylation transmembrane receptor protein tyrosine kinase signaling pathway	ATP binding nucleotide binding receptor activity receptor signaling protein tyrosine kinase activity transferase activity transmembrane receptor protein tyrosine kinase activity	
	2	154426292	NP_000688	ALOX12	arachidonate 12-lipoxygenase	1	1	cytosol sarcolemma	anti-apoptosis arachidonic acid metabolism cell motility electron transport fatty acid oxidation leukotriene biosynthesis negative regulation of cell volume oxygen and reactive oxygen species metabolism positive regulation of cell adhesion positive regulation of cell growth positive regulation of cell proliferation regulation of membrane potential superoxide release	arachidonate 12-lipoxygenase activity hepoxilin-epoxide hydrolase activity iron ion binding lipoxygenase activity metal ion binding oxidoreductase activity potassium channel inhibitor activity	
	2	13787193	NP_000469	ALPL	tissue non-specific alkaline phosphatase precursor	3	4	integral to membrane	metabolism ossification	alkaline phosphatase activity GPI anchor binding hydrolase activity magnesium ion binding zinc ion binding	Hypophostasia [MIM:241500]
Yes	2	13027388	NP_061041	ALS2CR2	amyotrophic lateral sclerosis 2 (juvenile) chr	1	1	not classified	protein amino acid phosphorylat	ATP binding nucleotide binding protein binding protein kinase activity transferase activity	
	1,2	4502067	NP_001624	AMBP	alpha-1-microglobulin/bikunin precursor	4	20	extracellular region plasma membrane	anti-inflammatory response cell adhesion heme catabolism negative regulation of immune response negative regulation of JNK cascade pregnancy protein-chromophore linkage transport	calcium channel inhibitor activity calcium oxalate binding heme binding IgA binding plasmin inhibitor activity protein homodimerization activity transporter activity trypsin inhibitor activity	
	2	110611172	NP_112205	AMN	amionless protein precursor	1	1	integral to membrane	development	not classified	Megaloblastic Anemia 1 [MIM:261100]
Yes	2	56549664	NP_001008220	AMY1C	salivary amylase alpha 1C precursor	5	7	not classified	not classified	not classified	
Yes	2	4502085	NP_000690	AMY2A	amylase, alpha 2A; pancreatic precursor	1	3	extracellular space	carbohydrate metabolism metabolism	alpha-amylase activity calcium ion binding catalytic activity cation binding chloride ion binding hydrolase activity, acting on glycosyl bonds	

Supplementary Table 1

Yes	2	10280622	NP_066188	AMY2B	amylase, pancreatic, alpha-2B precursor	4	6	extracellular space	carbohydrate metabolism digestion metabolism	alpha-amylase activity calcium ion binding chloride ion binding hydrolase activity, acting on glycosyl bonds	
Yes	2	4757752	NP_004664	ANGPTL1	angiotensin-like 1 precursor	1	1	extracellular space	not classified	receptor binding	
	2	6912236	NP_036230	ANGPTL2	angiotensin-like 2 precursor	1	1	extracellular space	development	receptor binding	
Yes	2	29893555	NP_114123	ANGPTL6	angiotensin-like 6	1	1	not classified	angiogenesis cell differentiation	not classified	
	2	31317212	NP_057460	ANKFY1	ankyrin repeat and FYVE domain containing	2	4	endosome membrane membrane	endocytosis	metal ion binding protein binding zinc ion binding	
	2	31317252	NP_065791	ANKFY1	ankyrin repeat and FYVE domain containing	2	2	endosome membrane membrane	endocytosis	metal ion binding protein binding zinc ion binding	
Yes	2	56676397	NP_037407	ANKRD11	ankyrin repeat domain 11	1	1	nucleus	not classified	not classified	
	1,2	4502095	NP_001141	ANPEP	membrane alanine aminopeptidase precursor	69	412	ER-Golgi intermediate compartment integral to plasma membrane	angiogenesis cell differentiation proteolysis	aminopeptidase activity membrane alanyl aminopeptidase activity metal ion binding metallopeptidase activity receptor activity zinc ion binding	Hypertension [MIM:151530]
	2	4502101	NP_000691	ANXA1	annexin I	19	168	cornified envelope cytoplasm	anti-apoptosis cell motility cell surface receptor linked signal transduction inflammatory response keratinocyte differentiation lipid metabolism peptide cross-linking	calcium ion binding calcium-dependent phospholipid binding phospholipase A2 inhibitor activity phospholipase inhibitor activity protein binding, bridging receptor binding structural molecule activity	
	2	22165431	NP_665875	ANXA11	annexin A11	31	167	cytoplasm nuclear envelope nucleoplasm nucleus	immune response	calcium ion binding calcium-dependent phospholipid binding protein binding	
	2	51896029	NP_00100395	ANXA13	annexin A13 isoform b	1	2	plasma membrane	cell differentiation	calcium ion binding calcium-dependent phospholipid binding	
	2	50845386	NP_00100285	ANXA2	annexin A2 isoform 2	29	204	plasma membrane soluble fraction	skeletal development	calcium ion binding calcium-dependent phospholipid binding cytoskeletal protein binding phospholipase inhibitor activity	
	2	50845388	NP_00100285	ANXA2	annexin A2 isoform 1	20	156	plasma membrane soluble fraction	skeletal development	calcium ion binding calcium-dependent phospholipid binding cytoskeletal protein binding phospholipase inhibitor activity	
	1,2	4826643	NP_005130	ANXA3	annexin A3	8	20	cytoplasm	signal transduction	calcium ion binding calcium-dependent phospholipid binding diphosphoinositol-polyphosphate diphosphatase activity phospholipase A2 inhibitor activity	
	1,2	4502105	NP_001144	ANXA4	annexin IV	24	84	cytoplasm	anti-apoptosis negative regulation of coagulation signal transduction	calcium ion binding calcium-dependent phospholipid binding phospholipase inhibitor activity	
	1,2	4502107	NP_001145	ANXA5	annexin 5	24	105	cytoplasm intracellular	anti-apoptosis blood coagulation negative regulation of coagulation signal transduction	calcium ion binding calcium-dependent phospholipid binding phospholipase inhibitor activity protein binding	

Supplementary Table 1

	1,2	71773329	NP_001146	ANXA6	annexin VI isoform 1	27	75	not classified	not classified	calcium ion binding calcium-dependent phospholipid binding protein binding	
	2	71773415	NP_004024	ANXA6	annexin VI isoform 2	20	66	not classified	not classified	calcium ion binding calcium-dependent phospholipid binding protein binding	
	2	4809279	NP_004025	ANXA7	annexin VII isoform 2	13	33	not classified	not classified	calcium ion binding calcium-dependent phospholipid binding voltage-gated calcium channel activity	
	2	71773480	NP_001150	AOX1	aldehyde oxidase 1	1	1	not classified	electron transport inflammatory response oxygen and reactive oxygen species metabolism	2 iron, 2 sulfur cluster binding aldehyde oxidase activity electron carrier activity iron ion binding metal ion binding molybdenum ion binding oxidoreductase activity xanthine dehydrogenase activity	
	2	14210504	NP_115882	AP1M1	adaptor-related protein complex 1, mu 1 subunit	1	1	clathrin vesicle coat coated pit	intracellular protein transport	protein binding	
Yes	2	4557469	NP_001273	AP2B1	adaptor-related protein complex 2, beta 1 subunit	1	2	clathrin coat of trans-Golgi network vesicle coated pit	intracellular protein transport protein complex assembly	binding protein binding	
Yes	2	71773106	NP_001025177	AP2B1	adaptor-related protein complex 2, beta 1 subunit	1	4	clathrin coat of trans-Golgi network vesicle coated pit	intracellular protein transport protein complex assembly	binding protein binding	
	2	14917109	NP_004059	AP2M1	adaptor-related protein complex 2, mu 1 subunit	1	1	clathrin vesicle coat coated pit	intracellular protein transport transport	lipid binding protein binding transporter activity	
	2	14917111	NP_004713	AP4M1	adaptor-related protein complex 4, mu 1 subunit	1	1	clathrin vesicle coat coated vesicle membrane Golgi trans cisterna membrane coat adaptor complex	intracellular protein transport transport vesicle-mediated transport	transporter activity	
	2	4502123	NP_001151	APAF1	apoptotic peptidase activating factor 1 isoform 1	1	2	cytosol intracellular	caspase activation via cytochrome c nervous system development regulation of apoptosis	ATP binding caspase activator activity nucleotide binding protein binding	
	1,2	4502133	NP_001630	APCS	serum amyloid P component precursor	5	12	extracellular region extracellular space	acute-phase response chaperone-mediated protein complex assembly protein folding	calcium ion binding sugar binding unfolded protein binding	
	2	4557321	NP_000030	APOA1	apolipoprotein A-I preproprotein	6	17	extracellular region extracellular space	cholesterol metabolism cholesterol transport circulation lipid metabolism lipid transport lipoprotein metabolism steroid metabolism	high-density lipoprotein binding lipid binding lipid transporter activity protein binding	Primary Hypoalphalipoproteinemia [MIM:604091]

Supplementary Table 1

	2	4502149	NP_001634	APOA2	apolipoprotein A-II preproprotein	1	1	extracellular region	glucose metabolism lipid transport negative regulation of lipid catabolism negative regulation of lipoprotein metabolism neutrophil activation positive regulation of interleukin-8 biosynthesis regulation of cytokine production response to glucose stimulus	lipid binding lipid transporter activity protein heterodimerization activity protein homodimerization activity	Apolipoprotein A-II Deficiency, Familial Hypercholesterolemia, Familial [MIM:143890]
	2	71773110	NP_000473	APOA4	apolipoprotein A-IV precursor	2	5	chylomicron extracellular region	circulation lipid metabolism lipid transport lipoprotein metabolism	lipid binding lipid transporter activity	
	1,2	4502163	NP_001638	APOD	apolipoprotein D precursor	7	151	extracellular region extracellular space	lipid metabolism transport	high-density lipoprotein binding lipid binding lipid transporter activity protein binding	
	2	4557325	NP_000032	APOE	apolipoprotein E precursor	6	32	chylomicron cytoplasm extracellular region	cholesterol homeostasis circulation cytoskeleton organization and biogenesis induction of apoptosis intracellular transport learning and/or memory lipid transport lipoprotein metabolism protein tetramerization regulation of axon extension regulation of ne	antioxidant activity apolipoprotein E receptor binding beta-amyloid binding heparin binding lipid transporter activity phospholipid binding tau protein binding	
	2	6912242	NP_036228	APPL	adaptor protein containing pH domain, PTB	1	1	cytoplasm endosome membrane membrane nucleus NuRD complex	cell cycle cell proliferation signal transduction	protein binding	
Yes	2	7657269	NP_055847	APRIN	androgen-induced prostate proliferative shu	1	2	membrane nucleus	carbohydrate metabolism negative regulation of cell proliferation	ATP binding binding DNA binding transferase activity, transferring hexosyl groups	
	1,2	4502171	NP_000476	APRT	adenine phosphoribosyltransferase isoform a	3	5	cytoplasm	adenine salvage nucleoside metabolism purine ribonucleoside salvage	adenine phosphoribosyltransferase activity AMP binding transferase activity, transferring glycosyl groups	2,8-Dihydroxyadenine urolithiasis [MIM:102600]
	2	71773201	NP_001025189	APRT	adenine phosphoribosyltransferase isoform b	3	10	cytoplasm	adenine salvage nucleoside metabolism purine ribonucleoside salvage	adenine phosphoribosyltransferase activity AMP binding transferase activity, transferring glycosyl groups	2,8-Dihydroxyadenine urolithiasis [MIM:102600]
	1,2	37694062	NP_932766	AQP1	aquaporin 1	3	35	integral to membrane integral to plasma membrane membrane outer membrane	excretion transport water transport	porin activity transporter activity water transporter activity	Aquaporin 1 Deficiency, Colton-Null [MIM:110450]

Supplementary Table 1

	1,2	4502179	NP_000477	AQP2	aquaporin 2	7	36	apical plasma membrane integral to membrane plasma membrane	excretion transport water transport	transporter activity water channel activity	Autosomal recessive nephrogenic diabetes insipidus, type 1 [MIM:222000] Autosomal dominant nephrogenic diabetes insipidus, type 1 [MIM:125800]
Yes	2	4502187	NP_001161	AQP7	aquaporin 7	1	1	integral to membrane integral to plasma membrane membrane outer membrane	excretion generation of precursor metabolites and energy glycerol transport transport water transport	glycerol channel activity porin activity transporter activity water channel activity	
Yes	1,2	4502203	NP_001650	ARF3	ADP-ribosylation factor 3	4	14	intracellular	ER to Golgi vesicle-mediated transport protein transport rRNA processing small GTPase mediated signal transduction	GTP binding nucleotide binding	
Yes	1,2	4502205	NP_001651	ARF4	ADP-ribosylation factor 4	2	4	intracellular	ER to Golgi vesicle-mediated transport protein transport rRNA processing small GTPase mediated signal transduction	enzyme activator activity GTP binding GTPase activity nucleotide binding	
Yes	1,2	4502209	NP_001653	ARF5	ADP-ribosylation factor 5	9	27	intracellular	ER to Golgi vesicle-mediated transport protein transport rRNA processing small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	1,2	4502211	NP_001654	ARF6	ADP-ribosylation factor 6	5	14	cell cortex endosome intracellular plasma membrane ruffle	cell adhesion cell motility cortical actin cytoskeleton organization and biogenesis ER to Golgi vesicle-mediated transport negative regulation of receptor mediated endocytosis positive regulation of actin filament polymerization protein transport regulati	GTP binding nucleotide binding protein binding	
Yes	2	51479145	NP_006412	ARFGEF1	brefeldin A-inhibited guanine nucleotide-exc	1	1	intracellular	exocytosis regulation of ARF protein signal transduction	ARF guanyl-nucleotide exchange factor activity binding	
	2	4757766	NP_004299	ARHGAP1	Rho GTPase activating protein 1	1	1	intracellular	cytoskeleton organization and biogenesis Rho protein signal transduction signal transduction	GTP binding protein binding Rho GTPase activator activity SH3/SH2 adaptor activity	
	2	4757768	NP_004300	ARHGDI	Rho GDP dissociation inhibitor (GDI) alpha	1	5	cytoplasm cytoskeleton	anti-apoptosis cell motility negative regulation of cell adhesion Rho protein signal transduction	GTPase activator activity protein binding Rho GDP-dissociation inhibitor activity	

Supplementary Table 1

	2	56676393	NP_001166	ARHGDI B	Rho GDP dissociation inhibitor (GDI) beta	1	1	cytoplasm cytoplasmic membrane-bound vesicle cytoskeleton	actin cytoskeleton organization and biogenesis cell motility development immune response negative regulation of cell adhesion Rho protein signal transduction	GTPase activator activity Rho GDP-dissociation inhibitor activity	
	2	7662088	NP_056128	ARHGEF12	Rho guanine nucleotide exchange factor (GEF)	1	3	intracellular	regulation of Rho protein signal transduction	GTPase activator activity protein binding Rho guanyl-nucleotide exchange factor activity	
	2	9506653	NP_061960	ARL15	ADP-ribosylation factor related protein 2	1	1	not classified	not classified	GTP binding	
	1,2	4757774	NP_004302	ARL3	ADP-ribosylation factor-like 3	2	4	intracellular	G-protein coupled receptor protein signaling pathway rRNA processing small GTPase mediated signal transduction	GTP binding nucleotide binding signal transducer activity	
	2	14149815	NP_115522	ARL6	ADP-ribosylation factor-like 6	3	4	cytoplasm intracellular membrane	rRNA processing small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	Bardet-Biedl Syndrome 3 [MIM:209900]
	2	20270343	NP_620150	ARL8A	ADP-ribosylation factor-like 10B	1	1	cytoplasm intracellular membrane midbody spindle midzone	chromosome segregation rRNA processing small GTPase mediated signal transduction	alpha-tubulin binding beta-tubulin binding GTP binding GTPase activity nucleotide binding	
	2	8922601	NP_060654	ARL8B	ADP-ribosylation factor-like 10C	1	3	cytoplasm intracellular membrane midbody spindle midzone	chromosome segregation rRNA processing small GTPase mediated signal transduction	alpha-tubulin binding beta-tubulin binding GDP binding GTP binding GTPase activity nucleotide binding	
	2	27436879	NP_775104	ARMC3	armadillo repeat containing 3	1	1	not classified	not classified	binding	
	2	156151430	NP_079415	ARMC9	armadillo repeat containing 9	1	1	not classified	not classified	binding	
	2	22907052	NP_006400	ARPC1A	actin related protein 2/3 complex subunit 1A	1	1	not classified	not classified	not classified	
	2	5031599	NP_005722	ARPC2	actin related protein 2/3 complex subunit 2	5	8	Arp2/3 protein complex cytoskeleton	cell motility regulation of actin filament polymerization	structural constituent of cytoskeleton	
	2	5031597	NP_005710	ARPC3	actin related protein 2/3 complex subunit 3	2	3	Arp2/3 protein complex	cell motility regulation of actin filament polymerization	structural constituent of cytoskeleton	
	1,2	5031595	NP_005709	ARPC4	actin related protein 2/3 complex subunit 4	2	6	Arp2/3 protein complex cytoskeleton	actin filament polymerization actin nucleation	actin filament binding protein binding, bridging structural constituent of cytoskeleton	
	2	68161509	NP_001020130	ARPC4	actin related protein 2/3 complex subunit 4	2	5	Arp2/3 protein complex cytoskeleton	actin filament polymerization actin nucleation	actin filament binding protein binding, bridging structural constituent of cytoskeleton	
	2	5031593	NP_005708	ARPC5	actin related protein 2/3 complex subunit 5	1	3	Arp2/3 protein complex cytoplasm cytoskeleton	actin cytoskeleton organization and biogenesis cell motility regulation of actin filament polymerization	structural constituent of cytoskeleton	
	2	13569956	NP_112240	ARPC5L	actin related protein 2/3 complex, subunit 5	1	2	cytoskeleton	regulation of actin filament polymerization	not classified	
	1,2	22748653	NP_689498	ARRDC1	arrestin domain containing 1	5	8	not classified	not classified	not classified	
	2	4502241	NP_000038	ARSE	arylsulfatase E precursor	1	2	not classified	metabolism skeletal development	arylsulfatase activity calcium ion binding hydrolase activity	Chondrodysplasia Punctata 1, X-linked Recessive [MIM:302950]

Supplementary Table 1

	2	31742482	NP_004033	ARSF	arylsulfatase F precursor	12	43	not classified	metabolism	arylsulfatase activity calcium ion binding hydrolase activity	
	1,2	30089928	NP_808592	ASAH1	N-acylsphingosine amidohydrolase (acid cerami	7	16	lysosome	carboxylic acid metabolism ceramide metabolism fatty acid metabolism	ceramidase activity hydrolase activity transferase activity, transferring acyl groups, acyl groups converted into alkyl on transfer	Farber Disease [MIM:228000]
	2	30089930	NP_004306	ASAH1	N-acylsphingosine amidohydrolase (acid cerami	9	34	lysosome	carboxylic acid metabolism ceramide metabolism fatty acid metabolism	ceramidase activity hydrolase activity transferase activity, transferring acyl groups, acyl groups converted into alkyl on transfer	Farber Disease [MIM:228000]
	2	68303549	NP_001020117	ASL	argininosuccinate lyase isoform 3	1	1	cytoplasm	amino acid biosynthesis arginine biosynthesis arginine catabolism urea cycle	argininosuccinate lyase activity catalytic activity lyase activity	Argoninosuccinic Aciduria [MIM:207900]
	2	50428938	NP_004308	ASNA1	arsA arsenite transporter, ATP-binding, hor	1	1	cytoplasm membrane nucleolus soluble fraction	anion transport response to arsenic transport	arsenite transporter activity arsenite-transporting ATPase activity ATP binding hydrolase activity nucleotide binding transporter activity	
	1,2	53759107	NP_000041	ASS1	argininosuccinate synthetase 1	20	59	cytoplasm	amino acid biosynthesis arginine biosynthesis urea cycle	argininosuccinate synthase activity ATP binding ligase activity nucleotide binding protein binding	Citrullinemia [MIM:215700]
	2	24497618	NP_054828	ATAD2	two AAA domain containing protein	1	1	not classified	not classified	ATP binding nucleoside-triphosphatase activity nucleotide binding	
	2	20127454	NP_004035	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	1	1	not classified	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	hydrolase activity IMP cyclohydrolase activity phosphoribosylaminoimidazolecarboxamide formyltransferase activity transferase activity	Aica-ribosiduria due to Atic Deficiency [MIM:608688]
Yes	2	21361181	NP_000692	ATP1A1	Na+/K+ -ATPase alpha 1 subunit isoform a	19	57	membrane membrane fraction sodium:potassium- exchanging ATPase complex	ATP hydrolysis coupled proton transport cation transport hydrogen ion homeostasis metabolism potassium ion transport sodium ion transport sperm motility	ATP binding ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism hydrolase activity hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances magnesium ion binding monovalent inorganic	

Supplementary Table 1

Yes	2	48762682	NP_001001586	ATP1A1	Na+/K+ -ATPase alpha 1 subunit isoform b	13	36	membrane membrane fraction sodium:potassium- exchanging ATPase complex	ATP hydrolysis coupled proton transport cation transport hydrogen ion homeostasis metabolism potassium ion transport sodium ion transport sperm motility	ATP binding ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism hydrolase activity hydrolase activity, acting on acid anhydrases, catalyzing transmembrane movement of substances magnesium ion binding monovalent inorganic cation transporter activity nucleotide binding potassium ion binding protien binding sodium ion binding sodium:potassium-exchanging ATPase activity	
	2	49574489	NP_001001787	ATP1B1	Na+/K+ -ATPase beta 1 subunit isoform b	1	9	integral to membrane membrane sodium:potassium- exchanging ATPase complex	ion transport potassium ion transport sodium ion transport	potassium ion binding sodium ion binding sodium:potassium-exchanging ATPase activity	
	2	50345984	NP_001001937	ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit precursor	3	5	membrane fraction mitochondrion proton-transporting ATP synthase complex (sensu Eukaryota) proton-transporting ATP synthase complex, catalytic core F(1) proton-transporting two- sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	ATP binding hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances meta	
	2	32189394	NP_001677	ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta subunit precursor	4	5	mitochondrion proton-transporting ATP synthase complex (sensu Eukaryota) proton-transporting ATP synthase complex, catalytic core F(1) proton-transporting ATP synthase, catalytic core (sensu Eukaryota)	ATP synthesis coupled proton transport generation of precursor metabolites and energy ion transport proton transport	ATP binding hydrogen-exporting ATPase activity, phosphorylative mechanism hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding nucleoside-triphos	
	2	17136148	NP_001174	ATP6AP1	ATPase, H+ transporting, lysosomal access	1	1	membrane proton-transportin two- sector ATPase complex vacuole	ATP synthesis coupled proton transport ion transport proton transport	ATP binding hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding nucleotide binding transporter activity	
	2	15011918	NP_005756	ATP6AP2	ATPase, H+ transporting, lysosomal access	1	5	integral to membrane membrane	not classified	receptor activity	
	2	19913418	NP_005168	ATP6V0A1	ATPase, H+ transporting, lysosomal V0 sub	1	1	integral to membrane membrane	ion transport proton transport	hydrogen ion transporter activity	

Supplementary Table 1

	2	85386053	NP_065683	ATP6V0A4	ATPase, H+ transporting, lysosomal V0 subunit	1	2	hydrogen-translocating V-type ATPase complex integral to membrane membrane	ion transport proton transport regulation of pH	hydrogen ion transporter activity	Renal Tubular Acidosis, Distal, Autosomal Recessive [MIM:602722]
	2	4502313	NP_001685	ATP6V0C	ATPase, H+ transporting, lysosomal, V0 subunit	1	1	membrane proton-transporting two-sector ATPase complex vacuole	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotation mechanism hydrolase activity metal ion binding	
	2	19913432	NP_004682	ATP6V0D1	ATPase, H+ transporting, lysosomal, V0 subunit	1	1	hydrogen-translocating V-type ATPase complex proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding	
	2	22749165	NP_689778	ATP6V0D2	ATPase, H+ transporting, lysosomal 38kDa	2	4	proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotation mechanism	
	2	19913424	NP_001681	ATP6V1A	ATPase, H+ transporting, lysosomal 70kDa	22	49	integral to plasma membrane proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	ATP binding hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding nucleotide binding	
	2	19913426	NP_001683	ATP6V1B1	ATPase, H+ transporting, lysosomal 56/58kDa	8	17	cytoplasm plasma membrane proton-transporting ATP synthase complex, catalytic core F(1) proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport energy coupled proton transport, against electrochemical gradient excretion ion transport regulation of pH sensory perception of sound	ATP binding hydrogen-exporting ATPase activity, phosphorylative mechanism hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding nucleotide binding	Renal Tubular Acidosis, Distal, with Progressive Deafness [MIM:267300]
	2	19913428	NP_001684	ATP6V1B2	vacuolar H+ATPase B2	12	27	cytoplasm proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport energy coupled proton transport, against electrochemical gradient ion transport	ATP binding hydrogen-exporting ATPase activity, phosphorylative mechanism hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding	
	2	4502315	NP_001686	ATP6V1C1	ATPase, H+ transporting, lysosomal 42kDa	1	1	proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	ATP binding hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances metal ion binding transporter activity	

Supplementary Table 1

	2	87159812	NP_001034451	ATP6V1C2	vacuolar H+ ATPase C2 isoform a	1	1	proton-transporting two-sector ATPase	ATP synthesis coupled proton transport	ATP binding hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	
	2	7706757	NP_057078	ATP6V1D	H(+)-transporting two-sector ATPase	2	3	proton-transporting two-sector ATPase	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding	
	2	4502317	NP_001687	ATP6V1E1	vacuolar H+ ATPase E1 isoform a	2	2	plasma membrane proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding protein binding	
	2	87159818	NP_001034456	ATP6V1E1	vacuolar H+ ATPase E1 isoform c	1	1	plasma membrane proton-transporting two-sector ATPase complex plasma membrane	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transportin ATPase activity, rotational mechanism hydrolase activity metal ion binding protein binding	
	2	87159816	NP_001034453	ATP6V1E1	vacuolar H+ ATPase E1 isoform b	1	1	plasma membrane proton-transporting two-sector ATPase complex plasma membrane	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transportin ATPase activity, rotational mechanism hydrolase activity metal ion binding protein binding	
	2	20357547	NP_004222	ATP6V1F	ATPase, H+ transporting, lysosomal 14kD, V1	1	1	membrane fraction proton-transporting two-sector ATPase complex	ATP synthesis coupled proton transport ion transport proton transport	hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding	
	2	4757818	NP_004879	ATP6V1G1	vacuolar H+ ATPase G1	1	2	not classified	ATP biosynthesis ion transport proton transport	hydrogen ion transporter activity hydrolase activity metal ion binding	
	2	47717100	NP_998784	ATP6V1H	ATPase, H+ transporting, lysosomal 50/57kD	1	34	hydrogen-transporting ATPase V1 domain peripheral to membrane of membrane fraction	ATP hydrolysis coupled proton transport ATP synthesis coupled proton transport endocytosis ion transport proton transport vacuolar acidification	ATP binding ATPase activity ATPase stimulator activity enzyme regulator activity hydrogen-transporting ATP synthase activity, rotational mechanism hydrogen-transporting ATPase activity, rotational mechanism hydrolase activity metal ion binding protein binding	
Yes	2	4502325	NP_001175	ATR	ataxia telangiectasia and Rad3 related protein	1	2	chromosome nucleus	cell cycle cell cycle checkpoint development DNA repair	binding DNA binding manganese ion binding phosphotransferase activity, alcohol group as acceptor protein serine/threonine kinase activity transferase activity	

Supplementary Table 1

	2	21450861	NP_647537	ATRN	attractin isoform 1	1	1	extracellular space integral to plasma membrane plasma membrane	development inflammatory response	receptor activity sugar binding	
	2	21450863	NP_647538	ATRN	attractin isoform 2	1	1	extracellular space integral to plasma membrane plasma membrane	development inflammatory response	receptor activity sugar binding	
	2	11342670	NP_001691	AZU1	azurocidin 1 preproprotein	1	1	azurophil granule extracellular region	anti-apoptosis cellular extravasation chemotaxis defense response to Gram- negative bacterium glial cell migration induction of positive chemotaxis macrophage chemotaxis microglial cell activation monocyte activation positive regulation of cell adhesion po	heparin binding serine-type endopeptidase activity toxin binding	
	2	4757826	NP_004039	B2M	beta-2-microglobulin precursor	1	1	extracellular region	antigen processing and presentation of endogenous antigen antigen processing and presentation of endogenous peptide antigen via MHC class I immune response	MHC class I receptor activity	Hypercatabolic Hypoproteinemia [MIM:241600]
	2	13929462	NP_001488	B4GALT1	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase 1, membrane-bound form	1	1	Golgi apparatus integral to membrane membrane	carbohydrate metabolism oligosaccharide biosynthesis	beta-N-acetylglucosaminylglycopeptide beta-1,4- galactosyltransferase activity galactosyltransferase activity lactose synthase activity manganese ion binding metal ion binding N-acetyllactosamine synthase activity transferase activity, transferring glycosy	Congenital Disorder of Glycosylation Type Iid [MIM:607091]
	1,2	9257197	NP_059344	BAIAP2	BAI1-associated protein 2 isoform 1	9	16	not classified	not classified	protein binding	
	2	9257199	NP_059345	BAIAP2	BAI1-associated protein 2 isoform 2	6	13	not classified	not classified	protein binding	
	1,2	32171238	NP_061330	BAIAP2L1	BAI1-associated protein 2-like 1	4	9	not classified	not classified	not classified	
	2	30795231	NP_006308	BASP1	brain abundant, membrane attached signal	6	20	cytoskeleton plasma membrane	not classified	not classified	
	2	18375626	NP_542417	BAT2	HLA-B associated transcript-2 isoform a	1	14	nucleus	not classified	protein binding	
	2	14670390	NP_075381	BAZ1B	bromodomain adjacent to zinc finger domain	1	1	nucleus	regulation of transcription, DNA- dependent transcription	metal ion binding protein binding transcription factor activity zinc ion binding	
	1,2	4502369	NP_003977	BBOX1	gamma-butyrobetaine dioxygenase	6	12	not classified	carnitine biosynthesis electron transport	gamma-butyrobetaine dioxygenase activity iron ion binding oxidoreductase activity oxidoreductase activity, acting on single donors with incorporation of molecular oxygen,incorporation of two atoms of oxygen	

Supplementary Table 1

	2	31543106	NP_005572	BCAM	basal cell adhesion molecule isoform 1 prec	1	1	cell surface integral to plasma membrane plasma membrane	cell adhesion signal transduction	transmembrane receptor activity	
	2	61742797	NP_001013275	BCAM	basal cell adhesion molecule isoform 2 prec	2	4	cell surface integral to plasma membrane plasma membrane	cell adhesion signal transduction	transmembrane receptor activity	
	2	20336332	NP_443074	BCL2L12	BCL2-like 12 isoform 2	1	1	not classified	apoptosis	not classified	
	2	82546843	NP_004318	BCR	breakpoint cluster region isoform 1	2	2	intracellular	intracellular signaling cascade protein amino acid phosphorylation regulation of Rho protein signal transduction signal transduction	GTPase activator activity kinase activity protein serine/threonine kinase activity Rho guanyl-nucleotide exchange factor activity transferase activity	
	2	82546845	NP_067585	BCR	breakpoint cluster region isoform 2	2	3	intracellular	intracellular signaling cascade protein amino acid phosphorylation regulation of Rho protein signal transduction signal transduction	GTPase activator activity kinase activity protein serine/threonine kinase activity Rho guanyl-nucleotide exchange factor activity transferase activity	
	2	66933014	NP_064524	BDH2	3-hydroxybutyrate dehydrogenase, type 2	5	10	cytoplasm mitochondrion	fatty acid beta-oxidation metabolism	3-hydroxybutyrate dehydrogenase activity NAD binding oxidoreductase activity	
	2	4502403	NP_001702	BGN	biglycan preproprotein	1	2	extracellular matrix (sensu Metazoa) transport vesicle	not classified	extracellular matrix structural constituent transferase activity	
	2	39930463	NP_085142	BHLHB9	basic helix-loop-helix domain containing, class 1	1	2	not classified	not classified	binding	
Yes	1,2	4502407	NP_001704	BHMT	betaine-homocysteine methyltransferase	12	37	not classified	protein amino acid methylation regulation of homocysteine metabolism	betaine-homocysteine S-methyltransferase activity homocysteine S-methyltransferase activity metal ion binding methyltransferase activity transferase activity zinc ion binding	
	2	13162290	NP_060084	BHMT2	betaine-homocysteine methyltransferase 2	1	1	not classified	not classified	homocysteine S-methyltransferase activity methyltransferase activity transferase activity	
	2	4557367	NP_000377	BLMH	bleomycin hydrolase	1	2	cytoplasm nucleus	proteolysis	aminopeptidase activity bleomycin hydrolase activity carboxypeptidase activity	
	2	33589854	NP_000703	BLVRA	biliverdin reductase A	1	1	not classified	electron transport heme catabolism metabolism	biliverdin reductase activity metal ion binding oxidoreductase activity zinc ion binding	
	2	4502419	NP_000704	BLVRB	biliverdin reductase B (flavin reductase (NADH))	1	4	not classified	cellular metabolism	biliverdin reductase activity coenzyme binding flavin reductase activity oxidoreductase activity	
	2	4502447	NP_001716	BPI	bactericidal/permeability-increasing protein	4	7	integral to plasma membrane membrane	defense response to bacterium immune response	Gram-negative bacterial binding lipid binding	

Supplementary Table 1

	1,2	4757876	NP_004326	BST2	bone marrow stromal cell antigen 2	1	2	integral to plasma membrane membrane	cell proliferation cell-cell signaling development humoral immune response positive regulation of I-kappaB kinase/NF-kappaB cascade defense response to bacterium	signal transducer activity	
	2	5802988	NP_006754	BTG2	B-cell translocation gene 2	1	3	not classified	DNA repair negative regulation of cell proliferation regulation of transcription, DNA-dependent transcription	transcription factor activity	
	2	5802976	NP_006820	C10orf116	adipose specific 2	2	7	not classified	not classified	not classified	
	2	51468120	XP_374765	C10orf18	PREDICTED: similar to retinoblastoma-assc	1	1	not classified	not classified	not classified	
	2	22749471	NP_689964	C10orf30	hypothetical protein LOC222389	1	1	not classified	not classified	not classified	
	2	42761479	NP_775860	C11orf47	hypothetical protein LOC283294	1	1	not classified	not classified	not classified	
	2	18087827	NP_542390	C11orf52	hypothetical protein LOC91894	1	1	not classified	not classified	not classified	
	2	21361495	NP_054758	C11orf54	hypothetical protein LOC28970	1	3	nucleus	not classified	hydrolase activity metal ion binding zinc ion binding	
	1,2	8923579	NP_060377	C11orf59	hypothetical protein LOC55004	3	5	not classified	not classified	not classified	
	2	27902289	NP_776174	C12orf51	hypothetical protein LOC283450	1	1	not classified	not classified	not classified	
	2	23308507	NP_694567	C12orf59	hypothetical protein LOC120939	1	1	not classified	not classified	not classified	
	2	8392875	NP_037374	C16orf80	transcription factor IIB	1	1	not classified	development	not classified	
	2	22749499	NP_689979	C17orf61	hypothetical protein LOC254863	2	10	not classified	not classified	not classified	
	2	21361664	NP_060411	C17orf80	lung cancer-related protein 8	1	3	not classified	not classified	not classified	
	2	22748999	NP_689687	C19orf18	hypothetical protein LOC147685	1	1	integral to membrane membrane	not classified	not classified	
	2	13027608	NP_076427	C1orf116	specifically androgen-regulated protein	4	6	not classified	not classified	not classified	
	2	28603818	NP_788954	C1orf179	hypothetical protein LOC338094	4	13	not classified	not classified	not classified	
	2	20149643	NP_060423	C1orf181	hypothetical protein LOC54680	1	1	not classified	not classified	not classified	
	2	68563515	NP_001020402	C1orf45	keratinocyte expressed, proline-rich protein	1	1	not classified	not classified	not classified	
	1,2	21389601	NP_653296	C1orf58	hypothetical protein LOC148362	17	59	not classified	not classified	not classified	
	2	22748799	NP_689584	C1orf93	hypothetical protein LOC127281	1	1	not classified	not classified	not classified	
	2	71043622	NP_001012989	C20orf106	hypothetical protein LOC200232	1	1	not classified	not classified	not classified	
	2	40807482	NP_149974	C20orf114	LPLUNC1 protein precursor	10	23	not classified	not classified	lipid binding	
	2	66773336	NP_542194	C20orf117	hypothetical protein LOC140710 isoform 1	1	1	not classified	not classified	catalytic activity	
	2	60685214	NP_115642	C2orf16	hypothetical protein LOC84226	1	1	not classified	not classified	not classified	
	2	31542711	NP_060347	C2orf18	hypothetical protein LOC54978	2	6	integral to membrane membrane	not classified	not classified	
Yes	1,2	4557385	NP_000055	C3	complement component 3 precursor	1	2	extracellular region	complement activation, alternative pathway complement activation, classical pathway G-protein coupled receptor protein signaling pathway inflammatory response innate immune response signal transduction	endopeptidase inhibitor activity receptor binding	
Yes	2	67190748	NP_009224	C4A	complement component 4A preproprotein	1	1	extracellular region	complement activation, complement activation, classical pathway inflammatory response innate immune response	endopeptidase inhibitor activity	

Supplementary Table 1

Yes	2	50345296	NP_001002029	C4B	complement component 4B preproprotein	1	1	extracellular region	complement activation complement activation, classical pathway inflammatory response innate immune response	endopeptidase inhibitor activity	
	2	38016947	NP_001726	C5	complement component 5	1	1	extracellular region extracellular space membrane attack complex	activation of MAPK activity chemotaxis complement activation, alternative pathway complement activation, classical pathway cytolysis G-protein coupled receptor protein signaling pathway inflammatory response innate immune response response to stress	chemokine activity endopeptidase inhibitor activity receptor binding	
	1,2	14165278	NP_115788	C5orf32	putative nuclear protein ORF1-FL49	2	28	nucleus	not classified	not classified	
	2	40354201	NP_954653	C6orf108	putative c-Myc-responsive isoform 2	1	1	nucleus	cell proliferation	not classified	
	1,2	21361741	NP_057569	C6orf55 (VTA1)	hypothetical protein LOC51534	6	12	not classified	not classified	not classified	
	2	58219024	NP_001010903	C6orf58	hypothetical protein LOC352999	1	1	not classified	not classified	not classified	
	2	13129018	NP_076956	C7orf24	hypothetical protein LOC79017	1	1	not classified	not classified	not classified	
	2	4502511	NP_001728	C9	complement component 9	1	1	integral to plasma membrane membrane membrane attack complex	complement activation, alternative pathway complement activation, classical pathway cytolysis hemolysis of host red blood cells innate immune response	not classified	
	1,2	11641247	NP_071738	C9orf19	chromosome 9 open reading frame 19	2	6	extracellular region membrane	not classified	not classified	
	2	8923251	NP_060208	C9orf39	hypothetical protein LOC54875	1	1	membrane	signal transduction	two-component sensor activity	
	2	13899241	NP_113614	C9orf58	ionized calcium binding adapter molecule 2	2	3	actin cytoskeleton focal adhesion	not classified	calcium ion binding	
	2	79750824	NP_00103061	C9orf88	hypothetical protein LOC64855 isoform 2	2	3	not classified	not classified	not classified	
	1,2	4557395	NP_000058	CA2	carbonic anhydrase II	9	25	cytoplasm	one-carbon compound metabolism	carbonate dehydratase activity lyase activity metal ion binding zinc ion binding	Autosomal recessive syndrome of osteopetrosis with renal tubular acidosis [MIM:259730]
	1,2	4502519	NP_000708	CA4	carbonic anhydrase IV precursor	2	2	membrane membrane fraction	one-carbon compound metabolism response to stimulus visual perception	carbonate dehydratase activity GPI anchor binding lyase activity metal ion binding zinc ion binding	Proximal renal tubular acidosis [MIM:114760]
	1,2	7706481	NP_057373	CAB39	calcium binding protein 39	9	21	not classified	not classified	protein binding	
	2	72534734	NP_001026894	CAB39L	calcium binding protein 39-like	1	1	not classified	not classified	binding	
	2	54112390	NP_000713	CACNA2D1	calcium channel, voltage-dependent, alpha 2D1	21	1	integral to membrane membrane membrane fraction voltage-gated calcium channel complex	calcium ion transport ion transport	calcium channel activity calcium ion binding dihydropyridine-sensitive calcium channel activity receptor activity voltage-gated ion channel activity	
	1,2	4826655	NP_004920	CALB1	calbindin 1	2	3	not classified	not classified	calcium ion binding protein binding vitamin D binding	
Yes	2	4502543	NP_001731	CALB2	calbindin 2 full length protein isoform	1	1	not classified	not classified	calcium ion binding	
Yes	2	58218968	NP_005175	CALM3	calmodulin 3	1	1	not classified	not classified	calcium ion binding	

Supplementary Table 1

	2	4885111	NP_005176	CALML3	calmodulin-like 3	1	1	not classified	not classified	calcium ion binding	
								nucleus	protein amino acid phosphorylation signal transduction	ATP binding calcium ion binding calcium- and calmodulin-dependent protein kinase activity calmodulin binding nucleotide binding protein serine/threonine kinase activity transferase activity	
	2	4502557	NP_001735	CAMK4	calcium/calmodulin-dependent protein kinase 4	1	1				
	2	39753970	NP_004336	CAMP	cathelicidin antimicrobial peptide	1	3	extracellular region	defense response to bacterium	not classified	
								nucleus ubiquitin ligase complex	negative regulation of enzyme activity protein ubiquitination regulation of transcription, DNA-dependent transcription	protein binding	
	2	21361794	NP_060918	CAND1	TIP120 protein	1	2				
								membrane	adenylate cyclase activation establishment and/or maintenance of cell polarity signal transduction	actin binding	
	2	5453595	NP_006358	CAP1	adenylyl cyclase-associated protein	1	1				
								F-actin capping protein complex nucleus	barbed-end actin filament capping protein complex assembly	actin binding	
	2	63252913	NP_001738	CAPG	gelsolin-like capping protein	8	20				
								intracellular	positive regulation of cell proliferation proteolysis	calcium ion binding calpain activity	
	2	12408656	NP_005177	CAPN1	calpain 1, large subunit	15	26				
								intracellular	proteolysis	calcium ion binding calpain activity	
	2	4502563	NP_001739	CAPN2	calpain 2, large subunit	1	1				
								intracellular	proteolysis signal transduction	calpain activity	
	2	37577157	NP_004046	CAPN5	calpain 5	10	20				
								intracellular nucleus	proteolysis	binding calpain activity cysteine-type peptidase activity	
	1,2	7656959	NP_055111	CAPN7	calpain 7	30	145				
								F-actin capping protein complex	actin cytoskeleton organization and biogenesis barbed-end actin filament capping cell motility protein complex assembly	actin binding	
	2	5453597	NP_006126	CAPZA1	F-actin capping protein alpha-1 subunit	2	2				
								F-actin capping protein complex	actin cytoskeleton organization and biogenesis barbed-end actin filament capping cell motility protein complex assembly	actin binding	
	2	5453599	NP_006127	CAPZA2	capping protein (actin filament) muscle Z-line	1	1				
								F-actin capping protein complex	actin cytoskeleton organization and biogenesis barbed-end actin filament capping cell motility	actin binding	
	2	4826659	NP_004921	CAPZB	F-actin capping protein beta subunit	2	3				
								cytosol	metabolism	15-hydroxyprostaglandin dehydrogenase (NADP+) activity carbonyl reductase (NADPH) activity oxidoreductase activity prostaglandin-E2 9-reductase activity	
	2	4502599	NP_001748	CBR1	carbonyl reductase 1	5	8				
Yes	2	4502601	NP_001227	CBR3	carbonyl reductase 3	2	5	cytosol	metabolism	carbonyl reductase (NADPH) activity oxidoreductase activity	

Supplementary Table 1

								nucleus	positive regulation of I-kappaB kinase/NF-kappaB cascade regulation of transcription, DNA-dependent transcription	DNA binding signal transducer activity	Mental Retardation Autosomal Recessive 3 [MIM:608443]
2	24431973	NP_060191	CC2D1A	coiled-coil and C2 domain containing 1A	6	6					
2	27735063	NP_775753	CCDC105	coiled-coil domain containing 105	2	2	not classified	not classified	not classified		
2	5453603	NP_006422	CCT2	chaperonin containing TCP1, subunit 2	1	1	cytosol	protein folding regulation of progression through cell cycle	ATP binding nucleotide binding unfolded protein binding		
2	58761484	NP_001008800	CCT3	chaperonin containing TCP1, subunit 3 isoform 1	1	2	cytoskeleton	protein folding	ATP binding nucleotide binding unfolded protein binding		
2	63162572	NP_005989	CCT3	chaperonin containing TCP1, subunit 3 isoform 3	3	6	cytoskeleton	protein folding	ATP binding nucleotide binding unfolded protein binding		
2	24307939	NP_036205	CCT5	chaperonin containing TCP1, subunit 5 (epsilon)	3	8	not classified	protein folding	ATP binding nucleotide binding unfolded protein binding		
2	58331171	NP_001009186	CCT6A	chaperonin containing TCP1, subunit 6A isoform 2	2	2	cytoplasm	protein folding	ATP binding nucleotide binding unfolded protein binding		
2	4502643	NP_001753	CCT6A	chaperonin containing TCP1, subunit 6A isoform 1	1	3	cytoplasm	protein folding	ATP binding nucleotide binding unfolded protein binding		
2	48762932	NP_006576	CCT8	chaperonin containing TCP1, subunit 8 (theta)	4	4	cytosol	protein folding	ATP binding ATPase activity, coupled nucleotide binding unfolded protein binding		
2	4557417	NP_000582	CD14	CD14 antigen precursor	7	21	extrinsic to membrane plasma membrane	apoptosis cell surface receptor linked signal transduction inflammatory response phagocytosis	GPI anchor binding peptidoglycan receptor activity transferase activity		
2	11321634	NP_036252	CD2AP	CD2-associated protein	14	21	actin cytoskeleton cytoplasm filamentous actin ruffle	protein complex assembly signal transduction substrate-bound cell migration, cell extension	protein binding SH3 domain binding structural constituent of cytoskeleton	Focal Segmental Glomerulosclerosis FSGS3 [MIM:607832]	
2	42761474	NP_976074	CD59	CD59 antigen p18-20	4	20	membrane fraction plasma membrane	blood coagulation cell surface receptor linked signal transduction immune response	GPI anchor binding protein binding		
1,2	4502679	NP_001771	CD63	CD63 antigen isoform A	1	5	endosome membrane integral to membrane integral to plasma membrane lysosomal membrane plasma membrane	not classified	not classified		
2	91199546	NP_001035123	CD63	CD63 antigen isoform B	2	50	endosome membrane integral to membrane integral to plasma membrane lysosomal membrane plasma membrane	not classified	not classified		

Supplementary Table 1

								integral to plasma membrane membrane	activation of MAPK activity entry of virus into host cell phosphatidylinositol biosynthesis phosphoinositide metabolism positive regulation of 1-phosphatidylinositol 4-kinase activity positive regulation of B cell proliferation positive regulation of cel	protein binding	
	1,2	4757944	NP_004347	CD81	CD81 antigen	4	18				
								integral to membrane integral to plasma membrane plasma membrane	cell adhesion cell motility fusion of sperm to egg plasma membrane paranodal junction assembly platelet activation	protein binding	
	1,2	4502693	NP_001760	CD9	CD9 antigen	7	67				
								midbody nucleus spindle microtubule	anti-apoptosis cell cycle cell division mitosis protein amino acid phosphorylation regulation of progression through cell cycle traversing start control point of mitotic cell cycle	ATP binding cyclin-dependent protein kinase activity nucleotide binding protein binding transferase activity	
	2	16306492	NP_203698	CDC2	cell division cycle 2 protein isoform 2	1	1				
								cytoplasm filopodium intracellular plasma membrane	actin cytoskeleton organization and biogenesis cell cycle cell division establishment and/or maintenance of cell polarity G1 phase macrophage differentiation negative regulation of protein complex assembly positive regulation of pseudopodium formation reg	GTP binding GTPase activity nucleotide binding protein binding	
Yes	1,2	4757952	NP_001782	CDC42	cell division cycle 42 isoform 1	3	10				
								cytoplasm filopodium intracellular plasma membrane	actin cytoskeleton organization and biogenesis cell cycle cell division establishment and/or maintenance of cell polarity G1 phase macrophage differentiation negative regulation of protein complex assembly positive regulation of pseudopodium formation reg	GTP binding GTPase activity nucleotide binding protein binding	
	2	16357472	NP_426359	CDC42	cell division cycle 42 isoform 2	7	15				

Supplementary Table 1

Yes	2	89903012	NP_00103489	CDC42	cell division cycle 42 isoform 1	7	27	cytoplasm filopodium intracellular plasma membrane	actin cytoskeleton organization and biogenesis cell cycle cell division establishment and/or maintenance of cell polarity G1 phase macrophase differentiation negative regulation of protein complex assembly positive regulation of pseudopodium formation regulation of cell growth regulation of progression through cell cycle small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	2	30089960	NP_055641	CDC42BPA	CDC42-binding protein kinase alpha isoform 2	2	3	intercellular junction leading edge	actin cytoskeleton reorganization intracellular signaling cascade protein amino acid phosphorylation regulation of small GTPase mediated signal transduction	ATP binding diacylglycerol binding identical protein binding kinase activity magnesium ion binding metal ion binding nucleotide binding protein serine/threonine kinase activity small GTPase regulator activity transferase activity zinc ion binding	
	2	16357474	NP_006026	CDC42BPB	CDC42-binding protein kinase beta	3	4	cytoskeleton intercellular junction leading edge	actin cytoskeleton reorganization cytoskeleton organization and biogenesis establishment and/or maintenance of cell polarity intracellular signaling cascade protein amino acid phosphorylation regulation of small GTPase mediated signal transduction	ATP binding diacylglycerol binding identical protein binding magnesium ion binding nucleotide binding protein serine/threonine kinase activity small GTPase regulator activity transferase activity zinc ion binding	
Yes	2	4826675	NP_004926	CDK5	cyclin-dependent kinase 5	1	4	axon cell soma cytoplasm dendrite growth cone membrane neuromuscular junction nucleus	axon extension cell cycle cell division cell proliferation embryonic development neurite development neuron differentiation positive regulation of neuron apoptosis protein amino acid phosphorylation striated muscle development	acetylcholine receptor activator activity ATP binding cyclin-dependent protein kinase activity ErbB-2 class receptor binding ErbB-3 class receptor binding nucleotide binding protein binding tau-protein kinase activity transferase activity	
	2	58535453	NP_001011649	CDK5RAP2	CDK5 regulatory subunit associated protein 1	1	1	centrosome cytoskeleton	brain development regulation of neuron differentiation	microtubule binding neuronal Cdc2-like kinase binding	
	2	11386171	NP_004354	CEACAM5	carcinoembryonic antigen-related cell adhesion molecule 5	1	1	integral to membrane integral to plasma membrane membrane	not classified	GPI anchor binding	

Supplementary Table 1

	2	21735548	NP_009117	CEP250	centrosomal protein 2 isoform 1	1	1	centrosome	mitotic cell cycle regulation of centriole-centriole cohesion	protein kinase binding	
	2	4557443	NP_000069	CETP	cholesteryl ester transfer protein, plasma precursor	7	14	not classified	cholesterol metabolism lipid metabolism lipid transport steroid metabolism	catalytic activity lipid binding	Cholesterol Ester Transfer Protein Deficiency [MIM:607322]
	2	42544239	NP_001919	CFD	complement factor D preproprotein	2	2	not classified	complement activation, alternative pathway innate immune response proteolysis	complement factor D activity peptidase activity	
	2	62739188	NP_001014975	CFH	complement factor H isoform b precursor	1	1	extracellular space	complement activation, alternative pathway innate immune response	not classified	Hemolytic Uremic Syndrome, Atypical [MIM: 235400]
	2	119392081	NP_000195	CFI	complement factor I	1	1	extracellular region membrane	complement activation, classical pathway innate immune response proteolysis	complement factor I activity peptidase activity scavenger receptor activity	Complement Factor I Deficiency [MIM: 610984]
Yes	1,2	5031635	NP_005498	CFL1	cofilin 1 (non-muscle)	5	13	cytoplasm cytoskeleton intracellular nucleus	actin cytoskeleton organization and biogenesis anti-apoptosis Rho protein signal transduction	actin binding protein binding	
Yes	2	14719392	NP_068733	CFL2	cofilin 2	2	8	cytoskeleton intracellular nucleus	not classified	actin binding protein binding	
Yes	2	33946278	NP_619579	CFL2	cofilin 2	3	11	cytoskeleton intracellular nucleus	not classified	actin binding protein binding	
	2	8923390	NP_060282	CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing	1	2	mitochondrion	not classified	protein binding	
	2	103485496	NP_002759	CHMP1A	chromatin modifying protein 1A isoform 2	1	3	condensed nuclear chromosome early endosome endomembrane system microtubule organizing center nuclear matrix nucleus	cell cycle gene silencing mitotic chromosome condensation negative regulation of S phase of mitotic cell cycle negative regulation of transcription by glucose negative regulation of transcription, DNA-dependent	transcription protein transport vesicle-mediated transport	protein binding
	1,2	31542306	NP_065145	CHMP1B	chromatin modifying protein 1B	6	45	not classified	protein transport	not classified	
	1,2	7656922	NP_055268	CHMP2A	chromatin modifying protein 2A	11	73	not classified	protein transport	not classified	
	1,2	40254866	NP_054762	CHMP2B	chromatin modifying protein 2B	4	15	intracellular	protein transport	not classified	Frontotemporal Dementia, Chromosome 3-linked [MIM:6000795]
	1,2	28827795	NP_789782	CHMP4B	chromatin modifying protein 4B	2	6	not classified	protein transport	not classified	
	1,2	20127558	NP_057494	CHMP5	chromatin modifying protein 5	2	7	not classified	protein transport	not classified	
	2	31542673	NP_078867	CHMP6	chromatin modifying protein 6	2	3	membrane	protein transport	not classified	
	2	6005731	NP_009167	CHP	calcium binding protein P22	1	1	not classified	potassium ion transport small GTPase mediated signal transduction	calcium ion binding potassium channel regulator activity	

Supplementary Table 1

	2	5454062	NP_006375	CIB1	calcium and integrin binding 1 (calmyrin)	2	3	cytoplasm endoplasmic reticulum membrane nucleoplasm	apoptosis cell adhesion double-strand break repair response to DNA damage stimulus	calcium ion binding protein binding	
	2	19920317	NP_006816	CKAP4	cytoskeleton-associated protein 4	1	3	integral to membrane membrane fraction	not classified	not classified	
	2	21536286	NP_001814	CKB	brain creatine kinase	8	28	cytoplasm	not classified	creatine kinase activity	
Yes	2	4502877	NP_001296	CLDN4	claudin 4	1	1	integral to plasma membrane plasma membrane tight junction	calcium-independent cell-cell adhesion pathogenesis	identical protein binding structural molecule activity transmembrane receptor activity	
	1,2	14251209	NP_001279	CLIC1	chloride intracellular channel 1	7	19	brush border cytoplasm membrane membrane fraction nuclear envelope nucleus soluble fraction	chloride transport ion transport signal transduction	chloride ion binding protein binding voltage-gated chloride channel activity	
	2	40288290	NP_004660	CLIC3	chloride intracellular channel 3	2	3	cytoplasm membrane nucleus	chloride transport ion transport signal transduction	chloride ion binding protein binding voltage-gated chloride channel activity	
	1,2	7330335	NP_039234	CLIC4	chloride intracellular channel 4	5	6	actin cytoskeleton cytoplasm intracellular membrane microvillus soluble fraction	cell differentiation chloride transport ion transport negative regulation of cell migration	chloride ion binding protein binding voltage-gated chloride channel activity	
	2	82617632	NP_058625	CLIC5	chloride intracellular channel 5	1	2	actin cytoskeleton insoluble fraction membrane microvillus	chloride transport ion transport pregnancy	chloride ion binding protein binding voltage-gated chloride channel activity	
	2	27894378	NP_444507	CLIC6	chloride intracellular channel 6	2	3	membrane	chloride transport ion transport	chloride ion binding voltage-gated chloride channel activity	
	2	4758012	NP_004850	CLTC	clathrin heavy chain 1	12	24	clathrin coat of coated pit clathrin vesicle coat	intracellular protein transport	protein binding structural molecule activity	Renal Cell Carcinoma [MIM:118955]
Yes	2	9257202	NP_009029	CLTCL1	clathrin, heavy polypeptide-like 1	4	6	clathrin coat of coated pit clathrin vesicle coat	morphogenesis receptor mediated endocytosis	binding signal transducer activity	
	1,2	42716297	NP_001822	CLU	clusterin isoform 1	6	13	extracellular space	apoptosis cell death complement activation, classical pathway innate immune response lipid metabolism	not classified	
	2	42740907	NP_976084	CLU	clusterin isoform 2	5	17	extracellular space	apoptosis cell death complement activation, classical pathway innate immune response lipid metabolism	not classified	
	2	20270371	NP_620164	CMBL	carboxymethylenebutenolidase-like (Pseud)	6	13	not classified	not classified	hydrolase activity	

Supplementary Table I

	2	7706497	NP_057392	CMPK	cytidylate kinase	3	4	cytoplasm nucleus	nucleobase, nucleoside, nucleotide and nucleic acid metabolism pyrimidine ribonucleotide biosynthesis	ATP binding cytidylate kinase activity kinase activity nucleotide binding phosphotransferase activity, phosphate group as acceptor transferase activity uridine kinase activity	
	2	8922699	NP_060705	CNDP2	CNDP dipeptidase 2 (metallopeptidase M20	15	33	not classified	proteolysis	carboxypeptidase activity hydrolase activity metallopeptidase activity protein dimerization activity	
	2	41393057	NP_055742	CNKSR2	connector enhancer of kinase suppressor of	2	1	cytoplasm membrane	regulation of signal transduction	protein binding	
	1,2	38570091	NP_149124	CNP	2',3'-cyclic nucleotide 3' phosphodiesterase	13	27	membrane	cyclic nucleotide catabolism synaptic transmission	2',3'-cyclic-nucleotide 3'-phosphodiesterase activity hydrolase activity	
	2	46048207	NP_079509	COASY	coenzyme A synthase isoform a	1	1	not classified	biosynthesis coenzyme A biosynthesis	ATP binding dephospho-CoA kinase activity nucleotide binding nucleotidyltransferase activity pantetheine-phosphate adenylyltransferase activity transferase activity	
	2	62821787	NP_055715	COBLL1	COBL-like 1	1	3	not classified	not classified	not classified	
	2	18641350	NP_001846	COL15A1	alpha 1 type XV collagen precursor	1	3	collagen type XV cytoplasm extracellular matrix (sensu Metazoa)	Protein binding Structural Molecule Activity	Angiogenesis Cell Adhesion Cell Differentiation Phosphate Transport	
	2	13385620	NP_085059	COL18A1	alpha 1 type XVIII collagen isoform 1 precursor	1	1	collagen cytoplasm extracellular matrix extracellular matrix (sensu Metazoa)	cell adhesion negative regulation of cell proliferation organ morphogenesis phosphate transport visual perception	extracellular matrix structural constituent metal ion binding protein binding structural molecule activity zinc ion binding	Knobloch Syndrome [MIM:267750]
	1,2	87196339	NP_001839	COL6A1	collagen, type VI, alpha 1 precursor	6	21	collagen collagen type VI cytoplasm	cell adhesion phosphate transport	extracellular matrix structural constituent protein binding	Bethlem Myopathy [MIM:158810] Ullrich Congenital Muscular Dystrophy, Autosomal Dominant [MIM:254090]
	2	55743106	NP_476508	COL6A3	alpha 3 type VI collagen isoform 5 precursor	1	2	collagen type VI cytoplasm extracellular matrix (sensu Metazoa)	cell adhesion muscle development phosphate transport	protein binding serine-type endopeptidase inhibitor activity structural molecule activity	Ullrich Congenital Muscular Dystrophy [MIM: 254090]
	2	4502969	NP_000745	COMT	catechol-O-methyltransferase isoform MB-C	1	1	integral to membrane membrane microsome soluble fraction	catecholamine metabolism neurotransmitter catabolism	catechol O-methyltransferase activity magnesium ion binding protein binding transferase activity	
	2	6466450	NP_009294	COMT	catechol-O-methyltransferase isoform S-CO	1	1	integral to membrane membrane microsome soluble fraction	catecholamine metabolism neurotransmitter catabolism	catechol O-methyltransferase activity magnesium ion binding protein binding transferase activity	
	2	38027914	NP_937832	COPS8	COP9 signalosome subunit 8 isoform 2	1	1	nucleus signalosome complex	metabolism	racemase and epimerase activity, acting on amino acids and derivatives	
	2	21624607	NP_066972	COTL1	coactosin-like 1	1	1	cytoskeleton intracellular	not classified	actin binding enzyme binding	

Supplementary Table I

	2	4557485	NP_000087	CP	ceruloplasmin precursor	6	15	extracellular region extracellular space	copper ion homeostasis copper ion transport ion transport iron ion homeostasis	copper ion binding copper ion transporter activity ferroxidase activity metal ion binding oxidoreductase activity	Aceruloplasminemia [MIM:604290]
Yes	2	53832021	NP_001005502	CPM	carboxypeptidase M precursor	1	3	plasma membrane	aromatic compound metabolism morphogenesis proteolysis	carboxypeptidase A activity ferric iron binding GPI anchor binding metal ion binding metallopeptidase activity zinc ion binding	
	2	88971372	XP_947165	CPN2	PREDICTED: similar to Carboxypeptidase	2	1	not classified	not classified	not classified	
Yes	2	23397708	NP_690908	CPNE1	copine I	1	2	not classified	lipid metabolism vesicle-mediated transport	calcium-dependent phospholipid binding phosphatidylserine binding transporter activity	
Yes	2	30181236	NP_689940	CPNE2	copine II	1	1	not classified	not classified	not classified	
	1,2	4503015	NP_003900	CPNE3	copine III	14	33	cytoplasm cytosol	lipid metabolism vesicle-mediated transport	calcium-dependent phospholipid binding kinase activity protein serine/threonine kinase activity transferase activity transporter activity	
	1,2	25141323	NP_065990	CPNE5	copine V	3	5	not classified	not classified	not classified	
Yes	2	5174421	NP_006023	CPNE6	copine 6	1	1	not classified	lipid metabolism nervous system development synaptic transmission vesicle-mediated transport	calcium ion binding transporter activity	
Yes	1,2	25121972	NP_705898	CPNE8	copine VIII	16	37	not classified	not classified	not classified	
Yes	2	88703066	NP_705899	CPNE9	copine-like protein	5	27	not classified	not classified	not classified	
	2	83641874	NP_112601	CPVL	serine carboxypeptidase vitellogenic-like	6	12	not classified	proteolysis	peptidase activity serine carboxypeptidase activity	
	1,2	86793036	NP_000564	CR1	complement receptor 1 isoform F precursor	17	87	integral to plasma membrane membrane	complement activation, classical pathway innate immune response	complement component C3b receptor activity receptor activity	
Yes	2	86793109	NP_000642	CR1	complement receptor 1 isoform S precursor	13	92	integral to plasma membrane membrane	complement activation, classical pathway innate immune response	complement component C3b receptor activity receptor activity	
	2	4503029	NP_001869	CRABP2	cellular retinoic acid binding protein 2	1	4	not classified	epidermis development regulation of transcription, DNA- dependent signal transduction transport	lipid binding retinal binding	
	2	63054825	NP_775960	CRB2	crumbs homolog 2	1	2	integral to membrane membrane	response to stimulus visual per	calcium ion binding	
	2	21040241	NP_631900	CRB3	crumbs 3 isoform a precursor	1	6	integral to membrane membrane tight junction	not classified	not classified	
	2	59938770	NP_878901	CREB5	cAMP responsive element binding protein 5	2	2	intracellular nucleus	positive regulation of transcription, DNA-dependent transcription transcription from RNA polymerase II promoter	metal ion binding protein dimerization activity sequence-specific DNA binding transcription factor activity zinc ion binding	

Supplementary Table I

	2	41327712	NP_058431	CRK	v-crk sarcoma virus CT10 oncogene homolog	1	1	cytoplasm nucleus	actin cytoskeleton organization and biogenesis cell motility intracellular signaling cascade regulation of transcription from RNA polymerase II promoter	protein binding SH2 domain binding SH3/SH2 adaptor activity	
Yes	2	62422571	NP_001014809	CRMP1	collapsin response mediator protein 1 isoform 1	1	1	not classified	nervous system development nucleobase, nucleoside, nucleotide and nucleic acid metabolism	dihydropyrimidinase activity hydrolase activity protein binding	
	2	7706635	NP_057274	CRNN	cornulin	3	4	cytoplasm membrane	cell-cell adhesion metabolism response to heat response to unfolded protein	calcium ion binding	
	2	153791497	NP_055490	CROCC	ciliary rootlet coiled-coil, rootletin	1	4	centriole ciliary rootlet	cell cycle centrosome organization and biogenesis	kinesin binding structural molecule activity	
	2	32171215	NP_859066	CRTC2	transducer of regulated cAMP response element 2	1	1	not classified	not classified	not classified	
	2	4503057	NP_001876	CRYAB	crystallin, alpha B	7	12	cytoplasm nucleus	muscle contraction protein folding transmembrane receptor protein tyrosine kinase signaling pathway visual perception	structural constituent of eye lens unfolded protein binding	Alpha-B Crystallinopathy [MIM:608810]
	2	7705744	NP_057058	CRYL1	lambda-crystallin	7	21	not classified	fatty acid metabolism	oxidoreductase activity	
	2	4503065	NP_001879	CRYM	crystallin, mu isoform 1	1	3	not classified	visual perception	ornithine cyclodeaminase activity	Autosomal Dominant Nonsyndromic Deafness [MIM:123740]
	2	62241008	NP_001014444	CRYM	crystallin, mu isoform 2	1	3	not classified	visual perception	ornithine cyclodeaminase activity	
	2	13236495	NP_001880	CRYZ	crystallin, zeta	4	7	not classified	visual perception	NADPH:quinone reductase activity oxidoreductase activity zinc ion binding	
	2	29029559	NP_001307	CSE1L	CSE1 chromosome segregation 1-like protein 1	1	1	cytoplasm nuclear pore nucleus	apoptosis cell proliferation protein import into nucleus, docking protein transport	binding importin-alpha export receptor activity protein transporter activity	
	2	4758078	NP_004374	CSK	c-src tyrosine kinase	1	1	cytoplasm plasma membrane	intracellular signaling cascade protein amino acid phosphorylation regulation of progression through cell cycle	ATP binding nucleotide binding protein C-terminus binding protein-tyrosine kinase activity transferase activity	
Yes	2	98986450	NP_071331	CSNK1G1	casein kinase 1, gamma 1	1	1	not classified	protein amino acid phosphorylation Wnt receptor signaling pathway	ATP binding nucleotide binding protein serine/threonine kinase activity transferase activity	
Yes	2	21314778	NP_001310	CSNK1G2	casein kinase 1, gamma 2	1	1	not classified	protein amino acid phosphorylation signal transduction Wnt receptor signaling pathway	ATP binding casein kinase I activity nucleotide binding protein kinase activity protein serine/threonine kinase activity transferase activity	
	2	4758086	NP_004069	CSRP1	cysteine and glycine-rich protein 1	2	2	nucleus	not classified	metal ion binding zinc ion binding	

Supplementary Table 1

	2	4503107	NP_000090	CST3	cystatin C precursor	1	3	not classified	not classified	cysteine protease inhibitor activity protein homodimerization activity	Icelandic-type cerebroarterial amyloidosis [MIM:105150]
	2	4503117	NP_000091	CTSB	cystatin B	2	10	intracellular nucleus	not classified	cysteine protease inhibitor activity	Myoclonic Epilepsy of Unverricht and Lundborg [MIM:254800]
	2	56549681	NP_005799	CTDSPL	small CTD phosphatase 3 isoform 2	1	3	nucleus	not classified	phosphoric monoester hydrolase activity	
	2	56549683	NP_00100839	CTDSPL	small CTD phosphatase 3 isoform 1	1	4	nucleus	not classified	phosphoric monoester hydrolase activity	
	2	25453487	NP_714964	CTH	cystathionase isoform 2	1	1	not classified	amino acid biosynthesis cysteine biosynthesis	cystathionine gamma-lyase activity lyase activity	Cystathioninuria [MIM:219500]
	2	119943118	NP_001026851	CTNS	cystinosis, nephropathic isoform 1	1	3	integral to membrane lysosomal membrane membrane	amino acid metabolism L-cystine transport transport immune response proteolysis	L-cystine transporter activity	Nephropathic Cystinosis [MIM:219800]
	1,2	4505989	NP_000299	CTSA	cathepsin A precursor	3	15	endoplasmic reticulum lysosome	intracellular protein transport proteolysis	carboxypeptidase C activity enzyme activator activity intracellular transporter activity peptidase activity serine carboxypeptidase activity	Galactosialidosis [MIM:256540]
	2	4503141	NP_001805	CTSC	cathepsin C isoform a preproprotein	1	2	lysosome	immune response proteolysis	chloride ion binding cysteine-type endopeptidase activity dipeptidyl-peptidase I activity	Papillon-Lefevre Syndrome [MIM:245000]
	2	22538440	NP_680475	CTSC	cathepsin C isoform b precursor	1	1	lysosome	immune response proteolysis	chloride ion binding cysteine-type endopeptidase activity dipeptidyl-peptidase I activity	Papillo-LeFevre Syndrome [MIM:245000]
	2	4503143	NP_001900	CTSD	cathepsin D preproprotein	1	2	extracellular region lysosome	proteolysis	cathepsin D activity pepsin A activity peptidase activity	Neuronal Ceroid Lipofuscinosis [MIM:610127]
	2	4503149	NP_001902	CTSG	cathepsin G preproprotein	9	74	insoluble fraction	immune response proteolysis	cathepsin G activity peptidase activity	
	1,2	4557503	NP_001072	CUBN	cubilin	104	672	brush border membrane extrinsic to external side of plasma membrane membrane	cholesterol metabolism cobalamin transport lipid metabolism physiological process protein transport receptor mediated endocytosis steroid metabolism tissue homeostasis	calcium ion binding cobalamin binding cobalt ion binding protein homodimerization activity receptor activity transporter activity	Megaloblastic Anemia 1, Finnish Type [MIM:261100]
	2	4503165	NP_003581	CUL3	cullin 3	1	1	nucleus	cell cycle cell cycle arrest G 1/S transition of mitotic cell cycle induction of apoptosis by intracellular signals positive regulation of cell proliferation ubiquitin cycle	protein binding	
	2	28372493	NP_003579	CUL4B	cullin 4B	1	1	not classified	cell cycle ubiquitin cycle	protein binding	Cabezas Syndrome [MIM:300354] Mental Retardation-Hypotonic Facies Syndrome [MIM:300639]
	2	62198241	NP_00101443	CUTA	cutA divalent cation tolerance homolog isoform 1	1	1	membrane	protein localization	enzyme binding	
	2	62526026	NP_00101484	CUTA	cutA divalent cation tolerance homolog isoform 2	1	1	membrane	protein localization	enzyme binding	
	2	54792785	NP_056082	CUTL2	cut-like 2	2	2	nucleus	regulation of transcription, DNA-dependent	sequence-specific DNA binding transcription factor activity	
	2	19923603	NP_079119	CYBRD1	cytochrome b reductase 1	2	2	integral to membrane	electron transport	ferric-chelate reductase activity	

Supplementary Table 1

Yes	2	74271907	NP_001028200	CYFIP1	cytoplasmic FMR1 interacting protein 1 isoform 1	4	5	not classified	not classified	not classified	
Yes	2	24307969	NP_055423	CYFIP1	cytoplasmic FMR1 interacting protein 1 isoform 2	1	1	not classified	not classified	not classified	
	2	82617630	NP_055191	CYFIP2	cytoplasmic FMR1 interacting protein 2	1	1	cytoplasm synaptosome	not classified	not classified	
	2	20149621	NP_056348	DAK	dihydroxyacetone kinase 2	15	37	not classified	glycerol metabolism	ATP binding glycerone kinase activity lyase activity nucleotide binding transferase activity	
	2	62198235	NP_001014436	DBNL	drebrin-like isoform b	1	1	cell cortex cytoplasm cytoskeleton intracellular lamellipodium	activation of JNK activity endocytosis immune response Rac protein signal transduction	actin binding enzyme activator activity protein binding	
	2	16751921	NP_444513	DCD	dermcidin preproprotein	1	1	extracellular region	defense response to bacterium defense response to fungus xenobiotic metabolism	manganese ion binding	
	2	5453629	NP_006391	DCTN2	dynactin 2	1	1	centrosome cytoskeleton dynactin complex dynein complex kinetochore membrane microtubule	cell proliferation microtubule-based process mitosis	motor activity	
	1,2	7705925	NP_057370	DCXR	dicarbonyl/L-xylulose reductase	2	5	membrane	carbohydrate metabolism D-xylulose metabolism glucose metabolism metabolism	L-xylulose reductase activity oxidoreductase activity	
	1,2	6912328	NP_036269	DDAH1	dimethylarginine dimethylaminohydrolase 1	1	5	not classified	arginine catabolism nitric oxide mediated signal transduction	dimethylargininase activity hydrolase activity metal ion binding zinc ion binding	Hypertension [MIM: 604743]
	2	7524354	NP_039268	DDAH2	dimethylarginine dimethylaminohydrolase 2	1	2	not classified	anti-apoptosis arginine catabolism nitric oxide biosynthesis nitric oxide mediated signal transduction	dimethylargininase activity hydrolase activity	
	2	148529014	NP_001914	DDB1	damage-specific DNA binding protein 1	1	1	nucleus	nucleotide-excision repair ubiquitin cycle	damaged DNA binding	
	2	4503281	NP_000781	DDC	dopa decarboxylase (aromatic L-amino acid decarboxylase)	11	31	not classified	amino acid and derivative metabolism carboxylic acid metabolism catecholamine biosynthesis	aromatic-L-amino-acid decarboxylase activity carboxy-lyase activity lyase activity	Aromatic L-Amino Acid Decarboxylase Deficiency [MIM:608643]
	1,2	4503291	NP_001346	DDT	D-dopachrome tautomerase	1	1	not classified	melanin biosynthesis from tyrosine	dopachrome isomerase activity lyase activity	
	2	62241024	NP_00101445	DDX19B	DEAD (Asp-Glu-Ala-As) box polypeptide 19B	1	1	cytoplasm nuclear pore nucleus	mRNA export from nucleus	ATP binding ATP-dependent helicase activity hydrolase activity nucleic acid binding nucleotide binding RNA binding	
	2	87196351	NP_001347	DDX3X	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3X	2	3	cytoplasm nucleus	not classified	ATP binding ATP-dependent RNA helicase activity DNA binding hydrolase activity nucleotide binding RNA binding	

Supplementary Table I

Yes	2	13514809	NP_004651	DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3	1	1	nucleus	not classified	ATP binding ATP-dependent helicase activity DNA binding hydrolase activity nucleotide binding RNA binding	
	2	4758138	NP_004387	DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	1	1	nucleus	cell growth	ATP binding ATP-dependent helicase activity hydrolase activity nucleic acid binding nucleotide binding RNA binding RNA helicase activity	
Yes	1,2	4885179	NP_005208	DEFA3	defensin, alpha 3 preproprotein	2	14	extracellular region	defense response to bacterium defense response to fungus response to virus xenobiotic metabolism	not classified	
	2	38158020	NP_919409	DHX34	DEAH (Asp-Glu-Ala-His) box polypeptide 34	1	1	not classified	electron transport	ATP binding ATP-dependent helicase activity electron transporter activity hydrolase activity metal ion binding nucleic acid binding nucleotide binding RNA binding	
	2	24586663	NP_060641	DIP13B	DIP13 beta	1	1	cytoplasm endosome membrane membrane nucleus NuRD complex	cell cycle cell proliferation signal transduction	protein binding	
Yes	2	55749758	NP_775873	DIP2B	DIP2 disco-interacting protein 2 homolog B	2	3	nucleus	metabolism	catalytic activity transcription factor binding	
	2	63055057	NP_001017992	DKFZp686D0972	hypothetical protein LOC345651	2	2	not classified	not classified	protein binding	
Yes	2	8923740	NP_060049	DMBT1	deleted in malignant brain tumors 1 isoform	1	1	cytoplasm extracellular region membrane phagocytic vesicle membrane	cell cycle epithelial cell differentiation induction of bacterial agglutination innate immune response negative regulation of progression through cell cycle	calcium-dependent protein binding pattern recognition receptor activity scavenger receptor activity	
Yes	2	17864092	NP_061720	DNAH7	axonemal dynein heavy chain 7	1	1	axonemal dynein complex	ciliary or flagellar motility	microtubule motor activity	
	1,2	15029526	NP_001362	DNAH8	dynein, axonemal, heavy polypeptide 8	1	1	axonemal dynein complex dynein complex	ciliary or flagellar motility microtubule-based movement	ATP binding ATPase activity microtubule motor activity nucleotide binding protein binding	
	2	4504511	NP_001530	DNAJA1	DnaJ (Hsp40) homolog, subfamily A, memb	1	1	not classified	protein folding response to unfolded protein	heat shock protein binding low-density lipoprotein receptor binding metal ion binding unfolded protein binding zinc ion binding	
	2	5031741	NP_005871	DNAJA2	DnaJ subfamily A member 2	1	1	membrane	G1 phase of mitotic cell cycle positive regulation of cell proliferation protein folding regulation of progression through cell cycle	heat shock protein binding metal ion binding unfolded protein binding zinc ion binding	

Supplementary Table I

	2	5453690	NP_006136	DNAJB1	DnaJ (Hsp40) homolog, subfamily B, memb	2	2	nucleus	protein folding response to unfolded protein	heat shock protein binding unfolded protein binding	
	2	74047534	NP_056083	DNAJC13	DnaJ (Hsp40) homolog, subfamily C, memb	1	1	not classified	protein folding	binding heat shock protein binding	
	2	4507713	NP_003306	DNAJC7	DnaJ (Hsp40) homolog, subfamily C, memb	1	1	not classified	protein folding	heat shock protein binding	
	2	56549119	NP_004936	DNM2	dynamain 2 isoform 3	1	1	cytoplasm microtubule postsynaptic membrane	endocytosis G2/M transition of mitotic cell cycle positive regulation of apoptosis regulation of transcription signal transduction synaptic vesicle transport	enzyme binding GTP binding GTPase activity hydrolase activity microtubule binding motor activity nucleotide binding transcriptional activator activity	Charcot-Marie-Tooth Disease, Dominant Intermediate B [MIM:606482]
	2	56549125	NP_001005362	DNM2	dynamain 2 isoform 4	1	1	cytoplasm microtubule postsynaptic membrane	endocytosis G2/M transition of mitotic cell cycle positive regulation of apoptosis regulation of transcription signal transduction synaptic vesicle transport	enzyme binding GTP binding GTPase activity hydrolase activity microtubule binding motor activity nucleotide binding transcriptional activator activity	Charcot-Marie-Tooth Neuropathy, Dominant Intermediate B [MIM: 606482]
	2	45827701	NP_005119	DOPEY2	pad-1-like	1	1	Golgi membrane	development endoplasmic reticulum organization and biogenesis Golgi to endosome transport	not classified	
	1,2	4758190	NP_004404	DPEP1	dipeptidase 1 (renal)	31	204	endoplasmic reticulum membrane microsome	carbohydrate metabolism proteolysis	cation binding dipeptidyl-peptidase activity GPI anchor binding membrane dipeptidase activity metal ion binding metallopeptidase activity zinc ion binding	
Yes	2	11641273	NP_071752	DPEP3	dipeptidase 3	1	1	membrane	proteolysis	dipeptidyl-peptidase activity GPI anchor binding membrane dipeptidase activity metal ion binding metallopeptidase activity zinc ion binding	
	1,2	18765694	NP_001926	DPP4	dipeptidylpeptidase IV	41	232	integral to membrane membrane	immune response proteolysis	aminopeptidase activity dipeptidyl-peptidase IV activity prolyl oligopeptidase activity	
	2	4503375	NP_001376	DPYS	dihydropyrimidinase	5	6	not classified	nucleobase, nucleoside, nucleotide and nucleic acid metabolism response to toxin	dihydropyrimidinase activity hydrolase activity	Dihydropyrimidinuria [MIM:222748]
	2	13435366	NP_004940	DSC2	desmocollin 2 isoform Dsc2b preproprotein	1	3	cytoskeleton integral to membrane intercellular junction membrane	cell adhesion homophilic cell adhesion	calcium ion binding protein binding	Arrhythmogenic right ventricular dysplasia-11 [MIM:610476]
	2	119964718	NP_001935	DSG3	desmoglein 3 preproprotein	1	1	cytoskeleton integral to membrane intercellular junction membrane	cell adhesion homophilic cell adhesion	calcium ion binding protein binding	

Supplementary Table I

	2	58530842	NP_001008844	DSP	desmoplakin isoform II	10	16	cell-cell adherens junction cornified envelope intermediate filament	epidermis development keratinocyte differentiation peptide cross-linking	protein binding, bridging structural constituent of cytoskeleton	Keratosis Palmoplantaris Striata II Dilated Cardiomyopathy With Woolly Hair and Keratoderma [MIM:605676]
	2	58530840	NP_004406	DSP	desmoplakin isoform I	4	11	cell-cell adherens junction cornified envelope intermediate filament	epidermis development keratinocyte differentiation peptide cross-linking	protein binding, bridging structural constituent of cytoskeleton	
	2	5802966	NP_006861	DSTN	destrin isoform a	1	3	intracellular	not classified	actin binding	
Yes	2	58530848	NP_001011546	DSTN	destrin isoform b	2	6	intracellular	not classified	actin binding	
	2	13128968	NP_076930	DUSP26	dual specificity phosphatase 26	1	1	not classified	protein amino acid dephosphorylation	hydrolase activity kinase activity protein tyrosine/serine/threonine phosphatase activity	
	2	33350932	NP_001367	DYNC1H1	dynein, cytoplasmic, heavy polypeptide 1	14	17	cytoplasmic dynein complex dynein complex microtubule	microtubule-based movement mitotic spindle organization and biogenesis	ATP binding ATPase activity, coupled microtubule motor activity nucleotide binding	
	2	51468606	NP_001073932	DYNC2H1	dynein, cytoplasmic, heavy polypeptide 2	2	2	dynein complex Golgi apparatus intracellular	Golgi organization and biogenesis microtubule-based movement	ATP binding ATPase activity ATPase activity, coupled ATPase activity, uncoupled hydrolase activity microtubule motor activity motor activity	
	2	4503431	NP_003485	DYSF	dysferlin	1	1	integral to membrane plasma membrane	muscle contraction	not classified	Miyoshi Myopathy [MIM: 254130]
	2	4503443	NP_001388	ECE1	endothelin converting enzyme 1	1	1	integral to membrane membrane membrane fraction	cell-cell signaling proteolysis	endothelin-converting enzyme 1 activity metal ion binding neprilysin activity zinc ion binding	Hirschsprung Disease [MIM: 142623]
	2	55770888	NP_003557	EEA1	early endosome antigen 1, 162kD	1	1	cytoplasm cytosol early endosome extrinsic to plasma membrane intracellular membrane fraction	early endosome to late endosome transport synaptic vesicle to endosome fusion vesicle fusion	calmodulin binding GTP-dependent protein binding metal ion binding nucleic acid binding phosphatidylinositol binding protein homodimerization activity zinc ion binding	
	1,2	4503471	NP_001393	EEF1A1	eukaryotic translation elongation factor 1 alpha	6	15	cytoplasm	protein biosynthesis translational elongation	GTP binding GTPase activity nucleotide binding protein binding translation elongation factor activity	
Yes	2	4503475	NP_001949	EEF1A2	eukaryotic translation elongation factor 1 alpha	7	65	cytoplasm nucleus	protein biosynthesis translational elongation	GTP binding nucleotide binding translation elongation factor activity	
	2	4503483	NP_001952	EEF2	eukaryotic translation elongation factor 2	4	7	not classified	protein biosynthesis	GTP binding nucleotide binding translation elongation factor activity	
	2	9665262	NP_004096	EFEMP1	EGF-containing fibulin-like extracellular mat	1	1	extracellular matrix (sensu	visual perception	calcium ion binding protein binding	Doyme Honeycomb Retinal Dystrophy [MIM: 126600]
	2	4758248	NP_004420	EFNB1	ephrin-B1 precursor	1	1	integral to plasma membrane membrane soluble fraction synapse	cell adhesion cell differentiation cell-cell signaling nervous system development	ephrin receptor binding protein binding	

Supplementary Table I

								extracellular region integral to membrane nucleus plasma membrane	activation of MAPK activity chromosome organization and biogenesis (sensu Eukaryota) DNA replication epidermal growth factor receptor signaling pathway positive regulation of cell proliferation	calcium ion binding epidermal growth factor receptor activating ligand activity growth factor activity protein binding	
	1,2	4503491	NP_001954	EGF	epidermal growth factor (beta-urogastrone)	30	295				
	2	38093661	NP_938024	EGFL11	EGF-like-domain, multiple 11	1	2	integral to membrane membrane	not classified	ATP binding	
	1,2	30240932	NP_006786	EHD1	EH-domain containing 1	14	34	not classified	not classified	ATP binding calcium ion binding GTP binding GTPase activity nucleotide binding protein binding	
	2	21361462	NP_055416	EHD2	EH-domain containing 2	1	3	nucleus	not classified	calcium ion binding GTP binding GTPase activity nucleic acid binding nucleotide binding protein binding	
Yes	1,2	7657056	NP_055415	EHD3	EH-domain containing 3	8	20	nucleus	not classified	ATP binding calcium ion binding GTP binding GTPase activity nucleic acid binding nucleotide binding	
	1,2	21264315	NP_644670	EHD4	EH-domain containing 4	17	36	endoplasmic reticulum nucleus	not classified	ATP binding calcium ion binding GTP binding GTPase activity nucleic acid binding nucleotide binding	
	2	4503515	NP_003747	EIF3H	eukaryotic translation initiation factor 3, subv	1	5	eukaryotic translation initiation factor 3 complex	protein biosynthesis regulation of translational initiation	translation initiation factor activity not classified	
Yes	2	83700235	NP_001958	EIF4A2	eukaryotic translation initiation factor 4A2	1	2	eukaryotic translation initiation factor 4F complex	protein biosynthesis regulation of translational initiation	ATP binding ATP-dependent helicase activity hydrolase activity nucleic acid binding nucleotide binding protein binding RNA binding translation initiation factor activity	
Yes	1,2	4503545	NP_001961	EIF5A	eukaryotic translation initiation factor 5A	1	1	not classified	protein biosynthesis translational initiation	nucleic acid binding protein binding translation initiation factor activity	
Yes	2	9966867	NP_065123	EIF5A2	eIF-5A2 protein	1	1	not classified	protein biosynthesis translational initiation	nucleic acid binding translation initiation factor activity	

Supplementary Table I

	2	4503549	NP_001963	ELA2	elastase 2, neutrophil preproprotein	1	1	cell surface extracellular region	anti-inflammatory response calcium ion homeostasis negative regulation of chemokine biosynthesis negative regulation of chemotaxis negative regulation of inflammatory response negative regulation of interleukin-8 biosynthesis positive regulation of interl	bacterial binding cytokine binding elastase activity peptidase activity serine-type endopeptidase activity	Cyclic Hematopoiesis [MIM:162800]
	2	34485720	NP_899243	EML5	echinoderm microtubule associated protein	1	1	not classified	not classified	lyase activity	
Yes	1,2	4503571	NP_001419	ENO1	enolase 1	22	103	cytoplasm nucleus phosphopyruvate hydratase complex	glycolysis regulation of transcription, DNA-dependent transcription	DNA binding lyase activity magnesium ion binding phosphopyruvate hydratase activity plasminogen activator activity	
Yes	2	5803011	NP_001966	ENO2	enolase 2	2	7	phosphopyruvate hydratase	glycolysis	lyase activity magnesium ion binding phosphopyruvate hydratase activity	
Yes	2	153267448	NP_443739	ENO3	enolase 3	2	3	phosphopyruvate hydratase	glycolysis	lyase activity magnesium ion binding phosphopyruvate hydratase activity	
	1,2	36951116	NP_001968	ENPEP	glutamyl aminopeptidase (aminopeptidase A)	25	100	integral to plasma membrane plasma membrane	cell migration cell proliferation cell-cell signaling proteolysis	calcium ion binding glutamyl aminopeptidase activity membrane alanyl aminopeptidase activity metallopeptidase activity zinc ion binding	Hypertension [MIM:138297]
	2	4826896	NP_005012	ENPP3	ectonucleotide pyrophosphatase/phosphodi	2	2	integral to plasma membrane membrane	nucleotide metabolism phosphate metabolism	endonuclease activity hydrolase activity nucleic acid binding nucleotide diphosphatase activity phosphodiesterase I activity	
	2	7662358	NP_055751	ENPP4	ectonucleotide pyrophosphatase/phosphodi	2	1	not classified	nucleotide metabolism	not classified	
	2	23503267	NP_699174	ENPP6	ectonucleotide pyrophosphatase/phosphodi	5	24	integral to membrane membrane	lipid catabolism nucleotide metabolism	hydrolase activity	
	2	27597073	NP_001970	EPHX2	epoxide hydrolase 2, cytoplasmic	1	1	cytosol peroxisome soluble fraction	aromatic compound catabolism blood pressure regulation calcium ion homeostasis drug metabolism inflammatory response metabolism oxygen and reactive oxygen species metabolism positive regulation of vasodilation response to toxin xenobiotic metabolism	epoxide hydrolase activity hydrolase activity magnesium ion binding protein homodimerization activity	
	2	8923678	NP_060427	EPN3	epsin 3	1	1	not classified	not classified	lipid binding	
	2	13876386	NP_112598	EPPK1	epiplakin 1	4	6	cytoskeleton	not classified	structural molecule activity	
	1,2	56682953	NP_004438	EPS8	epidermal growth factor receptor pathway s	14	30	not classified	cell proliferation epidermal growth factor receptor signaling pathway signal transduction	protein binding SH3/SH2 adaptor activity	
	2	21264608	NP_573441	EPS8L1	epidermal growth factor receptor pathway substrate 8-like protein 1 isoform a	6	8	not classified	not classified	not classified	
	2	148539855	NP_060199	EPS8L1	epidermal growth factor receptor pathway s	4	7	not classified	not classified	hydrolase activity	

Supplementary Table I

	2	21264616	NP_073609	EPS8L2	epidermal growth factor receptor pathway substrate 8-like protein 2	11	18	not classified	not classified	not classified	
Yes	2	4503595	NP_000493	EPX	eosinophil peroxidase	1	1	not classified	hydrogen peroxide catabolism	calcium ion binding iron ion binding oxidoreductase activity peroxidase activity	
	2	33413400	NP_001975	ESD	esterase D/formylglutathione hydrolase	1	1	cytoplasmic membrane-bound vesicle	not classified	S-formylglutathione hydrolase activity serine esterase activity thiosulfate sulfurtransferase activity	
	2	4503613	NP_001979	EVPL	envoplakin	18	43	cornified envelope plasmodesma	epidermis development keratinization keratinocyte differentiation peptide cross-linking	protein binding, bridging structural molecule activity	
	2	4503627	NP_000119	F11	platelet coagulation factor XI precursor	1	1	extracellular region membrane	blood coagulation proteolysis	coagulation factor IXa activity coagulation factor XIa activity heparin binding peptidase activity	
	2	4557577	NP_001434	FABP1	fatty acid binding protein 1, liver	2	3	cytoplasm	cell-cell signaling fatty acid metabolism organ morphogenesis transport	lipid binding lipid transporter activity	
	1,2	4758328	NP_004093	FABP3	fatty acid binding protein 3	2	6	cytoplasm soluble fraction	negative regulation of cell proliferation transport	fatty acid binding lipid transporter activity	
	2	4557587	NP_000128	FAH	fumarylacetoacetate hydrolase (fumarylacetoace	2	2	not classified	aromatic amino acid family metabolism L-phenylalanine catabolism metabolism tyrosine catabolism	calcium ion binding fumarylacetoacetase activity hydrolase activity magnesium ion binding	Tyrosinemia Type I [MIM:276700]
	1,2	24308440	NP_612410	FAM125A	hypothetical protein LOC93343	2	4	nucleus	not classified	not classified	
	2	58761488	NP_258257	FAM125B	hypothetical protein LOC89853 isoform 1	1	1	not classified	not classified	not classified	
	2	116174742	NP_064608	FAM20C	family with sequence similarity 20, member	1	1	not classified	not classified	not classified	
	2	56790935	NP_149079	FAM40A	hypothetical protein LOC85369	1	1	not classified	not classified	not classified	
	2	42734438	NP_057707	FAM49B	hypothetical protein LOC51571	1	2	not classified	not classified	not classified	
	2	39995080	NP_078795	FAM65A	hypothetical protein LOC79567	2	2	not classified	not classified	binding DNA-directed RNA polymerase activity	
	2	41872631	NP_004095	FASN	fatty acid synthase	20	38	not classified	fatty acid biosynthesis	3-hydroxypalmitoyl-[acyl-carrier protein] dehydratase activity 3-oxoacyl-[acyl-carrier protein] reductase activity 3-oxoacyl-[acyl-carrier protein] synthase activity catalytic activity cofactor binding enoyl-[acyl-carrier protein] reductase (NADPH, B-spec	
	2	66346693	NP_005236	FAT	FAT tumor suppressor 1 precursor	1	1	integral to plasma membrane membrane	cell adhesion cell-cell signaling homophilic cell adhesion morphogenesis	calcium ion binding protein binding	
	2	12056465	NP_001427	FBL	fibrillarlin	1	1	nucleus ribonucleoprotein complex	rRNA processing	RNA binding transferase activity not classified	

Supplementary Table 1

	2	16579888	NP_000498	FBP1	fructose-1,6-bisphosphatase 1	7	16	not classified	carbohydrate metabolism fructose metabolism gluconeogenesis	fructose-2,6-bisphosphate 2-phosphatase activity fructose-bisphosphatase activity hydrolase activity identical protein binding phosphoric ester hydrolase activity zinc ion binding	Fructose-1,6-Bisphosphatase Deficiency [MIM:229700]
	2	22907028	NP_003828	FBP2	fructose-1,6-bisphosphatase 2	1	2	not classified	carbohydrate metabolism fructose metabolism gluconeogenesis	fructose-2,6-bisphosphate 2-phosphatase activity fructose-bisphosphatase activity hydrolase activity phosphoric ester hydrolase activity	
	1,2	4503681	NP_003881	FCGBP	Fc fragment of IgG binding protein	14	34	not classified	not classified	not classified	
	2	61744445	NP_004099	FCN2	ficolin 2 isoform a precursor	2	4	cytoplasm	antimicrobial humoral response (sensu Vertebrata) opsonization phosphate transport	antigen binding calcium ion binding sugar binding	
	2	7305053	NP_038479	FER1L3	myoferlin isoform a	8	15	integral to membrane nuclear envelope plasma membrane	circulation muscle contraction	not classified	
	2	19718759	NP_579899	FER1L3	myoferlin isoform b	7	12	integral to membrane nuclear envelope plasma membrane	circulation muscle contraction	calcium ion binding	
	1,2	4503689	NP_000499	FGA	fibrinogen, alpha polypeptide isoform alpha-E p	5	17	extracellular region fibrinogen complex soluble fraction	blood coagulation blood pressure regulation platelet activation positive regulation of cell proliferation protein polymerization signal transduction	protein binding	Renal Amyloidosis [MIM:105200] Dysfibrinogenemia [MIM:134820]
Yes	2	70906435	NP_005132	FGB	fibrinogen, beta chain preproprotein	1	4	extracellular region fibrinogen complex soluble fraction	blood coagulation blood pressure regulation positive regulation of cell proliferation	not classified	
	2	70906437	NP_000500	FGG	fibrinogen, gamma chain isoform gamma-A pre	1	1	fibrinogen complex	blood pressure regulation platelet activation positive regulation of cell proliferation protein polymerization signal transduction	calcium ion binding	Dysfibrinogenemia[MIM:134850]
	2	70906439	NP_068656	FGG	fibrinogen, gamma chain isoform gamma-B	1	3	fibrinogen complex	blood pressure regulation platelet activation positive regulation of cell proliferation protein polymerization signal transduction	FK506 binding	
	2	5730075	NP_006673	FGL2	fibrinogen-like 2 precursor	4	7	fibrinogen complex	not classified	not classified	
	2	31742536	NP_071399	FIGNL1	fidetin-like 1	1	6	membrane	not classified	ATP binding nucleoside-triphosphatase activity nucleotide binding	
	2	4503729	NP_002005	FKBP4	FK506-binding protein 4	1	1	cytoplasm intracellular nucleus	protein folding	isomerase activity peptidyl-prolyl cis-trans isomerase activity protein binding, bridging actin binding	

Supplementary Table I

	2	4758384	NP_004108	FKBP5	FK506 binding protein 5	1	1	nucleus	protein folding	FK506 binding isomerase activity peptidyl-prolyl cis-trans isomerase activity protein binding	
Yes	2	62990121	NP_00101742	FKSG30	actin-like protein	6	135	actin filament	not classified	ATP binding protein binding structural constituent of cytoskeleton	
	2	62865871	NP_060137	FLJ20097	coiled-coil domain containing 132 isoform a	1	2	not classified	two-component signal transduction system (phosphorelay)	two-component sensor activity	
	2	47458827	NP_689704	FLJ32569	hypothetical protein LOC148811	2	2	not classified	proteolysis	carboxypeptidase activity Gly-X carboxypeptidase activity hydrolase activity metallopeptidase activity protein dimerization activity	
	2	22749369	NP_689894	FLJ38482	hypothetical protein LOC201931	1	1	not classified	not classified	not classified	
	2	116063573	NP_001447	FLNA	filamin A, alpha isoform 1	1	1	actin cytoskeleton	actin cytoskeleton organization and biogenesis cell motility cell surface receptor linked signal transduction nervous system development positive regulation of I-kappaB kinase/NF-kappaB cascade	actin filament binding protein binding signal transduce activity actin binding	
	2	105990514	NP_001448	FLNB	filamin B, beta (actin binding protein 278)	1	1	actin cytoskeleton integral to membrane	actin cytoskeleton organization and biogenesis cell differentiation cytoskeletal anchoring signal transduction striated muscle development	protein binding collagen binding	Spondylocapotalarsal Synostosis Syndrome [MIM: 272460]
	1,2	5031699	NP_005794	FLOT1	flotillin 1	12	34	caveola flotillin complex integral to membrane membrane	not classified	protein binding	
	2	4758394	NP_004466	FLOT2	flotillin 2	9	14	cell surface flotillin complex integral to membrane plasma membrane	cell adhesion epidermis development	protein binding	
	1,2	47132557	NP_997647	FN1	fibronectin 1 isoform 1 preproprotein	1		ER-Golgi intermediate compartment extracellular matrix (sensu Metazoa) extracellular region	acute-phase response cell adhesion cell migration metabolism response to wounding transmembrane receptor protein tyrosine kinase signaling pathway	collagen binding extracellular matrix structural constituent heparin binding oxidoreductase activity	
	2	16933542	NP_002017	FN1	fibronectin 1 isoform 3 preproprotein	9	18	ER-Golgi intermediate compartment extracellular matrix (sensu Metazoa) extracellular region	acute-phase response cell adhesion cell migration metabolism response to wounding transmembrane receptor protein tyrosine kinase signaling pathway	collagen binding extracellular matrix structural constituent heparin binding oxidoreductase activity	

Supplementary Table I

Yes	2	62548858	NP_001014986	FOLH1	folate hydrolase 1 isoform 2	1	1	integral to plasma membrane membrane fraction	proteolysis	carboxypeptidase activity dipeptidase activity glutamate carboxypeptidase II activity metal ion binding metallopeptidase activity zinc ion binding	
	2	4503787	NP_002022	FRK	fyn-related kinase	7	17	intracellular nucleus	intracellular signaling cascade negative regulation of cell proliferation protein amino acid phosphorylation regulation of progression through cell cycle	ATP binding non-membrane spanning protein tyrosine kinase activity nucleotide binding transferase activity	
	2	11140815	NP_006648	FTCD	formiminotransferase cyclodeaminase	6	6	cytoplasm	folic acid and derivative metabolism histidine metabolism metabolism	folic acid binding formimidoyltetrahydrofolate cyclodeaminase activity glutamate formimidoyltransferase activity lyase activity transferase activity	Glutamate Formiminotransferase Deficiency [MIM: 229100]
	2	56682959	NP_002023	FTH1	ferritin, heavy polypeptide 1	1	7	ferritin complex plasma membrane	cell proliferation immune response intracellular sequestering of iron ion iron ion homeostasis iron ion transport negative regulation of cell proliferation	binding ferric iron binding ferroxidase activity kinase binding metal ion binding oxidoreductase activity protein binding	Iron Overload, Autosomal Dominant [MIM:134770]
	1,2	20149498	NP_000137	FTL	ferritin, light polypeptide	5	14	ferritin complex	iron ion homeostasis iron ion transport	binding ferric iron binding identical protein binding metal ion binding	Hyperferritinemia-Cataract Syndrome [MIM:600886]
	2	4503803	NP_000138	FUCA1	fucosidase, alpha-L-1, tissue	1	1	cytoplasm lysosome	carbohydrate metabolism glycosaminoglycan catabolism metabolism	alpha-L-fucosidase activity hydrolase activity, acting on glycosyl bonds	Fucosidosis [MIM:230000]
	2	20149693	NP_079405	FUZ	fuzzy homolog	1	1	not classified	not classified	not classified	
	1,2	11125766	NP_001671	FXSD2	FXSD domain-containing ion transport regulator	1	7	integral to membrane membrane sodium:potassium-exchanging ATPase complex	ion transport potassium ion transport sodium ion transport	ion channel activity potassium ion binding sodium ion binding sodium:potassium-exchanging ATPase activity transporter activity	Hypomagnesemia 2, Renal [MIM:154020]
	2	21614520	NP_000393	G6PD	glucose-6-phosphate dehydrogenase isoform a	1	1	cytoplasm	carbohydrate metabolism glucose 6-phosphate utilization glucose metabolism pentose-phosphate shunt	glucose-6-phosphate 1-dehydrogenase activity oxidoreductase activity	Nonspherocytic Hemolytic Anemia due to G6PD Deficiency [MIM:305900]
	1,2	4503849	NP_000143	GAA	acid alpha-glucosidase preproprotein	4	8	lysosome	carbohydrate metabolism generation of precursor metabolites and energy glycogen catabolism metabolism	alpha-glucosidase activity hydrolase activity, hydrolyzing O-glycosyl compounds	Infantile-onset Glycogen Storage Disease Type II [MIM:232300]
	2	119393895	NP_001073272	GAA	acid alpha-glucosidase preproprotein	5	8	lysosome	carbohydrate metabolism generation of precursor metabolites and energy glycogen catabolism metabolism	hydrolase activity, hydrolyzing O-glycosyl compounds ATP binding	

Supplementary Table I

								integral to plasma membrane postsynaptic membrane	gamma-aminobutyric acid signaling pathway ion transport synaptic transmission	chloride channel activity chloride ion binding extracellular ligand-gated ion channel activity GABA-A receptor activity ion channel activity neurotransmitter receptor activity	
2	12548785	NP_068711	GABRB2	gamma-aminobutyric acid (GABA) A receptor	1	1					
Yes	2	109948279	NP_001035972	GAL7	galactin 7	1	4	not classified	not classified	not classified	
2	4503895	NP_000145	GALK1	galactokinase 1	1	1	cytoplasm membrane fraction	carbohydrate phosphorylation galactose metabolism metabolism	carbohydrate phosphorylation galactose metabolism metabolism	galactokinase activity galactose binding kinase activity nucleotide binding transferase activity hydrolase activity	Galactokinase Deficiency [MIM: 230200]
2	20270355	NP_620156	GALM	galactose mutarotase (aldose 1-epimerase)	2	3	not classified	carbohydrate metabolism galactose metabolism	carbohydrate metabolism galactose metabolism	aldose 1-epimerase activity isomerase activity	
2	38202257	NP_938148	GANAB	alpha glucosidase II alpha subunit isoform 2	1	1	endoplasmic reticulum	carbohydrate metabolism	carbohydrate metabolism	hydrolase activity hydrolase activity, hydrolyzing O-glycosyl compounds protein binding	
2	88900491	NP_938149	GANAB	alpha glucosidase II alpha subunit isoform 3	1	1	endoplasmic reticulum	carbohydrate metabolism	carbohydrate metabolism	hydrolase activity, hydrolyzing O-glycosyl compounds protein binding hydrolase activity, acting on glycosyl bonds	
1,2	7669492	NP_002037	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	10	114	cytoplasm	glucose metabolism glycolysis	glucose metabolism glycolysis	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity NAD binding oxidoreductase activity	
Yes	2	7657116	NP_055179	GAPDHS	glyceraldehyde-3-phosphate dehydrogenase	1	5	not classified	glucose metabolism glycolysis positive regulation of glycolysis sperm motility	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity NAD binding oxidoreductase activity	
2	4557619	NP_000149	GBE1	glucan (1,4-alpha-), branching enzyme 1	1	2	not classified	carbohydrate metabolism generation of precursor metabolites and energy glycogen biosynthesis	carbohydrate metabolism generation of precursor metabolites and energy glycogen biosynthesis	1,4-alpha-glucan branching enzyme activity cation binding hydrolase activity, hydrolyzing O-glycosyl compounds transferase activity, transferrin glycosyl groups	Type IV Glycogen Storage Disease [MIM:232500]
2	6912388	NP_036330	GCA	grancalcin, EF-hand calcium binding protein	1	1	cytoplasm plasma membrane	membrane fusion	membrane fusion	calcium ion binding	
2	149999606	NP_006293	GCS1	mannosyl-oligosaccharide glucosidase	1	1	endoplasmic reticulum integral to membrane membrane membrane fraction	metabolism oligosaccharide metabolism protein amino acid N-linked glycosylation	metabolism oligosaccharide metabolism protein amino acid N-linked glycosylation	mannosyl-oligosaccharide glucosidase activity not classified	Congenital Disorder of Glycosylation [MIM: 606056]
2	7705308	NP_057288	GDF2	growth differentiation factor 2	1	1	extracellular region extracellular space	growth	growth	cytokine activity growth factor activity	
Yes	2	4503971	NP_001484	GDI1	GDP dissociation inhibitor 1	6	17	cytoplasm	protein transport regulation of GTPase activity signal transduction	GTPase activator activity Rab GDP-dissociation inhibitor activity	
2	6598323	NP_001485	GDI2	GDP dissociation inhibitor 2	17	38	cell surface cytoplasm	protein transport regulation of GTPase activity signal transduction	protein transport regulation of GTPase activity signal transduction	GTPase activator activity Rab GDP-dissociation inhibitor activity	

Supplementary Table 1

	2	146198640	NP_077283	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3	1	2	not classified	glycerol metabolism	glycerophosphodiester phosphodiesterase activity	
	2	7657122	NP_056536	GEMIN4	gemin4	1	1	cytoplasm nucleolus nucleus small nuclear ribonucleoprotein complex spliceosome complex	nuclear mRNA splicing, via spliceosome rRNA processing	protein binding	
Yes	2	4503981	NP_002047	GFPT1	glucosamine-fructose-6-phosphate aminotransferase 1	1	1	cytoplasm	carbohydrate biosynthesis carbohydrate metabolism energy reserve metabolism fructose 6-phosphate metabolism glutamine metabolism metabolism	glutamine-fructose-6-phosphate transaminase (isomerizing) activity sugar binding transferase activity	
Yes	2	4826742	NP_005101	GFPT2	glutamine-fructose-6-phosphate transaminase 2	1	1	cytoplasm	carbohydrate biosynthesis carbohydrate metabolism energy reserve metabolism fructose 6-phosphate metabolism glutamine metabolism metabolism	glutamine-fructose-6-phosphate transaminase (isomerizing) activity sugar binding transferase activity	
	2	22035692	NP_665736	GFRA1	GDNF family receptor alpha 1 isoform b precursor	1	1	extrinsic to membrane membrane	cell surface receptor linked signaling pathway	glial cell line-derived neurotrophic factor receptor activity GPI anchor binding receptor activity receptor binding	
	2	4503987	NP_003869	GGH	gamma-glutamyl hydrolase precursor	5	11	lysosome	glutamine metabolism	exopeptidase activity gamma-glutamyl hydrolase activity hydrolase activity	
Yes	2	73915094	NP_001027536	GGT1	gamma-glutamyltransferase 1 precursor	1	5	integral to membrane membrane	amino acid metabolism glutathione biosynthesis	acyltransferase activity gamma-glutamyltransferase activity transferase activity	
Yes	2	62079287	NP_002049	GGT2 (GGT1)	gamma-glutamyltransferase 2	11	116	cellular component unknown	glutathione biosynthesis	gamma-glutamyl transferase activity	
	2	118200356	NP_055209	GHITM	growth hormone inducible transmembrane protein 1	1	1	integral to membrane membrane mitochondrion	not classified	protein binding	
	2	42544144	NP_974197	GIPC1	regulator of G-protein signalling 19 interacting protein 1	1	1	cytosol membrane membrane fraction soluble fraction	G-protein coupled receptor protein signaling pathway	protein binding receptor binding	
	2	42544149	NP_974196	GIPC1	regulator of G-protein signalling 19 interacting protein 1	1	1	cytosol membrane membrane fraction soluble fraction	G-protein coupled receptor protein signaling pathway	receptor binding ATP binding	
	2	41393579	NP_060125	GIPC2	PDZ domain protein GIPC2	1	1	not classified	not classified	protein binding	
	2	42794763	NP_976325	GK	glycerol kinase isoform a	1	1	cytoplasm mitochondrial outer membrane	carbohydrate metabolism glycerol-3-phosphate metabolism	glycerol kinase activity nucleotide binding transferase activity ATP binding	Glycerol Kinase Deficiency [MIM: 307030]
	2	41393575	NP_149991	GK2	glycerol kinase 2	1	1	cytoplasm mitochondrial outer membrane	carbohydrate metabolism glycerol-3-phosphate metabolism	glycerol kinase activity nucleotide binding transferase activity GTP binding	

Supplementary Table I

	1,2	10834966	NP_000395	GLB1	galactosidase, beta 1 isoform a	16	70	lysosome beta-galactosidase complex	carbohydrate metabolism metabolism	beta-galactosidase activity cation binding hydrolase activity, acting on glycosyl bonds	Gangliosidosis GM1 [MIM:230500]
Yes	2	119372312	NP_001073279	GLB1	galactosidase, beta 1 isoform b	17	73	lysosome beta-galactosidase complex	carbohydrate metabolism metabolism	beta-galactosidase activity cation binding hydrolase activity, acting on glycosyl bonds	
Yes	2	40255043	NP_078782	GLB1L	galactosidase, beta 1-like	1	2	lysosome beta-galactosidase complex	carbohydrate metabolism	beta-galactosidase activity catalytic activity cation binding	
	1,2	54633312	NP_036333	GLG1	golgi apparatus protein 1	2	2	membrane integral to membrane	not classified	fibroblast growth factor binding receptor binding	
	2	5729842	NP_006699	GLO1	glyoxalase I	1	1	cytoplasm	anti-apoptosis carbohydrate metabolism	lactoylglutathione lyase activity lyase activity metal ion binding zinc ion binding	
Yes	2	7705987	NP_057517	GLTP	glycolipid transfer protein	1	1	membrane	glycolipid transport	glycolipid binding glycolipid transporter activity	
	2	74271837	NP_001028216	GLUL	glutamine synthetase	2	2	not classified	glutamine biosynthesis nitrogen compound metabolism regulation of neurotransmitter levels	glutamate-ammonia ligase activity ligase activity	Congenital Glutamine Deficiency [MIM:610015]
	2	39995109	NP_000396	GM2A	GM2 ganglioside activator precursor	2	3	lysosome	glycolipid catabolism glycosphingolipid metabolism lipid metabolism sphingolipid catabolism	sphingolipid activator protein activity	Gangliosidosis GM2 AB Variant Tay-Sachs Disease [MIM:272750]
	1,2	4504037	NP_002058	GNA11	guanine nucleotide binding protein (G protein)	7	20	cytoplasm plasma membrane	G-protein coupled receptor protein signaling pathway protein amino acid ADP- ribosylation signal transduction	GTP binding GTPase activity nucleotide binding signal transducer activity	
Yes	2	42476111	NP_031379	GNA12	guanine nucleotide binding protein (G protein)	1	11	membrane	blood coagulation G-protein coupled receptor protein signaling pathway signal transduction	GTP binding GTPase activity nucleotide binding protein binding signal transducer activity	
Yes	2	24111250	NP_006563	GNA13	guanine nucleotide binding protein (G protein)	9	20	membrane	cell motility G-protein coupled receptor protein signaling pathway signal transduction	GTP binding GTPase activity nucleotide binding protein binding receptor signaling protein activity	
Yes	2	4758444	NP_004288	GNA14	guanine nucleotide binding protein (G protein)	5	11	plasma membrane heterotrimeric G-protein complex	G-protein coupled receptor protein signaling pathway protein amino acid ADP- ribosylation signal transduction	GTP binding GTPase activity nucleotide binding signal transducer activity	
	1,2	33946324	NP_002060	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	8	16	plasma membrane	G-protein coupled receptor protein signaling pathway signal transduction	GTP binding nucleotide binding protein binding signal transducer activity	

Supplementary Table I

	1,2	4504041	NP_002061	GNAI2	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2	8	21	not classified	G-protein coupled receptor protein signaling pathway negative regulation of adenylate cyclase activity response to nutrient signal transduction	GTP binding GTPase activity nucleotide binding signal transducer activity	
	1,2	5729850	NP_006487	GNAI3	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 3	8	16	not classified	G-protein coupled receptor protein signaling pathway negative regulation of adenylate cyclase activity signal transduction transport	GTP binding GTPase activity nucleotide binding signal transducer activity	
Yes	2	41281685	NP_620073	GNAO1	guanine nucleotide binding protein, alpha ac	2	4	not classified	axon guidance G-protein coupled receptor protein signaling pathway muscle contraction nervous system development signal transduction	GTP binding GTPase activity nucleotide binding receptor signaling protein activity	
	2	40254462	NP_002063	GNAQ	guanine nucleotide binding protein (G protein)	10	27	cytoplasm heterotrimeric G-protein complex plasma membrane	blood coagulation G-protein coupled receptor protein signaling pathway phospholipase C activation protein amino acid ADP-ribosylation signal transduction	GTP binding GTPase activity nucleotide binding signal transducer activity	
Yes	1,2	4504047	NP_000507	GNAS	GNAS complex locus isoform a	13	43	cytoplasm extracellular region heterotrimeric G-protein complex plasma membrane	G-protein coupled receptor protein signaling pathway G-protein signaling, adenylate cyclase activating pathway pregnancy protein secretion signal transduction	GTP binding GTPase activity guanyl nucleotide binding nucleotide binding signal transducer activity	
Yes	2	117938762	NP_001070956	GNAS	GNAS complex locus isoform f	1	1	cytoplasm extracellular region heterotrimeric G-protein complex plasma membrane	G-protein coupled receptor protein signaling pathway G-protein signaling, adenylate cyclase activating pathway pregnancy protein secretion signal transduction	GTP binding GTPase activity guanyl nucleotide binding nucleotide binding signal transducer activity	
Yes	2	117938765	NP_001070955	GNAS	GNAS complex locus isoform g	15	54	cytoplasm extracellular region heterotrimeric G-protein complex plasma membrane	G-protein coupled receptor protein signaling pathway G-protein signaling, adenylate cyclase activating pathway pregnancy protein secretion signal transduction	GTP binding GTPase activity guanyl nucleotide binding nucleotide binding signal transducer activity	
Yes	2	20330805	NP_005263	GNAT2	guanine nucleotide binding protein, alpha tra	1	3	heterotrimeric G-protein co	G-protein coupled receptor protein signaling pathway phototransduction response to stimulus signal transduction visual perception	G-protein coupled photoreceptor activity GTP binding nucleotide binding signal transducer activity	
	1,2	11321585	NP_002065	GNB1	guanine nucleotide-binding protein, beta-1 s	2	6	not classified	signal transduction	signal transducer activity	
Yes	1,2	20357529	NP_005264	GNB2	guanine nucleotide-binding protein, beta-2 s	9	67	not classified	signal transduction	signal transducer activity	
	2	5174447	NP_006089	GNB2L1	guanine nucleotide binding protein (G protein)	1	1	cytoplasm	protein kinase C activation signal transduction	protein kinase C binding receptor binding	

Supplementary Table I

Yes	2	4504053	NP_002066	GNB3	guanine nucleotide-binding protein, beta-3 s	1	1	not classified	blood pressure regulation G-protein coupled receptor protein signaling pathway signal transduction	GTPase activity signal transducer activity	
	1,2	51036603	NP_061329	GNG12	G-protein gamma-12 subunit	5	39	heterotrimeric G-protein complex membrane	G-protein coupled receptor protein signaling pathway signal transduction	signal transducer activity	
	2	4885287	NP_005265	GNG5	guanine nucleotide binding protein (G protei	1	2	heterotrimeric G-protein complex membrane	G-protein coupled receptor protein signaling pathway signal transduction	signal transducer activity	
	1,2	32698769	NP_443079	GNG7	guanine nucleotide binding protein (G protei	1	2	heterotrimeric G-protein complex membrane	regulation of G-protein coupled receptor protein signaling pathway signal transduction	signal transducer activity	
	2	13027378	NP_005462	GNPDA1	glucosamine-6-phosphate deaminase 1	1	1	not classified	carbohydrate metabolism fertilization (sensu Metazoa) generation of precursor metabolites and energy glucosamine catabolism N-acetylglucosamine metabolism	glucosamine-6-phosphate deaminase activity hydrolase activity	
	2	50541946	NP_001002296	GOLGA7	golgi autoantigen, golgin subfamily a, 7	1	4	membrane	not classified	not classified	
	2	4504067	NP_002070	GOT1	aspartate aminotransferase 1	9	15	cytosol	amino acid metabolism aspartate catabolism biosynthesis	aspartate transaminase activity	
	2	33695088	NP_005267	GPD1	glycerol-3-phosphate dehydrogenase 1 (sol	11	24	cytoplasm glycerol-3-phosphate dehydrogenase complex	carbohydrate metabolism glycerol-3-phosphate catabolism	glycerol-3-phosphate dehydrogenase (NAD+) activity NAD binding	
	1,2	18201905	NP_000166	GPI	glucose phosphate isomerase	9	19	extracellular space	carbohydrate metabolism gluconeogenesis glycolysis hemostasis humoral immune response nervous system development	cytokine activity glucose-6-phosphate isomerase activity growth factor activity isomerase activity	Chronic Hemolytic Anemia Duet to GPI Deficiency [MIM:172400]
Yes	2	42476105	NP_963886	GPM6A	glycoprotein M6A isoform 3	1	1	cell surface integral to plasma membrane membrane	not classified	not classified	
	2	42476108	NP_963885	GPM6A	glycoprotein M6A isoform 2	1	1	cell surface integral to plasma membrane membrane	not classified	glutathione peroxidase activity	
	2	74271834	NP_689742	GPR155	G protein-coupled receptor 155 isoform 10	1	1	integral to membrane membrane	intracellular signaling cascade	not classified	
	2	19882213	NP_115495	GPR98	G protein-coupled receptor 98 precursor	1	1	integral to membrane membrane	cell-cell adhesion nervous system development neuropeptide signaling pathway response to stimulus sensory perception of sound signal transduction visual perception	calcium ion binding G-protein coupled receptor activity receptor activity	Familial Febrile Seizures [MIM:604352] Usher Syndrome Type IIC [MIM:605472]
	1,2	4506403	NP_003970	GPRC5A	G protein-coupled receptor, family C, group	1	4	integral to plasma membrane membrane	G-protein coupled receptor protein signaling pathway signal transduction	metabotropic glutamate, GABA-B-like receptor activity receptor activity	

Supplementary Table 1

	1,2	7706451	NP_057319	GPRC5B	G protein-coupled receptor, family C, group 7	7	52	integral to membrane membrane	G-protein coupled receptor protein signaling pathway signal transduction visual perception	metabotropic glutamate, GABA-B-like receptor activity receptor activity sevenless binding	
	2	40217833	NP_061123	GPRC5C	G protein-coupled receptor family C, group 11	11	133	integral to plasma membrane membrane	G-protein coupled receptor protein signaling pathway signal transduction	metabotropic glutamate, GABA-B-like receptor activity protein binding receptor activity	
	2	4885351	NP_005300	GPT	glutamic-pyruvate transaminase (alanine aminotransferase)	1	1	cytoplasm soluble fraction	biosynthesis gluconeogenesis nitrogen compound metabolism	1-aminocyclopropane-1-carboxylate synthase activity alanine transaminase activity	
	2	75709200	NP_002076	GPX4	glutathione peroxidase 4 isoform A precursor	3	4	mitochondrion	development electron transport phospholipid metabolism response to oxidative stress	glutathione peroxidase activity oxidoreductase activity phospholipid-hydroperoxide glutathione peroxidase activity selenium binding	
	2	90903240	NP_00103493	GPX4	glutathione peroxidase 4 isoform C precursor	4	6	mitochondrion	development electron transport phospholipid metabolism response to oxidative stress	oxidoreductase activity phospholipid-hydroperoxide glutathione peroxidase activity selenium binding epidermal growth factor receptor binding	
	2	4504111	NP_002077	GRB2	growth factor receptor-bound protein 2 isoform 3	3	4	cytosol	cell-cell signaling epidermal growth factor receptor signaling pathway insulin receptor signaling pathway intracellular signaling cascade Ras protein signal transduction	epidermal growth factor receptor binding insulin receptor substrate binding SH3/SH2 adaptor activity	
	2	45359859	NP_987102	GRB2	growth factor receptor-bound protein 2 isoform 2	2	5	cytosol	cell-cell signaling epidermal growth factor receptor signaling pathway insulin receptor signaling pathway intracellular signaling cascade Ras protein signal transduction	insulin receptor substrate binding SH3/SH2 adaptor activity glycerate dehydrogenase activity	
	2	6912396	NP_036335	GRHPR	glyoxylate reductase/hydroxypyruvate reductase	1	4	cytoplasm	electron transport excretion L-serine biosynthesis metabolism	glyoxylate reductase (NADP) activity hydroxypyruvate reductase activity NAD binding oxidoreductase activity oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor protein binding glutathione transferase activity	
	2	55770852	NP_060021	GRID1	glutamate receptor, ionotropic, delta 1	1	3	integral to membrane membrane periplasmic space (sensu Proteobacteria) postsynaptic membrane	ion transport	glutamate-gated ion channel activity ion channel activity ionotropic glutamate receptor activity receptor activity transporter activity	
	2	38044288	NP_937895	GSN	gelsolin isoform b	10	21	actin cytoskeleton cytosol extracellular region	actin filament polymerization actin filament severing barbed-end actin filament capping	actin binding calcium ion binding protein binding	Finnish Type Familial Amyloidosis [MIM:105120]

Supplementary Table I

	2	50301238	NP_000628	GSR	glutathione reductase	1	1	cytoplasm mitochondrion	electron transport glutathione metabolism	disulfide oxidoreductase activity FAD binding glutathione-disulfide reductase activity	
	2	4504169	NP_000169	GSS	glutathione synthetase	1	3	not classified	amino acid metabolism glutathione biosynthesis nervous system development response to oxidative stress	ATP binding glutathione synthase activity ligase activity nucleotide binding	Glutathione Synthetase Deficiency [MIM:266130]
	1,2	22091454	NP_665683	GSTA1	glutathione S-transferase A1	1	2	cytoplasm	glutathione metabolism metabolism	glutathione transferase activity transferase activity	
Yes	1,2	20149504	NP_000837	GSTA2	glutathione S-transferase A2	10	70	not classified	metabolism	glutathione transferase activity transferase activity	
Yes	1,2	24430144	NP_000838	GSTA3	glutathione S-transferase A3	3	28	not classified	metabolism response to stress	glutathione transferase activity transferase activity	
Yes	2	24308514	NP_714543	GSTA5	glutathione S-transferase A5	1	3	not classified	metabolism	glutathione transferase activity transferase activity	
	2	72534726	NP_001026890	GSTCD	glutathione S-transferase, C-terminal domain	1	1	not classified	not classified	methyltransferase activity transferase activity	
Yes	2	23065544	NP_000552	GSTM1	glutathione S-transferase M1 isoform 1	1	2	cytoplasm	metabolism	glutathione transferase activity transferase activity	
Yes	2	23065547	NP_666533	GSTM1	glutathione S-transferase M1 isoform 2	3	6	cytoplasm	metabolism	glutathione transferase activity transferase activity	
	2	4504175	NP_000839	GSTM2	glutathione S-transferase M2	1	1	not classified	metabolism	glutathione transferase activity transferase activity	
Yes	1,2	23065552	NP_000840	GSTM3	glutathione S-transferase M3	9	28	not classified	establishment of blood-nerve barrier metabolism	glutathione transferase activity transferase activity	
	2	4758484	NP_004823	GSTO1	glutathione-S-transferase omega 1	1	1	cytoplasm	metabolism	glutathione transferase activity monodehydroascorbate reductase (NADH) activity transferase activity	
	2	38016131	NP_899062	GSTO2	glutathione S-transferase omega 2	1	1	cytoplasm	metabolism	glutathione transferase activity transferase activity	
	1,2	4504183	NP_000843	GSTP1	glutathione transferase	9	25	cytoplasm	anti-apoptosis central nervous system development metabolism	glutathione transferase activity transferase activity	
	2	58761500	NP_037473	GTPBP9	GTP-binding protein PTD004 isoform 1	2	3	cytoplasm intracellular	not classified	GTP binding nucleotide binding	
	1,2	4504223	NP_000172	GUSB	glucuronidase, beta	5	8	lysosome	carbohydrate metabolism glycosaminoglycan catabolism metabolism	beta-glucuronidase activity cation binding	
	2	24475863	NP_722575	H1FOO	oocyte-specific histone H1	1	1	nucleosome nucleus	nucleosome assembly	DNA binding	
Yes	2	41406067	NP_958844	H2AFV	H2A histone family, member V isoform 3	1	1	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	2	4504253	NP_002096	H2AFX	H2A histone family, member X	1	1	chromosome nucleosome nucleus	cell cycle chromosome organization and biogenesis (sensu Eukaryota) DNA recombination DNA repair meiosis nucleosome assembly	DNA binding protein binding	

Supplementary Table I

Yes	2	4504347	NP_000549	HBA1	alpha 1 globin	2	11	hemoglobin complex	oxygen transport transport	heme binding iron ion binding metal ion binding oxygen binding oxygen transporter activity	
	2	4504349	NP_000509	HBB	beta globin	2	14	hemoglobin complex	nitric oxide transport oxygen transport positive regulation of nitric oxide biosynthesis transport	heme binding hemoglobin binding iron ion binding metal ion binding oxygen binding oxygen transporter activity selenium binding	
Yes	2	4504351	NP_000510	HBD	delta globin	1	5	hemoglobin complex	oxygen transport transport	heme binding iron ion binding metal ion binding oxygen binding oxygen transporter activity	
Yes	2	28302131	NP_000550	HBG1	A-gamma globin	1	2	hemoglobin complex	oxygen transport transport	heme binding iron ion binding metal ion binding oxygen binding oxygen transporter activity	
	2	5729864	NP_006611	HBS1L	HBS1-like	1	1	not classified	protein biosynthesis signal transduction	GTP binding nucleotide binding translation elongation factor activity	
	2	14149777	NP_115500	HDHD2	haloacid dehalogenase-like hydrolase doma	1	1	not classified	metabolism	metal ion binding	
	1,2	20336761	NP_057071	HEBP1	heme binding protein 1	6	8	not classified	not classified	not classified	
	2	7657603	NP_055135	HEBP2	heme binding protein 2	1	1	not classified	not classified	not classified	
	2	4504381	NP_000178	HGD	homogentisate 1,2-dioxygenase	2	3	not classified	L-phenylalanine catabolism tyrosine catabolism	homogentisate 1,2-dioxygenase activity iron ion binding metal ion binding oxidoreductase activity	
	2	4758528	NP_004703	HGS	hepatocyte growth factor-regulated tyrosine	1	1	endosome membrane	intracellular protein transport negative regulation of cell proliferation signal transduction	zinc ion binding hydrolase activity	
	2	4885413	NP_005331	HINT1	histidine triad nucleotide binding protein 1	1	1	cytoskeleton nucleus	signal transduction	protein kinase C binding zinc ion binding nucleic acid binding	
Yes	2	4885373	NP_005316	HIST1H1A	histone cluster 1, H1a	1	3	chromosome nucleosome nucleus	nucleosome assembly spermatogenesis	DNA binding	
Yes	2	4885379	NP_005312	HIST1H1E	histone cluster 1, H1e	1	1	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	2	10800130	NP_066409	HIST1H2AD	histone 1, H2ad	1	2	not classified	not classified	not classified	
Yes	2	10645195	NP_066390	HIST1H2AE	H2A histone family, member A	2	8	chromosome nucleosome	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	2	20336754	NP_066402	HIST1H2BJ	H2B histone family, member R	1	2	nucleus chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) defense response to bacterium nucleosome assembly	DNA binding	
Yes	2	4504295	NP_003527	HIST1H3H	H3 histone family, member K	4	15	not classified	not classified	not classified	

Supplementary Table I

Yes	2	4504315	NP_003536	HIST1H4E	H4 histone family, member J	8	123	not classified	not classified	not classified	
	2	28195394	NP_778235	HIST2H2AB	histone H2A	1	1	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	2	4504277	NP_003519	HIST2H2BE	histone cluster 2, H2be	5	19	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) defense response to bacterium nucleosome assembly	DNA binding	
	2	66912162	NP_001019770	HIST2H2BF	histone 2, H2bf	4	10	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	2	77539758	NP_001029249	HIST2H4B	histone cluster 2, H4b	7	34	not classified	not classified	not classified	
Yes	2	4504299	NP_003484	HIST3H3	histone cluster 3, H3	2	11	chromosome nucleosome nucleus	chromosome organization and biogenesis (sensu Eukaryota) nucleosome assembly	DNA binding	
Yes	1,2	24797067	NP_002107	HLA-A	major histocompatibility complex, class I, A	1	1	integral to membrane integral to plasma membrane membrane MHC class I protein complex	antigen processing and presentation antigen processing and presentation of endogenous antigen antigen processing and presentation of endogenous peptide antigen	MHC class I receptor activity protein binding	
	2	17986001	NP_005505	HLA-B	major histocompatibility complex, class I, B	1	2	integral to membrane integral to plasma membrane membrane membrane fraction MHC class I protein complex	antigen processing and presentation antigen processing and presentation of endogenous antigen antigen processing and presentation of endogenous peptide antigen via MHC class I	MHC class I receptor activity	
	2	52426774	NP_061984	HLA-DRA	major histocompatibility complex, class II, D	1	1	integral to plasma membrane lysosome membrane plasma membrane	antigen processing and presentation of exogenous antigen antigen processing and presentation of exogenous peptide antigen via MHC class II immune response	MHC class II receptor activity	
Yes	2	17986005	NP_072049	HLA-DRB3	major histocompatibility complex, class II, D	1	1	integral to membrane integral to plasma membrane membrane	antigen processing and presentation of exogenous antigen antigen processing and presentation of exogenous peptide antigen via MHC class II immune response signal transduction	MHC class II receptor activity	
	2	5901970	NP_008826	HNMT	histamine N-methyltransferase isoform 1	1	1	not classified	respiratory gaseous exchange	histamine N-methyltransferase activity methyltransferase activity transferase activity	Susceptibility to Asthma [MIM:600807]

Supplementary Table I

	2	52632385	NP_001005338	HNRNPL	heterogeneous nuclear ribonucleoprotein L	1	1	heterogeneous nuclear ribonucleoprotein complex nucleoplasm nucleus	mRNA processing	nucleotide binding RNA binding ATP binding	
Yes	2	14110428	NP_112604	HNRPC	heterogeneous nuclear ribonucleoprotein C	1	1	nucleus spliceosome complex	nuclear mRNA splicing, via spliceosome RNA splicing	identical protein binding nucleotide binding RNA binding	
	2	4826762	NP_005134	HP	haptoglobin	2	7	extracellular region	defense response iron ion homeostasis proteolysis	hemoglobin binding serine-type endopeptidase activity	
Yes	2	19913446	NP_002134	HPCA	hippocalcin	1	1	not classified	not classified	actin binding calcium ion binding	
Yes	2	19913441	NP_002140	HPCAL1	hippocalcin-like 1	1	2	not classified	not classified	calcium ion binding	
	2	4504477	NP_002141	HPD	4-hydroxyphenylpyruvate dioxygenase	1	1	not classified	aromatic amino acid family metabolism L-phenylalanine catabolism tyrosine catabolism	4-hydroxyphenylpyruvate dioxygenase activity iron ion binding metal ion binding oxidoreductase activity	Tyrosinemia Type III [MIM:276710]
	1,2	31542939	NP_000851	HPGD	hydroxyprostaglandin dehydrogenase 15-(NAD)	6	21	cytosol	fatty acid metabolism lipid metabolism lipoxygenase pathway metabolism negative regulation of progression through cell cycle parturition pregnancy prostaglandin metabolism transforming growth factor beta receptor signaling pathway	15-hydroxyprostaglandin dehydrogenase (NAD+) activity NAD binding oxidoreductase activity prostaglandin E receptor activity protein homodimerization activity	Hypertension [MIM:601688]
Yes	2	45580723	NP_066275	HPR	haptoglobin-related protein	1	2	extracellular region	proteolysis	hemoglobin binding serine-type endopeptidase activity	
	2	11321561	NP_000604	HPX	hemopexin	1	1	extracellular region extracellular space	heme transport iron ion homeostasis transport	binding heme transporter activity iron ion binding metal ion binding	
Yes	2	34222246	NP_789765	HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene	2	23	cytoplasm intracellular membrane plasma membrane	cell motility cell surface receptor linked signal transduction chemotaxis organ morphogenesis regulation of progression through cell cycle small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein C-terminus binding	
	2	4504489	NP_000403	HRG	histidine-rich glycoprotein precursor	2	5	extracellular region	not classified	cysteine protease inhibitor activity heparin binding	
Yes	2	57864582	NP_001009931	HRNR	HRNR	1	1	not classified	development keratinization	calcium ion binding structural molecule activity	
	1,2	5032215	NP_005827	HRSP12	heat-responsive protein 12	2	4	cytoplasm nucleus	regulation of translational termin	endonuclease activity hydrolase activity	

Supplementary Table I

Yes	1,2	63029937	NP_001017963	HSP90AA1	heat shock protein 90kDa alpha (cytosolic)	4	7	cytosol	mitochondrial transport positive regulation of nitric oxide biosynthesis protein folding protein refolding response to unfolded protein signal transduction	ATP binding nitric-oxide synthase regulator activity nucleotide binding protein homodimerization activity TPR domain binding unfolded protein binding	
Yes	2	92859630	NP_001035231	HSP90AA2	heat shock 90kDa protein 1, alpha-like 3	4	10	not classified	protein folding response to unfolded protein	ATP binding unfolded protein binding	
Yes	1,2	20149594	NP_031381	HSP90AB1	heat shock 90kDa protein 1, beta	4	4	cytoplasm	positive regulation of nitric oxide biosynthesis protein folding response to unfolded protein	ATP binding nitric-oxide synthase regulator activity nucleotide binding TPR domain binding unfolded protein binding	
Yes	2	62122861	NP_001014390	HSP90AB2P	heat shock protein HSP 90-beta	1	2	not classified	protein folding response to unfolded protein	ATP binding nucleotide binding unfolded protein binding	
	2	62198220	NP_001014441	HSP90AB6P	heat shock protein 90Bf	1	1	not classified	protein folding response to unfolded protein	unfolded protein binding	
Yes	2	4507677	NP_003290	HSP90B1	HSP90B1	1	3	cytosol endoplasmic reticulum endoplasmic reticulum lumen endoplasmic reticulum membrane microsome perinuclear region	anti-apoptosis protein folding protein transport response to hypoxia sequestering of calcium ion	ATP binding calcium ion binding low-density lipoprotein receptor binding RNA binding unfolded protein binding viral binding	
	2	119874213	NP_079291	HSPA12A	heat shock 70kDa protein 12A	1	1	not classified	not classified	ATP binding nucleotide binding	
Yes	1,2	5123454	NP_005336	HSPA1A	heat shock 70kDa protein 1A	19	53	not classified	protein folding response to unfolded protein	ATP binding nucleotide binding protein binding	
Yes	2	124256496	NP_005518	HSPA1L	heat shock 70kDa protein 1-like	11	59	not classified	protein folding response to unfolded protein	ATP binding nucleotide binding	
	2	13676857	NP_068814	HSPA2	heat shock 70kDa protein 2	16	63	cell surface	male meiosis protein folding response to unfolded protein spermatid development	ATP binding nucleotide binding unfolded protein binding	
	1,2	16507237	NP_005338	HSPA5	heat shock 70kDa protein 5	1	1	caspase complex cell surface endoplasmic reticulum endoplasmic reticulum lumen ER-Golgi intermediate compartment integral to endoplasmic reticulum membrane nucleus perinuclear region	anti-apoptosis negative regulation of caspase activity	ATP binding calcium ion binding caspase inhibitor activity nucleotide binding protein binding, bridging unfolded protein binding	
Yes	1,2	34419635	NP_002146	HSPA6	heat shock 70kDa protein 6 (HSP70B')	6	31	not classified	protein folding response to unfolded protein	ATP binding nucleotide binding	

Supplementary Table 1

	1,2	5729877	NP_006588	HSPA8	heat shock 70kDa protein 8 isoform 1	19	49	cell surface intracellular nucleus	protein folding response to unfolded protein	ATP binding ATPase activity, coupled nucleotide binding protein binding	
Yes	2	24234686	NP_694881	HSPA8	heat shock 70kDa protein 8 isoform 2	1	1	cell surface intracellular nucleus	protein folding response to unfolded protein	ATP binding ATPase activity, coupled nucleotide binding protein binding	
	2	4504517	NP_001531	HSPB1	heat shock 27kDa protein 1	8	30	cell surface cytoplasm cytoskeleton	anti-apoptosis cell motility protein folding regulation of translational initiation response to unfolded protein	protein binding	Charcot-Marie-Tooth Disease, Type 2F [MIM:606595] Distal Hereditary Motor Neuropathy IIB [MIM:608634]
	1,2	41399285	NP_955472	HSPD1	chaperonin	1	1	cytoplasm mitochondrion	cellular protein metabolism protein folding protein import into mitochondrial matrix regulation of apoptosis response to unfolded protein	ATP binding chaperone binding nucleotide binding protein binding unfolded protein binding	
	2	62859979	NP_005520	HSPG2	heparan sulfate proteoglycan 2	18	41	basement membrane extracellular space	cell adhesion	not classified	Schwartz-Jampel Syndrome Type 1 [MIM: 255800]
	2	61676188	NP_113584	HUWE1	HECT, UBA and WWE domain containing 1	1	1	cytoplasm intracellular nucleus	histone ubiquitination mRNA transport protein polyubiquitination ubiquitin cycle	binding DNA binding ligase activity protein binding ubiquitin-protein ligase activity	
	2	4557878	NP_000192	ICAM1	intercellular adhesion molecule 1 precursor	1	1	integral to plasma membrane plasma membrane	cell-cell adhesion	transmembrane receptor activity insulin-like growth factor receptor activity	Graves Disease [MIM: 275000]
	1,2	28178825	NP_005887	IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	14	38	cytoplasm cytosol peroxisome	carbohydrate metabolism glyoxylate cycle isocitrate metabolism main pathways of carbohydrate metabolism metabolism tricarboxylic acid cycle	isocitrate dehydrogenase (NADP+) activity oxidoreductase activity oxidoreductase activity, acting on NADH or NADPH, NAD or NADP as acceptor	
Yes	2	28178832	NP_002159	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial	1	4	mitochondrion	carbohydrate metabolism glyoxylate cycle main pathways of carbohydrate metabolism metabolism tricarboxylic acid cycle	isocitrate dehydrogenase (NADP+) activity oxidoreductase activity oxidoreductase activity, acting on NADH or NADPH, NAD or NADP as acceptor	
Yes	2	10835238	NP_006426	IFITM2	interferon induced transmembrane protein 2	1	6	integral to membrane membrane	immune response response to biotic stimulus	protein binding	
Yes	2	148612842	NP_066362	IFITM3	interferon-induced transmembrane protein 3	1	7	integral to membrane plasma membrane	immune response response to biotic stimulus	not classified	
	2	119964726	NP_000867	IGF2R	insulin-like growth factor 2 receptor precursor	1	1	endosome integral to plasma membrane lysosome membrane trans-Golgi network transport vesicle	receptor mediated endocytosis signal transduction transport	receptor activity transporter activity metal ion binding	
	2	4826772	NP_004961	IGFALS	insulin-like growth factor binding protein, acidic	6	12	extracellular region extracellular space soluble fraction	cell adhesion signal transduction	insulin-like growth factor binding protein binding transferase activity	

Supplementary Table 1

	2	116812565	NP_071911	IKZF5	zinc finger protein, subfamily 1A, 5	1	1	intracellular nucleus	not classified	nucleic acid binding zinc ion binding interleukin-1 receptor antagonist activity	
	2	27894319	NP_776214	IL1RN	interleukin 1 receptor antagonist isoform 1 precursor	2	2	extracellular region extracellular space intracellular	inflammatory response	interleukin-1 receptor antagonist activity	Gastic Cancer Risk [MIM:137215]
	2	27894321	NP_776215	IL1RN	interleukin 1 receptor antagonist isoform 4	1	1	extracellular region extracellular space intracellular	inflammatory response	calcium ion binding	
Yes	2	28610147	NP_002175	IL6ST	interleukin 6 signal transducer isoform 1 precursor	1	2	integral to plasma membrane plasma membrane	cell surface receptor linked signal transduction immune response	interleukin-6 receptor activity oncostatin-M receptor activity receptor activity	
	2	29029546	NP_795352	INADL	InaD-like protein	7	11	membrane tight junction	not classified	protein binding	
	2	14165288	NP_115639	IQCG	IQ motif containing G	1	2	not classified	not classified	not classified	
	2	4506787	NP_003861	IQGAP1	IQ motif containing GTPase activating protein	20	40	actin filament intracellular membrane	negative regulation of Ras protein signal transduction regulation of small GTPase mediated signal transduction signal transduction	calmodulin binding GTPase inhibitor activity Ras GTPase activator activity	
	2	5729887	NP_006624	IQGAP2	IQ motif containing GTPase activating protein	5	6	actin cytoskeleton intracellular	regulation of small GTPase mediated signal transduction signal transduction	actin binding calmodulin binding GTPase inhibitor activity Ras GTPase activator activity	
	2	5453700	NP_006138	IRF6	interferon regulatory factor 6	1	1	nucleus	regulation of transcription, DNA- dependent transcription	transcription factor activity	Van der Woude Syndrome [MIM:119300] Popliteal Pterygium Syndrom [MIM:119500]
	2	31657142	NP_852478	ITGA1	integrin, alpha 1 precursor	1	1	integral to membrane integrin complex membrane	cell adhesion cell-matrix adhesion integrin-mediated signaling pathway	calcium ion binding collagen binding magnesium ion binding receptor activity	
	2	4504747	NP_002195	ITGA3	integrin alpha 3 isoform a precursor	1	1	integral to membrane integrin complex membrane	cell adhesion cell-matrix adhesion integrin-mediated signaling pathway	calcium ion binding protein binding receptor activity	
	2	6006011	NP_005492	ITGA3	integrin alpha 3 isoform b, precursor	1	1	integral to membrane integrin complex membrane	cell adhesion cell-matrix adhesion integrin-mediated signaling pathway	protein binding receptor activity protein binding	
	2	4504763	NP_002201	ITGAV	integrin alpha-V precursor	1	1	integral to membrane integrin complex membrane	cell adhesion cell-matrix adhesion integrin-mediated signaling pathway	calcium ion binding protein binding receptor activity	
	2	19743813	NP_002202	ITGB1	integrin beta 1 isoform 1A precursor	1	1	endoplasmic reticulum integral to membrane integrin complex membrane ruffle	cell adhesion cell migration cell-matrix adhesion cellular defense response development homophilic cell adhesion integrin-mediated signaling pathway	identical protein binding protein binding protein heterodimerization activity receptor activity	
	2	31563378	NP_852133	ITGB4BP	integrin beta 4 binding protein isoform a	1	1	nucleus	integrin-mediated signaling pathway mature ribosome assembly protein biosynthesis	protein binding ribosome binding translation initiation factor activity	

Supplementary Table I

	2	4504779	NP_002205	ITGB8	integrin, beta 8 precursor	1	1	integral to membrane integrin complex membrane	cell adhesion cell-matrix adhesion integrin-mediated signaling pahtwya	receptor activity not classified	
	2	31542984	NP_002209	ITIH4	inter-alpha (globulin) inhibitor H4	2	5	extracellular region	acute-phase response hyaluronan metabolism	serine-type endopeptidase inhibitor activity	
	2	31542986	NP_060095	ITLN1	intelectin	4	7	membrane	not classified	GPI anchor binding sugar binding	
	1,2	11527402	NP_068839	ITM2B	integral membrane protein 2B	5	21	integral to plasma membrane membrane membrane fraction	nervous system development sensory perception of sound	not classified	Familial Dementia [MIM:176500]
	2	13569885	NP_112188	ITM2C	integral membrane protein 2C isoform 1	1	1	Golgi apparatus integral to membrane membrane	not classified	not classified	
	2	60302914	NP_00101253	ITM2C	integral membrane protein 2C isoform 2	1	1	Golgi apparatus integral to membrane membrane	not classified endocytosis regulation of Rho protein signal transduction	calcium ion binding	
Yes	2	47717123	NP_003015	ITSN1	intersectin 1 isoform ITSN-1	1	1	intracellular	regulation of Rho protein signal transduction synaptic vesicle endocytosis	calcium ion binding protein binding Rho guanyl-nucleotide exchange factor activity	
	2	22325385	NP_006268	ITSN2	intersectin 2 isoform 1	1	1	intracellular	endocytosis regulation of Rho protein signal transduction	protein binding Rho guanyl-nucleotide exchange factor activity SH3/SH2 adaptor activity cytoskeletal protein binding	
	2	12056468	NP_068831	JUP	junction plakoglobin	9	15	cytoplasm cytoskeleton membrane fraction soluble fraction	cell adhesion	cytoskeletal protein binding structural molecule activity	Naxos Disease [MIM:601214]
	2	68362740	NP_008995	KALRN	kalirin, RhoGEF kinase isoform 3	1	1	actin cytoskeleton intracellular	intracellular signaling cascade nervous system development protein amino acid phosphorylation regulation of Rho protein signal transduction signal transduction vesicle-mediated transport	ATP binding nucleotide binding protein binding protein serine/threonine kinase activity Rho guanyl-nucleotide exchange factor activity transferase activity	Coronary Heart Disease [MIM:608901]
	2	6912444	NP_036415	KCNG2	potassium voltage-gated channel, subfamily 1	1	2	integral to membrane membrane voltage-gated potassium channel complex	ion transport potassium ion transport regulation of heart contraction	delayed rectifier potassium channel activity potassium ion binding protein binding	
	2	4557693	NP_000212	KHK	ketoheksokinase isoform a	2	4	cytoplasm	carbohydrate metabolism	ketoheksokinase activity kinase activity transferase activity	Essential Fructosuria [MIM:229800]
	2	5670342	NP_006479	KHK	ketoheksokinase isoform b	1	1	cytoplasm	carbohydrate metabolism	ketoheksokinase activity kinase activity transferase activity	
	1,2	41281489	NP_055576	KIAA0174	putative MAPK activating protein PM28	19	229	ER-Golgi intermediate compartment	not classified	protein binding	
	2	14767307	XP_049237	KIAA0841	PREDICTED: similar to Protein KIAA0841 is	1	1	not classified	not classified	not classified	
	2	58331177	NP_00100995	KIAA1189	hypothetical protein LOC57471 isoform a	1	1	cytoplasm membrane	not classified	not classified	
	2	55741661	NP_065944	KIAA1529	hypothetical protein LOC57653	1	5	not classified	not classified	not classified	

Supplementary Table I

	2	113420564	XP_946410	KIAA1833	PREDICTED: similar to c11.1 CG12132-PA	2	1	not classified	not classified	ATP binding	
	2	19923949	NP_612433	KIF12	kinesin family member 12	16	39	microtubule microtubule associated complex	microtubule-based movement	ATP binding microtubule motor activity nucleotide binding	
Yes	2	46852172	NP_056069	KIF13B	kinesin family member 13B	1	3	cytoplasm cytoskeleton microtubule microtubule associated complex	microtubule-based movement protein targeting signal transduction T cell activation	ATP binding microtubule motor activity nucleotide binding protein binding protein kinase binding	
	2	46852174	NP_008985	KIF3A	kinesin family member 3A	1	1	microtubule microtubule associated complex	microtubule-based movement organelle organization and biogenesis	ATP binding microtubule motor activity nucleotide binding	
	2	4758646	NP_004789	KIF3B	kinesin family member 3B	1	1	microtubule microtubule associated complex plus-end kinesin complex	anterograde axon cargo transport determination of left/right symmetry microtubule-based movement	nucleotide binding plus-end-directed microtubule motor activity ATP binding	
Yes	2	45446749	NP_004975	KIF5A	kinesin family member 5A	3	6	kinesin complex membrane fraction microtubule	microtubule-based movement synaptic transmission	ATP binding microtubule motor activity nucleotide binding	
Yes	2	4758648	NP_004512	KIF5B	kinesin family member 5B	1	1	kinesin complex microtubule	microtubule-based movement	ATP binding microtubule motor activity nucleotide binding	
	2	33589863	NP_878906	KIF9	kinesin family member 9 isoform 3	1	1	microtubule microtubule associated complex	microtubule-based movement	microtubule motor activity nucleotide binding ATP binding	
	2	94536858	NP_005541	KIFC3	kinesin family member C3	1	1	kinesin complex microtubule microtubule associated complex	microtubule-based movement visual perception	microtubule motor activity nucleotide binding beta-glucuronidase activity	
	2	24497614	NP_004786	KL	klotho	1	1	integral to membrane membrane	aging carbohydrate metabolism metabolism positive regulation of bone mineralization	cation binding glucosidase activity hormone activity vitamin D binding IgE binding	Hyperphosphatemic Tumoral Calcinosis [MIM: 211900]
	2	4504875	NP_002248	KLK1	kallikrein 1 preproprotein	1	1	not classified	proteolysis	peptidase activity tissue kallikrein activity	Decreased Urinary Activity of Kallikrein [MIM:147910]
	1,2	4504893	NP_000884	KNG1	kininogen 1	17	81	extracellular region	blood coagulation diuresis inflammatory response natriuresis negative regulation of blood coagulation negative regulation of cell adhesion positive regulation of apoptosis smooth muscle contraction vasodilation	cysteine protease inhibitor activity heparin binding receptor binding zinc ion binding	
	2	19923142	NP_002256	KPNB1	karyopherin beta 1	2	3	cytoplasm nuclear pore nucleus	NLS-bearing substrate import into nucleus protein import into nucleus, docking protein import into nucleus, translocation protein transport	nuclear localization sequence binding protein binding protein transporter activity zinc ion binding	

Supplementary Table 1

Yes	2	15718761	NP_004976	KRAS	c-K-ras2 protein isoform b	3	5	intracellular membrane	regulation of progression through cell cycle sensory perception of sound small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
Yes	2	15718763	NP_203524	KRAS	c-K-ras2 protein isoform a	1	1	intracellular	regulation of progression through cell cycle sensory perception of sound small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	2	21264602	NP_005551	LAMA5	laminin alpha 5	1	2	basal lamina extracellular matrix (sensu Metazoa) integral to membrane laminin-1 membrane	cell adhesion regulation of cell adhesion regulation of cell migration regulation of embryonic development	protein binding receptor binding structural molecule activity	
	2	9845498	NP_002284	LAMC1	laminin, gamma 1 precursor	1	2	basement membrane laminin-1	cell adhesion endoderm development positive regulation of epithelial cell proliferation protein complex assembly	extracellular matrix structural constituent protein binding	
	1,2	7669501	NP_005552	LAMP1	lysosomal-associated membrane protein 1	3	25	integral to plasma membrane late endosome lysosome membrane membrane fraction	not classified	not classified	
	2	7669503	NP_054701	LAMP2	lysosomal-associated membrane protein 2 precursor	4	26	integral to membrane late endosome lysosomal membrane membrane membrane fraction platelet dense granule membrane	not classified	not classified	Danon Disease [MIM:300257]
	2	41393561	NP_056991	LAP3	leucine aminopeptidase 3	1	1	cytoplasm intracellular	protein metabolism proteolysis	leucyl aminopeptidase activity magnesium ion binding manganese ion binding peptidase activity prolyl aminopeptidase activity zinc ion binding	
	2	31652249	NP_004130	LBP	lipopolysaccharide-binding protein precursor	2	4	extracellular region extracellular space integral to membrane membrane	acute-phase response cellular defense response defense response to bacterium lipid transport	lipid binding	

Supplementary Table I

Yes	2	112789548	NP_001036236	LCK	lymphocyte-specific protein tyrosine kinase	1	4	lipid raft pericentriolar material plasma membrane	caspase activation hemopoiesis induction of apoptosis intracellular signaling cascade positive regulation of T cell activation positive regulation of T cell receptor signaling pathway protein amino acid phosphorylation Ras protein signal transduction regulation of lymphocyte activation regulation of progression through cell cycle release of sequestered calcium ion into cytosol response to drug T cell differentiation zinc ion homeostasis	ATP binding ATPase binding CD4 receptor binding CD8 receptor binding glycoprotein binding nucleotide binding phosphoinositide 3-kinase binding protein C-terminus binding protein kinase binding protein serine/threonine phosphatase activity protein-tyrosine kinase activity SH2 domain binding transferase activity	
	2	38455402	NP_005555	LCN2	lipocalin 2	2	4	cytoplasm soluble fraction	transport	binding transporter activity	
	2	4504965	NP_002289	LCP1	L-plastin	1	1	cytoplasm cytosol	not classified	actin binding calcium ion binding protein binding	
	1,2	5031857	NP_005557	LDHA	lactate dehydrogenase A	6	24	cytoplasm cytosol	anaerobic glycolysis tricarboxylic acid cycle intermediate metabolism	L-lactate dehydrogenase activity oxidoreductase activity protein binding	
Yes	2	47059044	NP_659409	LDHAL6A	lactate dehydrogenase A-like 6A	1	4	cytoplasm	anaerobic glycolysis tricarboxylic acid cycle intermediate metabolism	L-lactate dehydrogenase activity oxidoreductase activity	
Yes	2	15082234	NP_149972	LDHAL6B	lactate dehydrogenase A-like 6B	1	14	cytoplasm	anaerobic glycolysis tricarboxylic acid cycle intermediate metabolism	L-lactate dehydrogenase activity oxidoreductase activity	
	1,2	4557032	NP_002291	LDHB	lactate dehydrogenase B	14	95	cytoplasm	anaerobic glycolysis tricarboxylic acid cycle intermediate metabolism	L-lactate dehydrogenase activity oxidoreductase activity	
	2	21361918	NP_071751	LEPRE1	leucine proline-enriched proteoglycan (lepre	1	2	endoplasmic reticulum	protein metabolism	binding iron ion binding L-ascorbic acid binding metal ion binding oxidoreductase activity oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, each of oxygen into both donors oxidoreductase activity, actin	
	2	115430223	NP_002297	LGALS3	galectin 3	1	1	cytoplasm extracellular region nucleus plasma membrane	not classified	sugar binding not classified	Lymphocyte Function-Associated Antigen 1 [MIM: 116920]
	1,2	5031863	NP_005558	LGALS3BP	galectin 3 binding protein	26	201	extracellular matrix (sensu Metazoa) extracellular space membrane	cell adhesion cellular defense response signal transduction	protein binding scavenger receptor activity	
Yes	2	4504985	NP_002298	LGALS7	galectin 7	1	4	extracellular space nucleus	apoptosis heterophilic cell adhesion	sugar binding	

Supplementary Table 1

Yes	2	13324690	NP_077294	LILRA6	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 6	1	1	not classified	not classified	not classified	
	1,2	5803023	NP_006807	LMAN2	lectin, mannose-binding 2	10	16	endoplasmic reticulum ER-Golgi intermediate compartment integral to membrane membrane	protein transport	calcium ion binding sugar binding	
	2	21687060	NP_660295	LOC124220	hypothetical protein LOC124220	4	6	not classified	not classified	not classified	
	2	41107482	XP_060887	LOC128192	PREDICTED: similar to peptidylprolyl isom	1	2	not classified	not classified	not classified	
	2	88969090	XP_946711	LOC131691	PREDICTED: similar to peptidylprolyl isom	1	4	not classified	not classified	not classified	
	2	51474268	NP_00107391	LOC146909	hypothetical protein LOC146909	1	3	not classified	not classified	not classified	
	2	32401447	NP_861450	LOC283537	hypothetical protein LOC283537	1	1	not classified	not classified	not classified	
Yes	2	42661306	XP_375443	LOC284100	PREDICTED: similar to 14-3-3 protein epsilon (14-3-3E) (Mitochondrial import stimulation factor L subunit) (MSF L)	1	2	not classified	not classified	not classified	
	2	27500852	XP_209196	LOC284422	PREDICTED: hypothetical protein	1	2	not classified	not classified	not classified	
	2	47825361	NP_00100141	LOC342897	similar to F-box only protein 2	2	2	not classified	protein catabolism	protein binding receptor activity	
	2	113417317	XP_946024	LOC345630	PREDICTED: similar to fibrillarin	2	1	not classified	not classified	not classified	
Yes	2	51493085	XP_370729	LOC387934	PREDICTED: similar to Fatty acid-binding protein, epidermal (E-FABP) (Psoriasis-associated fatty acid-binding protein homolog) (PA-FABP)	2	4	not classified	not classified	not classified	
	2	51458856	XP_371330	LOC388720	PREDICTED: similar to ubiquitin and riboso	1	2	not classified	not classified	not classified	
	2	56847632	NP_00100874	LOC388817	peptidylprolyl isomerase A	5	28	not classified	protein folding	not classified	
	2	113429505	XP_948564	LOC388849	PREDICTED: hypothetical protein	1	1	not classified	not classified	not classified	
Yes	2	51466608	XP_372002	LOC389599	PREDICTED: similar to amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 2	1	1	not classified	not classified	not classified	
	2	113423709	XP_945177	LOC390299	PREDICTED: similar to peptidylprolyl isom	1	1	not classified	not classified	not classified	
	2	89040598	XP_950835	LOC440335	PREDICTED: hypothetical protein isoform 4	1	1	not classified	not classified	not classified	
Yes	2	122937516	NP_00107400	LOC441282	hypothetical protein LOC441282	1	4	not classified	not classified	not classified	
Yes	2	113419293	XP_941326	LOC642103	PREDICTED: similar to Maltase-glucoamyla	9	24	not classified	not classified	not classified	
	2	89042981	XP_951274	LOC643008	PREDICTED: hypothetical protein LOC6430	1	1	not classified	not classified	not classified	
	2	148233728	NP_00108306	LOC643751	hypothetical protein LOC643751	1	4	not classified	not classified	not classified	
	2	148227764	NP_00108317	LOC643752	hypothetical protein LOC643752	1	3	not classified	not classified	not classified	
	2	88951501	XP_942241	LOC646821	PREDICTED: similar to cytoplasmic beta-ac	1	1	not classified	not classified	not classified	
	2	89057283	XP_943300	LOC649125	PREDICTED: similar to actin-related protei	1	1	not classified	not classified	not classified	
	2	89062025	XP_945805	LOC651536	PREDICTED: similar to immunoglobulin iote	1	1	not classified	not classified	transferase activity	
Yes	2	89037862	XP_945963	LOC651665	PREDICTED: similar to retrotransposon-like	1	1	not classified	not classified	not classified	
	2	51472697	XP_290660	LOC654259	PREDICTED: similar to Programmed cell death 6-interacting protein (ALG-2-interacting protein X) (ALG-2-interacting protein 1) (E2F1-inducible protein) (Eig2)	6	33	not classified	not classified	not classified	
Yes	2	113412128	XP_00113425	LOC732165	PREDICTED: similar to Triosephosphate isc	1	1	not classified	not classified	not classified	
	2	13569913	NP_112203	LOC81691	exonuclease NEF-sp	1	1	intracellular nucleolus	not classified	exonuclease activity nucleic acid binding nucleotide binding	
	1,2	6806919	NP_004516	LRP2	low density lipoprotein-related protein 2	84	470	coated pit integral to membrane lysosome membrane	lipid metabolism protein amino acid glycosylation receptor mediated endocytosis	calcium ion binding protein binding receptor activity	
	2	18677767	NP_570843	LRRC15	leucine rich repeat containing 15	1	1	integral to membrane membrane	not classified	transferase activity	

Supplementary Table I

	2	23397554	NP_694992	LRRRC57	leucine rich repeat containing 57	1	1	not classified	not classified	protein binding	
								not classified	protein amino acid phosphorylation small GTPase mediated signal transduction	ATP binding binding GTP binding GTPase activator activity nucleotide binding protein serine/threonine kinase activity transferase activity	Parkinson Disease [MIM:607060]
	2	83722282	NP_940980	LRRK2	leucine-rich repeat kinase 2	4	5				
								not classified	inflammatory response leukotriene biosynthesis proteolysis	epoxide hydrolase activity membrane alanyl aminopeptidase activity metal ion binding metallopeptidase activity peptidase activity zinc ion binding	
	2	4505029	NP_000886	LTA4H	leukotriene A4 hydrolase	6	9				
								cytoplasm	leukotriene metabolism	15-oxoprostaglandin 13-oxidase activity alcohol dehydrogenase activity oxidoreductase activity zinc ion binding	
	1,2	28570172	NP_036344	LTB4DH	NADP-dependent leukotriene B4 12-hydrox	5	10				
Yes	2	54607120	NP_002334	LTF	lactotransferrin	1	3	extracellular region extracellular space	defense response to bacterium humoral immune response ion transport iron ion homeostasis iron ion transport	ferric iron binding metal ion binding peptidase activity serine-type endopeptidase activity	
	2	9966764	NP_009191	LYPLA2	lysophospholipase II	1	1	not classified	fatty acid metabolism lipid metabolism	hydrolase activity	
								extracellular region	cell wall catabolism cytolysis defense response to bacterium metabolism	hydrolase activity, acting on glycosyl bonds lysozyme activity	Familial Visceral Amyloidosis [MIM:105200]
	2	4557894	NP_000230	LYZ	lysozyme precursor	1	3				
								membrane	apoptosis intracellular signaling cascade	guanylate kinase activity kinase activity protein binding	
	2	27544925	NP_066016	MAGI3	membrane-associated guanylate kinase-rela	2	2				
	2	16418397	NP_443118	MAL2	MAL2 proteolipid protein	1	1	integral to membrane membrane	not classified	ferric iron binding	
								endoplasmic reticulum ER-Golgi intermediate compartment integral to membrane membrane membrane fraction	metabolism protein amino acid glycosylation	calcium ion binding hydrolase activity, acting on glycosyl bonds mannosidase activity mannosyl-oligosaccharide 1,2-alpha-mannosidase activity	
	1,2	24497519	NP_005898	MAN1A1	mannosidase, alpha, class 1A, member 1	15	63				
								Golgi apparatus integral to membrane membrane membrane fraction	cell motility carbohydrate metabolism glycoprotein biosynthesis mannose metabolism metabolism	alpha-mannosidase activity hydrolase activity, acting on glycosyl bonds mannosyl-oligosaccharide 1,3-1,6-alpha-mannosidase activity	
	2	51477714	NP_002363	MAN2A1	mannosidase, alpha, class 2A, member 1	1	1				
	2	11496277	NP_068805	MAP2K1IP1	mitogen-activated protein kinase kinase 1 in	1	1	membrane	not classified	not classified	
Yes	2	66932916	NP_002736	MAPK1	mitogen-activated protein kinase 1	1	1	not classified	cell cycle chemotaxis induction of apoptosis protein amino acid phosphorylation response to stress signal transduction synaptic transmission	ATP binding MAP kinase activity nucleotide binding protein serine/threonine kinase activity transferase activity	
	2	11125772	NP_002347	MARCKS	myristoylated alanine-rich protein kinase C s	4	6	actin cytoskeleton plasma membrane	cell motility	actin filament binding calmodulin binding	

Supplementary Table 1

								extracellular region	antimicrobial humoral response (sensu Vertebrata) complement activation, classical pathway innate immune response proteolysis	calcium ion binding serine-type endopeptidase activity	
	1,2	21264363	NP_006601	MASP2	mannan-binding lectin serine protease 2 iso	12	124				
	2	38202209	NP_060798	MBD5	methyl-CpG binding domain protein 5	1	1	nucleus	not classified	DNA binding	
	1,2	5174539	NP_005908	MDH1	cytosolic malate dehydrogenase	5	18	cytosol	glycolysis malate metabolism tricarboxylic acid cycle	L-lactate dehydrogenase activity L-malate dehydrogenase activity malic enzyme activity oxidoreductase activity	
	2	21735621	NP_005909	MDH2	mitochondrial malate dehydrogenase precu	1	1	mitochondrial matrix mitochondrion	glycolysis malate metabolism tricarboxylic acid cycle	L-lactate dehydrogenase activity L-malate dehydrogenase activity oxidoreductase activity	
	2	46195707	NP_001401	MEGF8	EGF-like-domain, multiple 4	2	4	integral to membrane membrane	development	calcium ion binding receptor activity structural molecule activity	
	2	5174557	NP_005919	MFGE8	milk fat globule-EGF factor 8 protein	2	2	lipid particle	cell adhesion fertilization (sensu Metazoa)	protein binding	
	2	5174559	NP_005920	MF12	melanoma-associated antigen p97 isoform	2	3	extracellular region integral to plasma membrane membrane	ion transport iron ion homeostasis iron ion transport	ferric iron binding GPI anchor binding metal ion binding zinc ion binding	
	2	16163666	NP_201573	MF12	melanoma-associated antigen p97 isoform	2	3	extracellular region integral to plasma membrane membrane	ion transport iron ion homeostasis iron ion transport	GPI anchor binding metal ion binding zinc ion binding cytokine activity	
Yes	1,2	4758712	NP_004659	MGAM	maltase-glucoamylase	39	114	integral to membrane membrane	carbohydrate metabolism metabolism starch catabolism	alpha-glucosidase activity glucan 1,4-alpha-glucosidase activity hydrolase activity, hydrolyzing O-glycosyl compounds	
Yes	2	113419289	XP_00113016	MGAM	PREDICTED: similar to Maltase-glucoamyla	1	3	not classified	not classified	not classified	
Yes	2	4505163	NP_002399	MGAT2	mannosyl (alpha-1,6-)-glycoprotein beta-1,2	1	1	Golgi apparatus Golgi stack integral to membrane membrane membrane fraction	oligosaccharide biosynthesis protein amino acid N-linked glycosylation	alpha-1,6-mannosylglycoprotein 2-beta-N- acetylglucosaminyltransferase activity transferase activity, transferring glycosyl groups	
	2	31543193	NP_689672	MGC45438	hypothetical protein LOC146556	1	1	not classified	not classified	not classified	
Yes	2	113427454	XP_943120	MGC57346	PREDICTED: similar to ADP-ribosylation fac	1	4	not classified	not classified	not classified	
	2	39752687	NP_945352	MGC59937	hypothetical protein LOC375791	1	1	not classified	not classified	not classified	
	2	44917608	NP_056061	MGRN1	mahogunin, ring finger 1	1	2	not classified	ubiquitin cycle	ligase activity metal ion binding protein binding zinc ion binding	
	2	16445409	NP_438112	MID2	midline 2 isoform 2	1	2	intracellular microtubule microtubule associated complex	not classified	metal ion binding protein binding zinc ion binding	
	2	4505185	NP_002406	MIF	macrophage migration inhibitory factor (glycosy	1	6	extracellular region extracellular space	cell proliferation cell surface receptor linked signal transduction inflammatory response negative regulation of apoptosis prostaglandin biosynthesis regulation of macrophage activation	isomerase activity phenylpyruvate tautomerase activity protein binding ATP binding	Rheumatoid Arthritis [MIM: 604302]

Supplementary Table I

									development JNK cascade protein amino acid phosphorylation protein kinase cascade response to stress	nucleotide binding protein serine/threonine kinase activity small GTPase regulator activity transferase activity metal ion binding	
	2	68161546	NP_001020108	MINK1	misshapen/NIK-related kinase isoform 4	2	3	not classified			
	1,2	20270349	NP_620153	MITD1	MIT, microtubule interacting and transport, domain 1	9	19	not classified	not classified	not classified	
	2	4758720	NP_004520	MLLT3	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3	1	1	nucleus	regulation of transcription, DNA-dependent transcription	not classified	
	2	5174575	NP_005927	MLLT4	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 4 isoform 3	1	1	intercellular junction myosin	cell adhesion cell-cell signaling signal transduction	motor activity protein binding protein C-terminus binding	
	1,2	6042202	NP_009219	MME	membrane metallo-endopeptidase	48	311	integral to plasma membrane plasma membrane	cell-cell signaling proteolysis	metal ion binding metallopeptidase activity neprilysin activity zinc ion binding	Hypertension Important cell surface marker in the diagnostic of human acute lymphocytic leukemia [MIM:120520]
	2	5729929	NP_006681	MMP24	matrix metalloproteinase 24 preproprotein	1	2	extracellular matrix (sensu Metazoa) integral to plasma membrane membrane	peptidoglycan metabolism proteolysis	calcium ion binding enzyme activator activity metalloendopeptidase activity zinc ion binding	
	2	13027809	NP_073209	MMP25	matrix metalloproteinase 25 preproprotein	1	1	cell surface extracellular matrix (sensu Metazoa) membrane	inflammatory response peptidoglycan metabolism proteolysis	calcium ion binding GPI anchor binding metalloendopeptidase activity zinc ion binding	
	1,2	13376091	NP_079032	MMRN2	multimerin 2	1	1	extracellular matrix (sensu Metazoa)	not classified	not classified	
	2	4505227	NP_002423	MNDA	myeloid cell nuclear differentiation antigen	1	3	nucleus	cellular defense response regulation of transcription, DNA-dependent transcription	3-mercaptopyruvate sulfurtransferase activity	
	2	4557759	NP_000241	MPO	myeloperoxidase	7	32	lysosome	anti-apoptosis defense response hydrogen peroxide catabolism response to oxidative stress	calcium ion binding chromatin binding iron ion binding oxidoreductase activity peroxidase activity	Myeloperoxidase Deficiency [MIM:254600]
	2	38570142	NP_071919	MPP5	membrane protein, palmitoylated 5	1	1	cytoplasm membrane tight junction	not classified	protein binding	
	2	61835204	NP_001013454	MPST	3-mercaptopyruvate sulfurtransferase	3	6	mitochondrial matrix	cyanate catabolism response to toxin sulfate transport	3-mercaptopyruvate sulfurtransferase activity thiosulfate sulfurtransferase activity transferase activity	
Yes	2	15809016	NP_291024	MRLC2	myosin regulatory light chain MRCL2	3	5	not classified	not classified	calcium ion binding	
	1,2	4505257	NP_002435	MSN	moesin	27	76	cytoplasm cytoskeleton plasma membrane	cell motility	cytoskeletal protein binding receptor binding structural constituent of cytoskeleton	
	2	6912516	NP_036463	MSRA	methionine sulfoxide reductase A	1	2	not classified	methionine metabolism protein metabolism protein modification response to oxidative stress	oxidoreductase activity protein-methionine-S-oxide reductase activity	
	2	47132622	NP_002442	MTAP	5'-methylthioadenosine phosphorylase	1	1	not classified	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	S-methyl-5-thioadenosine phosphorylase activity transferase activity, transferring pentosyl groups actin binding	

Supplementary Table I

	2	13699868	NP_005947	MTHFD1	methylenetetrahydrofolate dehydrogenase 1	4	5	mitochondrion	amino acid biosynthesis folic acid and derivative biosynthesis histidine biosynthesis methionine biosynthesis one-carbon compound metabolism purine nucleotide biosynthesis	ATP binding formate-tetrahydrofolate ligase activity hydrolase activity ligase activity methylenetetrahydrofolate cyclohydrolase activity methylenetetrahydrofolate dehydrogenase (NADP+) activity nucleotide binding oxidoreductase activity	Spina Bifida [MIM:601634]
	2	67189069	NP_001018017	MUC1	mucin 1 isoform 3 precursor	7	92	cytoskeleton extracellular region integral to membrane membrane membrane	not classified	actin binding hormone activity	
	2	113206028	NP_001037857	MUC1	mucin 1 isoform 7 precursor	2	1	cytoskeleton extracellular region integral to membrane membrane membrane	not classified	hormone activity extracellular matrix structural constituent	
Yes	2	116284392	NP_002448	MUC2	mucin 2 precursor	1	2	extracellular region extracellular space	not classified	extracellular matrix structural constituent	
Yes	2	153945878	NP_002449	MUC5B	mucin 5, subtype B, tracheobronchial	4	8	extracellular matrix (sensu Metazoa) extracellular region	cell adhesion	extracellular matrix structural constituent protein binding	
	2	12313881	NP_060187	MUCDHL	mu-protocadherin isoform 2	1	1	integral to membrane membrane	calcium-dependent cell-cell adhesion cell adhesion homophilic cell adhesion	calcium ion binding protein binding	
	2	24432077	NP_689636	MUM1L1	melanoma associated antigen (mutated) 1-L1	2	6	not classified	not classified	not classified	
	2	110618250	NP_112554	MUPCDH	mucin and cadherin-like isoform 3	1	1	integral to membrane membrane	calcium-dependent cell-cell adhesion cell adhesion homophilic cell adhesion	protein binding actin binding	
	2	19913412	NP_005106	MVP	major vault protein	6	14	cytoplasm nucleus ribonucleoprotein complex	response to drug	not classified	
	2	18390319	NP_056234	MXRA5	adlican	1	1	not classified	protein amino acid phosphorylation	ATP binding transferase activity vascular endothelial growth factor receptor activity	
	1,2	14150145	NP_115724	MXRA8	matrix-remodelling associated 8	7	28	not classified	not classified	not classified	
Yes	2	41406064	NP_005955	MYH10	myosin, heavy polypeptide 10, non-muscle	4	10	myosin	cytokinesis regulation of cell shape	actin binding ATP binding calmodulin binding motor activity nucleotide binding protein binding	
	2	11321579	NP_003793	MYH13	myosin, heavy polypeptide 13, skeletal muscle	1	1	muscle myosin myosin striated muscle thick filament	striated muscle contraction	actin binding ATP binding calmodulin binding microfilament motor activity nucleotide binding	

Supplementary Table 1

	2	33563340	NP_079005	MYH14	myosin, heavy chain 14 isoform 2	8	17	myosin	regulation of cell shape sensory perception of sound	actin binding ATP binding calmodulin binding motor activity nucleotide binding	Autosomal Dominant nonsyndromic sensorineural hearing loss [MIM:600652]
	2	116284396	NP_001070654	MYH14	myosin, heavy chain 14 isoform 1	1	1	myosin	regulation of cell shape sensory perception of sound	ATP binding calmodulin binding motor activity nucleotide binding ATP binding	Autosomal Dominant Nonsyndromic Sensorineural Deafness [MIM: 600652]
	2	11342672	NP_002461	MYH3	myosin, heavy chain 3, skeletal muscle, embryonic	1	1	myosin striated muscle thick filament	actin filament-based movement muscle development striated muscle contraction	actin binding ATP binding ATPase activity, coupled calmodulin binding microfilament motor activity nucleotide binding	Freeman-Sheldon Syndrome [MIM:193700]
	1,2	12667788	NP_002464	MYH9	myosin, heavy polypeptide 9, non-muscle	19	51	myosin	regulation of cell shape sensory perception of sound	actin binding ATP binding calmodulin binding motor activity nucleotide binding	Fechtner syndrome [MIM:153640] Epstein syndrome [MIM:153650]
Yes	2	17986258	NP_066299	MYL6	myosin, light chain 6, alkali, smooth muscle	1	3	unconventional myosin	muscle filament sliding striated muscle development	calcium ion binding motor activity structural constituent of muscle	
Yes	2	88999583	NP_524147	MYL6	myosin, light chain 6, alkali, smooth muscle	3	10	unconventional myosin	muscle filament sliding striated muscle development	calcium ion binding motor activity structural constituent of muscle	
	2	22547229	NP_057323	MYO15A	myosin XV	1	3	cytoskeleton myosin	sensory perception of sound	actin binding ATP binding calmodulin binding motor activity nucleotide binding structural molecule activity	Recessive Congenital Deafness [MIM:600316]
Yes	2	4885503	NP_005370	MYO1A	myosin IA	1	2	brush border muscle myosin myosin	actin filament-based movement sensory perception of sound	actin binding ATP binding ATPase activity, coupled calmodulin binding microfilament motor activity nucleotide binding	
	2	44889481	NP_036355	MYO1B	myosin IB	2	4	myosin	not classified	actin binding ATP binding calmodulin binding motor activity nucleotide binding	
	2	46430642	NP_203693	MYO1C	myosin IC	33	120	myosin unconventional myosin	not classified	actin binding ATP binding calmodulin binding motor activity nucleotide binding	
	2	51100974	NP_056009	MYO1D	myosin ID	16	24	myosin	not classified	actin binding ATP binding calmodulin binding motor activity nucleotide binding	

Supplementary Table I

Yes	2	10835119	NP_000250	MYO5A	myosin VA (heavy polypeptide 12, myosin)	2	3	growth cone myosin neuron projection	actin filament-based movement transport	actin binding ATP binding ATPase activity, coupled calmodulin binding microfilament motor activity nucleotide binding	
	2	122937345	NP_001073936	MYO5B	Myosin VB	1	1	myosin	not classified	actin binding ATP binding calmodulin binding motor activity nucleotide binding	
	2	50582540	NP_004990	MYO6	myosin VI	7	21	unconventional myosin	actin filament-based movement sensory perception of sound striated muscle contraction	actin binding ATP binding ATPase activity, coupled calmodulin binding microfilament motor activity motor activity nucleotide binding structural constituent of muscle	Autosomal Recessive Congenital Sensorineural Deafness [MIM:607821] Autosomal Dominant Nonsyndromic Sensorineural Deafness [MIM:606346]
	2	49574508	NP_060037	NAGK	N-Acetylglucosamine kinase	1	1	not classified	N-acetylglucosamine metabolism N-acetylmannosamine metabolism	kinase activity N-acetylglucosamine kinase activity nucleotide binding protein binding transferase activity N-acetyltransferase activity	
	1,2	66346698	NP_000254	NAGLU	alpha-N-acetylglucosaminidase precursor	21	63	lysosome	glycosaminoglycan metabolism metabolism nervous system development	alpha-N-acetylglucosaminidase activity hydrolase activity, acting on glycosyl bonds	Mucopolysaccharidosis Type IIIB [MIM:252920]
	2	47933379	NP_003818	NAPA	N-ethylmaleimide-sensitive factor attachmer	1	2	endoplasmic reticulum Golgi apparatus	ER to Golgi vesicle-mediated transport intra-Golgi vesicle-mediated transport intracellular protein transport membrane fusion	binding intracellular transporter activity protein binding	
Yes	2	44917606	NP_071363	NAPB	N-ethylmaleimide-sensitive factor attachmer	1	2	endoplasmic reticulum Golgi apparatus	ER to Golgi vesicle-mediated transport intracellular protein transport	binding intracellular transporter activity	
	2	48675830	NP_945341	NAPE-PLD	N-acyl-phosphatidylethanolamine-hydrolyzif	1	1	not classified	not classified	hydrolase activity	
	2	4505331	NP_003817	NAPG	N-ethylmaleimide-sensitive factor attachmer	1	1	endoplasmic reticulum Golgi apparatus membrane fraction mitochondrion	ER to Golgi vesicle-mediated transport intra-Golgi vesicle-mediated transport intracellular protein transport membrane fusion protein complex assembly protein stabilization	intracellular transporter activity protein binding	
	2	40255089	NP_660202	NAPRT1	nicotinate phosphoribosyltransferase domain	2	2	not classified	pyridine nucleotide biosynthesis	nicotinate phosphoribosyltransferase activity transferase activity, transferring glycosyl groups	
	1,2	4758754	NP_004842	NAPSA	napsin A preproprotein	8	106	not classified	proteolysis	pepsin A activity peptidase activity	

Supplementary Table 1

	2	33356128	NP_078837	NARG1L	NMDA receptor regulated 1-like protein isoform 1	1	1	cytoplasm nucleus transcription factor complex	N-terminal protein amino acid acetylation positive regulation of transcription, DNA-dependent	acetyltransferase activity protein binding ribosome binding	
	2	4758762	NP_004530	NARS	asparaginyl-tRNA synthetase	1	1	cytoplasm soluble fraction	asparaginyl-tRNA aminoacylation aspartyl-tRNA aminoacylation DNA replication protein biosynthesis	asparagine-tRNA ligase activity aspartate-tRNA ligase activity ATP binding DNA-directed DNA polymerase activity ligase activity nucleic acid binding nucleotide binding protein binding transferase activity	
	2	13376735	NP_079422	NAT13	N-acetyltransferase 13	1	1	cytoplasm	N-terminal protein amino acid acetylation	protein binding transferase activity protein binding	
Yes	2	98985795	NP_001035720	NCALD	neurocalcin delta	1	3	clathrin coat of trans-Golgi network vesicle cytosol intracellular	vesicle-mediated transport	actin binding calcium ion binding clathrin binding tubulin binding	
	2	7305303	NP_038464	NCKAP1	NCK-associated protein 1 isoform 1	7	12	integral to membrane membrane	apoptosis central nervous system development	protein binding	
	2	45545411	NP_995314	NCKAP1	NCK-associated protein 1 isoform 2	6	7	integral to membrane membrane	apoptosis central nervous system development	calcium ion binding	
	2	24638433	NP_056146	NCSTN	nicastrin precursor	2	2	endoplasmic reticulum Golgi apparatus integral to plasma membrane membrane	amyloid precursor protein catabolism membrane protein ectodomain proteolysis Notch receptor processing positive regulation of enzyme activity protein processing	protein binding	
	2	37655183	NP_006087	NDRG1	N-myc downstream regulated gene 1	2	5	nucleus	cell differentiation response to metal ion	not classified	Charcot-Marie-Tooth Disease Type 4D [MIM:601455]
	2	32130514	NP_071922	NDRG3	N-myc downstream-regulated gene 3 isoform 1	1	1	cytoplasm	cell differentiation negative regulation of cell growth spermatogenesis	not classified	
	2	4758794	NP_004534	NEB	nebulin	2	4	actin cytoskeleton striated muscle thin filament Z disc	regulation of actin filament length somatic muscle development	actin binding protein binding structural constituent of muscle	Nemaline Myopathy [MIM:256030]
	2	5453758	NP_006384	NEBL	nebulin sarcomeric isoform	1	1	actin cytoskeleton	regulation of actin filament length	actin binding structural constituent of muscle zinc ion binding	
	2	115298674	NP_002499	NID1	nidogen 1 precursor	1	1	basement membrane membrane	cell adhesion cell-matrix adhesion	extracellular matrix structural constituent protein binding hedgehog receptor activity	
	2	47578107	NP_056199	NIPBL	delangin isoform B	1	1	nucleus	cell cycle	binding	
	2	9910460	NP_064587	NIT2	nitrilase family, member 2	1	1	not classified	nitrogen compound metabolism	hydrolase activity, acting on carbon- nitrogen (but not peptide) bonds	
	2	5453788	NP_006159	NKX6-1	NK6 transcription factor related, locus 1	1	14	nucleus	development organ morphogenesis regulation of transcription, DNA- dependent	sequence-specific DNA binding transcription factor activity	

Supplementary Table I

Yes	2	38045913	NP_937818	NME1	non-metastatic cells 1, protein (NM23A) exp	2	3	cytoplasm nucleus	cell cycle CTP biosynthesis GTP biosynthesis negative regulation of cell proliferation negative regulation of progression through cell cycle nucleoside triphosphate biosynthesis nucleotide metabolism regulation of apoptosis UTP biosynthesis	ATP binding deoxyribonuclease activity DNA binding kinase activity magnesium ion binding nucleoside diphosphate kinase activity nucleotide binding transferase activity	
Yes	2	66392227	NP_001018145	NME2	non-metastatic cells 2, protein (NM23B) exp	2	6	lamellipodium nucleus ruffle	cell adhesion cell cycle CTP biosynthesis GTP biosynthesis negative regulation of cell proliferation negative regulation of progression through cell cycle nucleoside triphosphate biosynthesis nucleotide metabolism regulation of transcription, DNA-dependent transcription UTP biosynthesis	ATP binding kinase activity magnesium ion binding nucleoside diphosphate kinase activity nucleotide binding protein binding transcription factor activity transferase activity	
	2	11136626	NP_056533	NOX3	NADPH oxidase 3	1	1	integral to membrane membrane	electron transport	FAD binding iron ion binding oxidoreductase activity	
	2	4557801	NP_000261	NP	purine nucleoside phosphorylase	7	14	not classified	DNA modification nucleobase, nucleoside, nucleotide and nucleic acid metabolism	purine-nucleoside phosphorylase activity	
	2	4557803	NP_000262	NPC1	Niemann-Pick disease, type C1 precursor	1	1	endoplasmic reticulum integral to membrane lysosome membrane nuclear envelope perinuclear region	bile acid metabolism cholesterol homeostasis cholesterol transport lysosomal transport	intracellular transporter activity sterol transporter activity sequence-specific DNA binding	
	2	15451907	NP_006301	NPEPPS	aminopeptidase puromycin sensitive	4	5	cytosol nucleus	proteolysis	aminopeptidase activity membrane alanyl aminopeptidase activity metal ion binding metallopeptidase activity zinc ion binding	
	2	4758822	NP_004637	NPHS1	nephrin	1	1	cell adhesion excretion	cell adhesion excretion	protein binding	
	1,2	7657615	NP_055440	NPHS2	podocin	6	9	integral to plasma membrane membrane	excretion	protein binding	Autosomal recessive steroid-resistant nephrotic syndrome [MIM:600995]

Supplementary Table 1

	2	4505417	NP_000895	NQO2	NAD(P)H dehydrogenase, quinone 2	1	2	not classified	electron transport	electron carrier activity metal ion binding NAD(P)H dehydrogenase (quinone) activity NADPH dehydrogenase (quinone) activity zinc ion binding	
Yes	2	4505451	NP_002515	NRAS	neuroblastoma RAS viral (v-ras) oncogene	5	10	intracellular membrane	regulation of progression through cell cycle small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	2	93141039	NP_001035199	NRF1	nuclear respiratory factor 1	1	4	nucleus	generation of precursor metabolites and energy regulation of transcription from RNA polymerase II promoter regulation of transcription, DNA-dependent transcription	transcription factor activity ADP reductase activity	
	2	7657033	NP_055410	NT5C	5',3'-nucleotidase, cytosolic	2	2	cytoplasm nucleus	dephosphorylation nucleotide metabolism pyrimidine deoxyribonucleotide catabolism	hydrolase activity magnesium ion binding nucleotidase activity pyrimidine nucleotide binding	
	2	4505467	NP_002517	NT5E	5' nucleotidase, ecto	1	2	membrane membrane fraction	DNA metabolism nucleotide catabolism	5'-nucleotidase activity GPI anchor binding hydrolase activity, acting on ester bonds metal ion binding nucleotide binding zinc ion binding	
	2	37594464	NP_054861	NUDT5	nudix-type motif 5	2	1	intracellular	D-ribose catabolism nucleotide metabolism ribonucleoside diphosphate catabolism	ADP-ribose diphosphatase activity ADP-sugar diphosphatase activity hydrolase activity magnesium ion binding protein binding olfactory receptor activity	
Yes	2	5031985	NP_005787	NUTF2	nuclear transport factor 2	1	1	intracellular	protein transport	protein binding	
	1,2	32313593	NP_006409	OLFM4	olfactomedin 4 precursor	19	63	membrane	not classified	latrotoxin receptor activity	
	2	50080203	NP_001001959	OR11L1	olfactory receptor, family 11, subfamily L, member 1	1	1	integral to membrane membrane	G-protein coupled receptor protein signaling pathway response to stimulus sensory perception of smell signal transduction	receptor activity cysteine-type peptidase activity	
Yes	2	52317120	NP_001004738	OR5L1	olfactory receptor, family 5, subfamily L, member 1	1	1	integral to membrane membrane	G-protein coupled receptor protein signaling pathway response to stimulus sensory perception of smell signal transduction	olfactory receptor activity receptor activity	
	2	19718746	NP_060500	OSBPL1A	oxysterol-binding protein-like 1A isoform A	1	1	intracellular	cholesterol metabolism lipid transport steroid metabolism vesicle-mediated transport	phospholipid binding	
	2	109148508	NP_060140	OTUB1	otubain 1	2	5	not classified	immune response ubiquitin cycle	protein binding	

Supplementary Table I

	2	4826878	NP_005100	OXR1	oxidative-stress responsive 1	2	2	not classified	protein amino acid phosphorylation protein kinase cascade response to oxidative stress	ATP binding magnesium ion binding nucleotide binding protein binding protein serine/threonine kinase activity transferase activity	
	1,2	28416927	NP_002551	P2RX4	purinergic receptor P2X4	2	2	integral to membrane integral to plasma membrane membrane	ion transport signal transduction	ATP binding ATP-gated cation channel activity ion channel activity receptor activity	
	2	6005826	NP_009160	PACSN2	protein kinase C and casein kinase substrat	4	9	cytoplasm	actin cytoskeleton organization and biogenesis endocytosis intracellular protein transport	transporter activity	
	1,2	19224660	NP_057307	PACSN3	protein kinase C and casein kinase substrat	4	8	cytoplasm	endocytosis negative regulation of endocytosis	cytoskeletal protein binding enzyme binding	
	2	15042937	NP_031391	PADI2	peptidyl arginine deiminase, type II	1	1	not classified	protein modification	calcium ion binding hydrolase activity protein-arginine deiminase activity	
	2	4557741	NP_000421	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform	1	1	astral microtubule cell cortex cytoskeleton kinetochore microtubule microtubule associated complex nuclear envelope	cell cycle cell differentiation cell division cell motility establishment of mitotic spindle orientation lipid catabolism microtubule-based process mitosis nervous system development signal transduction transport	dynein binding	Miller-Dieker Lissencephaly Syndrome [MIM:607432]
	2	4505585	NP_002563	PAFAH1B2	platelet-activating factor acetylhydrolase, isoform	1	2	not classified	lipid catabolism	1-alkyl-2-acetyl-glycerophosphocholine esterase activity hydrolase activity, acting on ester bonds	
	2	62955042	NP_115910	PARD6B	PAR-6 beta	1	1	plasma membrane tight junction	axonogenesis cell cycle cell division establishment and/or maintenance of cell polarity intercellular junction assembly regulation of cell migration	not classified	
	2	31543380	NP_009193	PARK7	DJ-1 protein	1	1	cytoplasm nucleus	protein folding Ras protein signal transduction	catalytic activity	Parkinson Disease 7, Autosomal Recessive [MIM: 606324]
	2	5915660	NP_006428	PARP4	poly (ADP-ribose) polymerase family, member	1	1	intracellular nucleus ribonucleoprotein complex	cell death DNA repair inflammatory response protein amino acid ADP-ribosylation response to drug transport	DNA binding enzyme binding NAD+ ADP-ribosyltransferase activity transferase activity, transferring glycosyl groups	
	2	62177133	NP_071412	PBLD	MAWD binding protein isoform a	4	6	not classified	biosynthesis	catalytic activity isomerase activity	
	2	74316009	NP_001028253	PBLD	MAWD binding protein isoform b	1	1	not classified	biosynthesis	isomerase activity calcium ion binding	

Supplementary Table I

	2	4557831	NP_000272	PCBD1	pterin-4 alpha-carbinolamine dehydratase precursor	1	1	not classified	tetrahydrobiopterin biosynthesis	4-alpha-hydroxytetrahydrobiopterin dehydratase activity identical protein binding lyase activity	Hyperphenylalaninemic [MIM:264070]
	2	5453854	NP_006187	PCBP1	poly(rC) binding protein 1	1	1	cytoplasm nucleus ribonucleoprotein complex	mRNA metabolism	DNA binding protein binding RNA binding single-stranded DNA binding	
	2	32483401	NP_002582	PCK1	cytosolic phosphoenolpyruvate carboxykinase	4	6	cytosol	gluconeogenesis	GTP binding lyase activity nucleotide binding phosphoenolpyruvate carboxykinase (GTP) activity	
	2	16933559	NP_060145	PCLKC	protocadherin LKC precursor	1	1	integral to membrane membrane	cell adhesion homophilic cell adhesion negative regulation of cell growth	protein binding not classified	
	2	33620751	NP_057381	PCYOX1	prenylcysteine oxidase 1	4	6	lysosome	prenylated protein catabolism prenylcysteine metabolism	lyase activity oxidoreductase activity prenylcysteine oxidase activity	
	2	20127517	NP_009148	PDCD10	programmed cell death 10	2	3	not classified	apoptosis	not classified	Cerebral Cavemous Malformations [MIM:603285]
	2	21735592	NP_002589	PDCD2	programmed cell death 2 isoform 1	1	1	cytoplasm nucleus	apoptosis	DNA binding metal ion binding protein binding zinc ion binding	
	1,2	7019485	NP_037364	PDCD6	programmed cell death 6	11	41	not classified	apoptosis induction of apoptosis by extracellular signals	calcium ion binding	
	1,2	22027538	NP_037506	PDCD6IP	programmed cell death 6 interacting protein	42	183	cytosol	apoptosis protein transport	signal transducer activity	
	2	32699045	NP_872418	PDDC1	Parkinson disease 7 domain containing 1	1	1	not classified	not classified	not classified	
	2	47132537	NP_775656	PDE8A	phosphodiesterase 8A isoform 2	1	1	not classified	cyclic nucleotide metabolism regulation of transcription, DNA-dependent two-component signal transduction system (phosphorelay)	hydrolase activity magnesium ion binding manganese ion binding two-component response regulator activity 6-phosphofructokinase activity	
	2	4505701	NP_003672	PDXK	pyridoxal kinase	1	3	not classified	not classified	ATP binding metal ion binding nucleotide binding pyridoxal kinase activity transferase activity zinc ion binding	
	2	21361142	NP_002605	PDZK1	PDZ domain containing 1	16	35	not classified	drug transport	PDZ domain binding protein C-terminus binding	
	1,2	5031657	NP_005755	PDZK1IP1	PDZK1 interacting protein 1	4	11	extracellular region integral to membrane membrane	not classified	not classified	
	1,2	4505621	NP_002558	PEBP1	prostatic binding protein	8	21	not classified	not classified	ATP binding lipid binding nucleotide binding phosphatidylethanolamine binding serine-type endopeptidase inhibitor activity	
	1,2	6912582	NP_036524	PEF1	penta-EF hand domain containing 1	5	17	membrane	not classified	calcium ion binding protein binding	

Supplementary Table I

	2	4557835	NP_000276	PEPD	Xaa-Pro dipeptidase	1	8	not classified	amino acid metabolism collagen catabolism proteolysis	manganese ion binding metal ion binding metallocarboxypeptidase D activity X-Pro dipeptidase activity	
	2	7656971	NP_055702	PFAAP5	phosphonoformate immuno-associated prot	1	1	not classified	not classified	not classified	
	2	48762920	NP_002617	PFKL	liver phosphofructokinase isoform b	11	14	6-phosphofructokinase complex cytoplasm	fructose 6-phosphate metabolism glycolysis regulation of glycolysis	6-phosphofructokinase activity ATP binding kinase activity magnesium ion binding nucleotide binding transferase activity	
	2	50346005	NP_00100202	PFKL	liver phosphofructokinase isoform a	6	10	6-phosphofructokinase complex cytoplasm	fructose 6-phosphate metabolism glycolysis regulation of glycolysis	ATP binding kinase activity magnesium ion binding nucleotide binding transferase activity hydrolase activity	
Yes	2	4505749	NP_000280	PFKM	phosphofructokinase, muscle	1	2	6-phosphofructokinase complex cytoplasm	glycogen metabolism glycolysis regulation of glycolysis	6-phosphofructokinase activity ATP binding kinase activity magnesium ion binding nucleotide binding transferase activity	
	1,2	4826898	NP_005013	PFN1	profilin 1	5	23	actin cytoskeleton	actin cytoskeleton organization and biogenesis cytoskeleton organization and biogenesis	actin monomer binding	
	1,2	16753215	NP_444252	PFN2	profilin 2 isoform a	1	2	actin cytoskeleton	actin cytoskeleton organization and biogenesis cytoskeleton organization and biogenesis regulation of actin polymerization and/or depolymerization	actin binding phosphatidylinositol-4,5-bisphosphate binding protein binding	
Yes	1,2	4505753	NP_002620	PGAM1	phosphoglycerate mutase 1 (brain)	2	3	cytosol	glycolysis metabolism	bisphosphoglycerate mutase activity bisphosphoglycerate phosphatase activity catalytic activity hydrolase activity intramolecular transferase activity, phosphotransferases isomerase activity phosphoglycerate mutase activity	
	2	40068518	NP_002622	PGD	phosphogluconate dehydrogenase	6	15	not classified	electron transport pentose-phosphate shunt, oxidative branch	NADP binding oxidoreductase activity phosphogluconate dehydrogenase (decarboxylating) activity protein binding	
Yes	1,2	4505763	NP_000282	PGK1	phosphoglycerate kinase 1	23	58	not classified	glycolysis phosphorylation	ATP binding nucleotide binding phosphoglycerate kinase activity transferase activity	
Yes	2	31543397	NP_620061	PGK2	phosphoglycerate kinase 2	3	10	cytosol	glycolysis phosphorylation	ATP binding nucleotide binding phosphoglycerate kinase activity transferase activity	
	1,2	6912586	NP_036220	PGLS	6-phosphogluconolactonase	4	4	not classified	carbohydrate metabolism pentose-phosphate shunt	6-phosphogluconolactonase activity hydrolase activity	

Supplementary Table I

	2	4827036	NP_005082	PGLYRP1	peptidoglycan recognition protein 1	1	1	extracellular region	defense response to Gram-positive bacterium detection of bacterium innate immune response peptidoglycan catabolism	bacterial binding N-acetylmuramoyl-L-alanine amidase activity peptidoglycan receptor activity protein binding	
	2	21361621	NP_002624	PGM1	phosphoglucomutase 1	2	3	cytoplasm	carbohydrate metabolism gluconeogenesis glucose metabolism glycolysis	intramolecular transferase activity, phosphotransferases isomerase activity magnesium ion binding phosphoglucomutase activity	
	2	40556370	NP_056103	PHF15	PHD finger protein 15	1	2	not classified	regulation of transcription, DNA-	metal ion binding protein binding zinc ion binding	
	2	23308577	NP_006614	PHGDH	phosphoglycerate dehydrogenase	2	2	extracellular region	amino acid biosynthesis brain development L-serine biosynthesis metabolism transport	electron carrier activity NAD binding oxidoreductase activity oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor phosphoglycerate dehydrogenase activity transporter activity	Phosphoglycerate Dehydrogenase Deficiency [MIM:601815]
	2	24475861	NP_054891	PHPT1	phosphohistidine phosphatase 1	2	1	cytoplasm cytosol	dephosphorylation	phosphohistidine phosphatase activity phosphoprotein phosphatase activity kinase activity	
	1,2	31377806	NP_002635	PIGR	polymeric immunoglobulin receptor	17	32	integral to plasma membrane membrane	protein secretion	protein transporter activity	
	2	4505799	NP_002636	PIK3C2A	phosphoinositide-3-kinase, class 2, alpha p	1	1	nucleus phosphoinositide 3-kinase complex	intracellular signaling cascade phosphatidylinositol biosynthesis	inositol or phosphatidylinositol kinase activity phosphatidylinositol 3-kinase activity phosphatidylinositol-4-phosphate 3-kinase activity phosphoinositide binding protein binding transferase activity	
	2	4505821	NP_002643	PIP	prolactin-induced protein	2	2	extracellular region	not classified	actin binding	
	2	21322230	NP_079055	PIP4K2C	phosphatidylinositol-5-phosphate 4-kinase,	1	3	not classified	not classified	kinase activity magnesium ion binding	
	1,2	58331146	NP_001009944	PKD1	polycystin 1 isoform 1 precursor	1	1	extracellular matrix (sensu Metazoa) integral to membrane integral to plasma membrane membrane	calcium-independent cell-matrix adhesion homophilic cell adhesion morphogenesis neuropeptide signaling pathway	sugar binding transferase activity	Polycystic Kidney Disease, adult, type I [MIM:601313]
	2	31559825	NP_853514	PKD1L3	polycystin 1-like 3	1	1	integral to membrane membrane	cation transport neuropeptide signaling pathway	calcium ion binding cation channel activity	
	2	4505835	NP_000288	PKD2	polycystin 2	1	2	actin cytoskeleton basal body integral to membrane membrane plasma membrane	cation transport cell-matrix adhesion organ morphogenesis	calcium ion binding cytoskeletal protein binding ion channel activity protein C-terminus binding voltage-gated chloride channel activity voltage-gated sodium channel activity	Polycystic kidney disease, adult, type II [MIM:173910]

Supplementary Table I

	2	25777667	NP_733842	PKHD1	polyductin isoform 2	6	9	anchored to external side of plasma membrane apical plasma membrane basal body integral to membrane membrane	cell-cell adhesion homeostasis negative regulation of cell motility	protein binding receptor activity	Autosomal Recessive Polycystic Kidney Disease [MIM:263200]
	2	10835121	NP_000289	PKLR	pyruvate kinase, liver and RBC isoform 1	1	1	not classified	glycolysis	magnesium ion binding pyruvate kinase activity	Pyruvate Kinase Deficiency [MIM:266200]
	2	32967597	NP_870986	PKLR	pyruvate kinase, liver and RBC isoform 2	1	1	not classified	glycolysis	pyruvate kinase activity magnesium ion binding	
Yes	1,2	33286418	NP_002645	PKM2	pyruvate kinase 3 isoform 1	1	5	cytoplasm cytosol	glycolysis	magnesium ion binding protein binding pyruvate kinase activity transferase activity	
	2	33286420	NP_872270	PKM2	pyruvate kinase 3 isoform 2	17	40	cytoplasm cytosol	glycolysis	magnesium ion binding protein binding pyruvate kinase activity transferase activity	
	2	5453910	NP_006216	PLCD1	phospholipase C, delta 1	8	18	not classified	intracellular signaling cascade lipid catabolism lipid metabolism phospholipid metabolism	calcium ion binding hydrolase activity phosphoinositide phospholipase C activity signal transducer activity	
	2	72534684	NP_001026866	PLD3	phospholipase D3 isoform 1	1	1	membrane	metabolism	catalytic activity	
	2	50234893	NP_001001974	PLEKHA1	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1	1	1	membrane nucleus	not classified	phospholipid binding	
	2	48675827	NP_778228	PLEKHA7	pleckstrin homology domain containing, fam	1	1	not classified	not classified	not classified	
	2	32307144	NP_000293	PLOD1	lysyl hydroxylase precursor	1	1	endoplasmic reticulum membrane	electron transport epidermis development hydroxylysine biosynthesis protein metabolism protein modification	iron ion binding L-ascorbic acid binding metal ion binding oxidoreductase activity procollagen-lysine 5-dioxygenase activity protein homodimerization activity	Ehlers-Danlos Syndrome, Type VIA [MIM:225400]
	2	4505891	NP_001075	PLOD3	procollagen-lysine, 2-oxoglutarate 5-dioxyge	1	2	endoplasmic reticulum membrane	protein metabolism protein modification	iron ion binding L-ascorbic acid binding metal ion binding oxidoreductase activity oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen procollagen-lysine 5-dioxygenase activity prot	
	2	4505897	NP_002661	PLS1	plastin 1	1	2	actin cytoskeleton brush border	not classified	actin filament binding calcium ion binding structural constituent of cytoskeleton	
Yes	2	7549809	NP_005023	PLS3	plastin 3	9	21	actin cytoskeleton	not classified	actin binding calcium ion binding protein binding	
	1,2	10863877	NP_066928	PLSCR1	phospholipid scramblase 1	2	3	integral to membrane plasma membrane	phospholipid scrambling platelet activation response to virus	calcium ion binding phospholipid scramblase activity protein binding	

Supplementary Table I

	2	5729980	NP_006547	PMVK	phosphomevalonate kinase	1	1	cytoplasm peroxisome	cholesterol biosynthesis lipid metabolism protein amino acid phosphorylation	phosphomevalonate kinase activity transferase activity	
Yes	1,2	66277202	NP_00101812	PODXL	podocalyxin-like precursor isoform 1	9	52	integral to plasma membrane membrane	not classified	not classified	
	2	33598950	NP_005388	PODXL	podocalyxin-like precursor isoform 2	8	57	integral to plasma membrane membrane	not classified	not classified	
	2	29171738	NP_795714	PPAP2A	phosphatidic acid phosphatase type 2A isoform 2	2	3	integral to plasma membrane membrane fraction plasma membrane	androgen receptor signaling pathway germ cell migration negative regulation of cell proliferation phospholipid dephosphorylation protein kinase C activation regulation of lipid metabolism	hydrolase activity phosphatidate phosphatase activity	
	2	29171755	NP_003616	PPFIA2	PTPRF interacting protein alpha 2	1	1	cytoplasm	cell-matrix adhesion	protein binding	
	2	10863927	NP_066953	PPIA	peptidylprolyl isomerase A	5	17	not classified	protein folding regulation of viral genome replication	isomerase activity peptide binding peptidyl-prolyl cis-trans isomerase activity unfolded protein binding virion binding	
Yes	2	30102944	NP_839944	PPIAL4	peptidylprolyl isomerase A (cyclophilin A)-like 2	2	3	not classified	protein folding	isomerase activity peptidyl-prolyl cis-trans isomerase activity	
	2	4758950	NP_000933	PPIB	peptidylprolyl isomerase B precursor	1	1	endoplasmic reticulum endoplasmic reticulum lumen	protein folding	isomerase activity peptide binding peptidyl-prolyl cis-trans isomerase activity unfolded protein binding	
	2	45439327	NP_002696	PPL	periplakin	18	28	cytoskeleton	keratinization	protein binding structural constituent of cytoskeleton	
Yes	2	56790945	NP_001008709	PPP1CA	protein phosphatase 1, catalytic subunit, alpha 2	2	9	not classified	carbohydrate metabolism cell cycle cell division glycogen metabolism	hydrolase activity iron ion binding manganese ion binding metal ion binding phosphoprotein phosphatase activity	
Yes	2	4506005	NP_002700	PPP1CB	protein phosphatase 1, catalytic subunit, beta 1	1	5	not classified	carbohydrate metabolism cell cycle cell division glycogen metabolism	hydrolase activity iron ion binding manganese ion binding metal ion binding phosphoprotein phosphatase activity	
	2	4506013	NP_002703	PPP1R7	protein phosphatase 1, regulatory subunit 7	1	1	cytoplasm nucleus	not classified	transferase activity oxidoreductase activity	

Supplementary Table I

Yes	2	21361399	NP_055040	PPP2R1A	alpha isoform of regulatory subunit A, protein	7	10	cytosol membrane microtubule cytoskeleton mitochondrion nucleus protein phosphatase type 2A complex soluble fraction	ceramide metabolism inactivation of MAPK activity induction of apoptosis negative regulation of cell growth negative regulation of tyrosine phosphorylation of Stat3 protein protein amino acid dephosphorylation protein complex assembly regulation of cell a	antigen binding binding protein binding protein heterodimerization activity protein phosphatase type 2A regulator activity	
Yes	2	4506019	NP_002708	PPP2R2A	alpha isoform of regulatory subunit B55, protein	1	3	protein phosphatase type 2	protein amino acid dephosphorylation signal transduction	protein binding protein phosphatase type 2A activity protein phosphatase type 2A regulator activity	
Yes	2	32307117	NP_858062	PPP2R2B	beta isoform of regulatory subunit B55, protein	1	2	protein phosphatase type 2	signal transduction	protein phosphatase type 2A regulator activity	
	2	40805841	NP_955450	PRCP	prolylcarboxypeptidase isoform 2 precursor	1	4	lysosome	proteolysis	serine-type peptidase activity serine-type Pro-X carboxypeptidase activity	
Yes	2	32455264	NP_859047	PRDX1	peroxiredoxin 1	6	23	not classified	cell proliferation skeletal development	oxidoreductase activity peroxiredoxin activity	
Yes	1,2	32189392	NP_005800	PRDX2	peroxiredoxin 2 isoform a	2	9	cytoplasm	anti-apoptosis regulation of apoptosis response to oxidative stress	oxidoreductase activity peroxiredoxin activity thioredoxin peroxidase activity	
Yes	2	33188452	NP_859427	PRDX2	peroxiredoxin 2 isoform b	4	14	cytoplasm	anti-apoptosis regulation of apoptosis response to oxidative stress	oxidoreductase activity peroxidase activity peroxiredoxin activity thioredoxin peroxidase activity	
	2	6912238	NP_036226	PRDX5	peroxiredoxin 5 precursor, isoform a	2	4	mitochondrion peroxisome	inflammatory response response to oxidative stress	oxidoreductase activity peroxiredoxin activity	
	2	32455262	NP_857635	PRDX5	peroxiredoxin 5 precursor, isoform c	2	7	mitochondrion peroxisome	inflammatory response response to oxidative stress	oxidoreductase activity peroxiredoxin activity	
	2	32455260	NP_857634	PRDX5	peroxiredoxin 5 precursor, isoform b	1	3	mitochondrion peroxisome	regulation of apoptosis response to oxidative stress	peroxiredoxin activity ATP binding	
	1,2	4758638	NP_004896	PRDX6	peroxiredoxin 6	7	25	cytosol lysosome	lipid catabolism phospholipid catabolism response to oxidative stress	hydrolase activity oxidoreductase activity peroxiredoxin activity phospholipase A2 activity	
Yes	2	154800441	NP_006241	PRH1	proline-rich protein HaellI subfamily 1	1	1	integral to membrane	G-protein coupled receptor protein	rhodopsin-like receptor activity	
	2	4758958	NP_004148	PRKAR2A	cAMP-dependent protein kinase, regulatory	1	1	cAMP-dependent protein kinase complex cytoplasm membrane fraction plasma membrane	intracellular signaling cascade protein amino acid phosphorylation signal transduction	cAMP binding cAMP-dependent protein kinase regulator activity kinase activity nucleotide binding	
	2	4506067	NP_002728	PRKCA	protein kinase C, alpha	1	1	membrane fraction	cell surface receptor linked signal transduction induction of apoptosis by extracellular signals intracellular signaling cascade protein amino acid phosphorylation regulation of progression through cell cycle	ATP binding calcium ion binding diacylglycerol binding nucleotide binding protein binding protein kinase C activity transferase activity zinc ion binding	

Supplementary Table I

Yes	2	47157322	NP_997700	PRKCB1	protein kinase C, beta isoform 1	1	2	cytoplasm plasma membrane	intracellular signaling cascade protein amino acid phosphorylation	ATP binding calcium ion binding diacylglycerol binding nucleotide binding protein kinase C activity transferase activity zinc ion binding	
	2	31377782	NP_006245	PRKCD	protein kinase C, delta	1	1	cytoplasm membrane nucleus	intracellular signaling cascade protein amino acid phosphorylation protein stabilization	ATP binding diacylglycerol binding diacylglycerol-activated phospholipid- dependent protein kinase C activity enzyme activator activity enzyme binding metal ion binding nucleotide binding protein C-terminus binding protein serine/threonine kinase activity	
	2	28557781	NP_006246	PRKCH	protein kinase C, eta	1	2	not classified	intracellular signaling cascade protein amino acid phosphorylation	ATP binding diacylglycerol binding metal ion binding nucleotide binding protein kinase C activity transferase activity zinc ion binding	Cerebral Infarction [MIM:601367]
	2	48255885	NP_002731	PRKCI	protein kinase C, iota	1	1	cytosol nucleus polarisome	cytoskeleton organization and biogenesis establishment and/or maintenance of epithelial cell polarity intercellular junction assembly and maintenance intracellular signaling cascade membrane organization and biogenesis protein amino acid phosphorylation p	ATP binding atypical protein kinase C activity diacylglycerol binding metal ion binding nucleotide binding phospholipid binding protein binding protein serine/threonine kinase activity transferase activity zinc ion binding	
	2	52486327	NP_002735	PRK CZ	protein kinase C, zeta isoform 1	1	2	cytoplasm membrane fraction plasma membrane	anti-apoptosis intracellular signaling cascade protein amino acid phosphorylation	ATP binding atypical protein kinase C activity diacylglycerol binding metal ion binding nucleotide binding protein binding protein serine/threonine kinase activity transferase activity zinc ion binding	
	2	75709228	NP_001028754	PRK CZ	protein kinase C, zeta isoform 2	1	3	cytoplasm membrane fraction plasma membrane	anti-apoptosis intracellular signaling cascade protein amino acid phosphorylation	atypical protein kinase C activity diacylglycerol binding metal ion binding nucleotide binding protein binding protein serine/threonine kinase activity transferase activity zinc ion binding copper ion binding	
	2	13375901	NP_078929	PRKRIP1	PRKR interacting protein 1 (IL11 inducible)	1	1	not classified	not classified	not classified	

Supplementary Table 1

	2	34335270	NP_898902	PRNP	prion protein preproprotein	1	1	cytoplasm endoplasmic reticulum extrinsic to membrane Golgi apparatus lipid raft plasma membrane	copper ion homeostasis metabolism response to oxidative stress	copper ion binding GPI anchor binding microtubule binding	Creutzfeldt-Jakob Disease [MIM: 123400]
	1,2	5174387	NP_006008	PROM1	prominin 1	23	174	integral to plasma membrane membrane	response to stimulus visual perception	not classified	Autosomal Recessive Retinal Degeneration [MIM:604365]
	2	21389623	NP_653308	PROM2	prominin 2	18	92	integral to membrane	not classified	not classified	
	2	4506121	NP_003882	PROZ	protein Z, vitamin K-dependent plasma glyco	4	6	extracellular region extracellular space	blood coagulation proteolysis	calcium ion binding protein binding serine-type endopeptidase activity	
	2	4506153	NP_002764	PRSS8	prostasin preproprotein	4	6	extracellular region extracellular space integral to membrane plasma membrane	proteolysis	serine-type endopeptidase activity	
	2	71361688	NP_002768	PRTN3	proteinase 3 (serine proteinase, neutrophil, Wegener granulomatosis autoantigen)	1	1	not classified	collagen catabolism positive regulation of cell proliferation proteolysis	protein binding serine-type endopeptidase activity	
	1,2	11386147	NP_002769	PSAP	prosaposin isoform a preproprotein	3	6	extracellular space integral to membrane lysosome	glycosphingolipid metabolism lipid metabolism lipid transport lysosome organization and biogenesis sphingolipid metabolism	alpha-galactosidase activity beta-galactosidase activity cerebroside-sulfatase activity enzyme activator activity galactosylceramidase activity glucosylceramidase activity lipid binding sphingomyelin phosphodiesterase activity	Metachromatic Leukodystrophy Due to SAP1 Deficiency [MIM:249900] Gaucher Disease, Atypical Due to SAP2 Deficiency [MIM:610539]
	2	110224479	NP_001035931	PSAP	prosaposin isoform c preproprotein	1	4	extracellular space integral to membrane lysosome	glycosphingolipid metabolism lipid metabolism lipid transport lysosome organization and biogenesis sphingolipid metabolism	beta-galactosidase activity cerebroside-sulfatase activity enzyme activator activity galactosylceramidase activity glucosylceramidase activity lipid binding sphingomyelin phosphodiesterase activity phosphoserine transaminase activity	Metachromatic Leukodystrophy [MIM: 249900]
	2	17402893	NP_478059	PSAT1	phosphoserine aminotransferase isoform 1	2	4	not classified	amino acid biosynthesis L-serine biosynthesis metabolism pyridoxine biosynthesis	transferase activity RNA binding	Phosphoserine Aminotransferase Deficiency [MIM: 610992]
	1,2	5031995	NP_005663	PSCA	prostate stem cell antigen preproprotein	1	7	plasma membrane	not classified	GPI anchor binding	
	2	23110935	NP_683877	PSMA1	proteasome alpha 1 subunit isoform 1	2	1	cytosol nucleus polysome proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	threonine endopeptidase activity	
	2	4506181	NP_002778	PSMA2	proteasome alpha 2 subunit	2	3	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein cata	threonine endopeptidase activity	

Supplementary Table I

	2	23110939	NP_687033	PSMA3	proteasome alpha 3 subunit isoform 2	1	1	cytoplasm cytosol nucleus proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	protein binding threonine endopeptidase activity	
	2	4506185	NP_002780	PSMA4	proteasome alpha 4 subunit	1	2	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	threonine endopeptidase activity	
	2	23110942	NP_002781	PSMA5	proteasome alpha 5 subunit	2	4	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	hydrolase activity, acting on ester bonds	
Yes	2	23110944	NP_002782	PSMA6	proteasome alpha 6 subunit	1	1	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	protein binding RNA binding threonine endopeptidase activity	
Yes	2	68303565	NP_001020268	PSMA8	proteasome (prosome, macropain) subunit	1	1	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	threonine endopeptidase activity	
	2	4506193	NP_002784	PSMB1	proteasome beta 1 subunit	1	1	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	protein binding threonine endopeptidase activity	
	2	4506195	NP_002785	PSMB2	proteasome beta 2 subunit	1	4	cytosol nucleus proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	peptidase activity threonine endopeptidase activity	
	1,2	22538465	NP_002786	PSMB3	proteasome beta 3 subunit	4	6	cytosol nucleus proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	threonine endopeptidase activity	
	2	4506201	NP_002788	PSMB5	proteasome beta 5 subunit	3	6	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	protein binding threonine endopeptidase activity	
	2	23110925	NP_002789	PSMB6	proteasome beta 6 subunit	1	1	cytosol proteasome core complex (sensu Eukaryota)	ubiquitin-dependent protein catabolism	threonine endopeptidase activity	
	2	24497435	NP_002796	PSMC5	proteasome 26S ATPase subunit 5	1	1	cytoplasm cytosol nucleus proteasome complex (sensu Eukaryota)	negative regulation of programmed cell death positive regulation of gene-specific transcription proteasomal ubiquitin-dependent protein catabolism protein catabolism	ATP binding ATPase activity hydrolase activity nucleotide binding thyrotropin-releasing hormone receptor binding transcription factor binding	
	2	24430160	NP_002797	PSMC6	proteasome 26S ATPase subunit 6	1	1	cytoplasm cytosol nucleus proteasome complex (sensu Eukaryota)	protein catabolism ubiquitin-dependent protein catabolism	ATP binding ATPase activity hydrolase activity nucleic acid binding nucleotide binding protein binding, bridging	

Supplementary Table I

	2	4506221	NP_002807	PSMD12	proteasome 26S non-ATPase subunit 12	1	1	cytosol proteasome regulatory particle (sensu Eukaryota)	not classified	not classified	
	2	25777612	NP_002800	PSMD3	proteasome 26S non-ATPase subunit 3	1	1	cytosol proteasome complex (sensu Eukaryota) protein complex	not classified	binding	
	2	30410792	NP_002809	PSME2	proteasome activator subunit 2	1	1	cytosol proteasome activator complex proteasome complex (sensu Eukaryota)	immune response	proteasome activator activity	
	2	20070186	NP_109589	PTER	phosphotriesterase related	2	4	not classified	catabolism	hydrolase activity, acting on ester bonds metal ion binding zinc ion binding	
Yes	2	18104971	NP_536316	PTP4A2	protein tyrosine phosphatase type IVA, mem	2	3	membrane	protein amino acid dephosphorylation	hydrolase activity prenylated protein tyrosine phosphatase activity	
	2	18104973	NP_536317	PTP4A2	protein tyrosine phosphatase type IVA, mem	1	1	membrane	protein amino acid dephosphorylation	prenylated protein tyrosine phosphatase activity hydrolase activity	
	2	18375650	NP_542416	PTPN13	protein tyrosine phosphatase, non-receptor	9	14	cytoplasm cytoskeleton	protein amino acid dephosphorylation	hydrolase activity non-membrane spanning protein tyrosine phosphatase activity protein binding structural molecule activity	
	2	24308073	NP_056281	PTPN23	protein tyrosine phosphatase, non-receptor	4	5	not classified	protein amino acid dephosphorylation	hydrolase activity protein tyrosine phosphatase activity	
	2	18104993	NP_536859	PTPN6	protein tyrosine phosphatase, non-receptor	1	1	cytoskeleton membrane	apoptosis G-protein coupled receptor protein signaling pathway intracellular signaling cascade negative regulation of cell proliferation protein amino acid dephosphorylation	hydrolase activity protein binding protein tyrosine phosphatase activity	
	1,2	18860900	NP_002834	PTPRJ	protein tyrosine phosphatase, receptor type, J pr	1	1	integral to plasma membrane	cell-cell signaling protein amino acid dephosphorylation transmembrane receptor protein tyrosine kinase signaling pathway	hydrolase activity transmembrane receptor protein tyrosine phosphatase activity	Somatic Colon Cancer [MIM:114500]
	2	13677220	NP_109595	PTPRO	receptor-type protein tyrosine phosphatase	2	2	integral to membrane integral to plasma membrane membrane	protein amino acid dephosphorylation	hydrolase activity protein tyrosine phosphatase activity receptor activity transmembrane receptor protein tyrosine phosphatase activity	
	2	13677222	NP_109596	PTPRO	receptor-type protein tyrosine phosphatase	2	6	integral to membrane integral to plasma membrane membrane	protein amino acid dephosphorylation	protein tyrosine phosphatase activity receptor activity transmembrane receptor protein tyrosine phosphatase activity not classified	
	2	4757886	NP_004330	PTTG1IP	pituitary tumor-transforming gene 1 protein-1	1	1	cytoplasm integral to membrane membrane nucleus	development protein import into nucleus	GTP binding	

Supplementary Table I

	2	4506359	NP_000311	QDPR	quinoid dihydropteridine reductase	1	1	cytoplasm	amino acid metabolism dihydrobiopterin metabolism L-phenylalanine catabolism metabolism tetrahydrobiopterin biosynthesis	6,7-dihydropteridine reductase activity electron carrier activity	
	1,2	33695095	NP_057215	RAB10	ras-related GTP-binding protein RAB10	10	23	membrane	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
Yes	1,2	4758984	NP_004654	RAB11A	Ras-related protein Rab-11A	1	4	Golgi trans face	plasma membrane to endosome transport protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding syntaxin binding transporter activity	
Yes	1,2	4758986	NP_004209	RAB11B	RAB11B, member RAS oncogene family	7	33	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	50363228	NP_079427	RAB11FIP1	Rab coupling protein isoform 1	1	1	not classified	protein transport	not classified	
	1,2	4506363	NP_002861	RAB13	RAB13, member RAS oncogene family	2	4	intracellular tight junction	cell adhesion protein transport regulation of transcription, DNA- dependent small GTPase mediated signal transduction vesicle-mediated transport	ATP binding GTP binding GTPase activity nucleotide binding transcription factor binding	
	1,2	19923483	NP_057406	RAB14	GTPase Rab14	6	21	cytosol early endosome Golgi stack intracellular late endosome lysosome membrane fraction nuclear envelope- endoplasmic reticulum network perinuclear region plasma membrane rough endoplasmic reticulum trans-Golgi network transport vesicle	Golgi to endosome transport intracellular transport neurotransmitter secretion protein transport small GTPase mediated signal transduction vesicle-mediated transport	GTP binding GTPase activity nucleotide binding	
	2	38371739	NP_941959	RAB15	RAB15, member RAS oncogene family	1	3	membrane	protein transport small GTPase mediated signal transduction	nucleotide binding GTP binding	
	2	11967981	NP_071894	RAB17	RAB17, member RAS oncogene family	1	1	intracellular membrane	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
Yes	2	10880989	NP_067075	RAB18	RAB18, member RAS oncogene family	1	1	intracellular membrane	endocytosis protein transport regulation of transcription, DNA- dependent small GTPase mediated signal transduction	ATP binding GTP binding GTPase activity nucleotide binding transcription factor binding	
	2	56847628	NP_001008749	RAB19B	GTP-binding protein RAB19B	1	1	not classified	small GTPase mediated signal t	GTP binding nucleotide binding	

Supplementary Table I

Yes	1,2	4758988	NP_004152	RAB1A	RAB1A, member RAS oncogene family	5	10	cytoplasm endoplasmic reticulum intracellular membrane	ER to Golgi vesicle-mediated transport protein transport regulation of transcription, DNA-dependent small GTPase mediated signal transduction	ATP binding GTP binding nucleotide binding transcription factor binding	
Yes	1,2	13569962	NP_112243	RAB1B	RAB1B, member RAS oncogene family	3	11	Golgi apparatus intracellular membrane	protein transport regulation of transcription, DNA-dependent small GTPase mediated signal transduction	ATP binding GTP binding nucleotide binding transcription factor binding	
	1,2	4506365	NP_002856	RAB2	RAB2, member RAS oncogene family	6	12	endoplasmic reticulum	ER to Golgi vesicle-mediated transport protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	1,2	7661922	NP_055814	RAB21	RAB21, member RAS oncogene family	2	3	endosome	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	2	10190714	NP_065724	RAB22A	RAS-related protein RAB-22A	6	8	early endosome plasma membrane	endocytosis endosome organization and biogenesis protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	2	9966861	NP_065120	RAB25	RAB25	1	2	membrane	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	2	19923264	NP_004571	RAB27A	Ras-related protein Rab-27A	1	1	not classified	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	1,2	5729997	NP_004154	RAB27B	RAB27B, member RAS oncogene family	2	2	not classified	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	33589861	NP_006859	RAB31	RAB31, member RAS oncogene family	1	1	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	5803133	NP_006825	RAB32	RAB32, member RAS oncogene family	1	1	mitochondrion	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	2	31543537	NP_114140	RAB34	RAB39	1	1	membrane	protein transport small GTPase mediated signal transduction	nucleotide binding GTP binding	
	1,2	5803135	NP_006852	RAB35	RAB35, member RAS oncogene family	4	8	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	11641237	NP_071732	RAB38	RAB38	1	5	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	4506367	NP_002857	RAB3A	RAB3A, member RAS oncogene family	1	1	membrane synaptic vesicle	exocytosis neurotransmitter secretion protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	1,2	19923750	NP_002858	RAB3B	RAB3B, member RAS oncogene family	2	5	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	

Supplementary Table I

Yes	2	19923985	NP_612462	RAB3C	RAB3C, member RAS oncogene family	2	4	membrane	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	1,2	4759000	NP_004274	RAB3D	RAB3D, member RAS oncogene family	1	1	membrane	exocytosis hemocyte development (sensu Arthropoda) protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	2	55743118	NP_036365	RAB3GAP1	RAB3 GTPase-activating protein	1	1	soluble fraction	regulation of GTPase activity	Rab GTPase binding GTP binding	Warburg Micro Syndrome [MIM: 600118]
	2	50234889	NP_940892	RAB43	RAB43 protein	1	1	exosome (RNase complex) membrane	intracellular protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
Yes	1,2	19923262	NP_004153	RAB5A	RAB5A, member RAS oncogene family	6	13	early endosome	endocytosis protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	1,2	4506371	NP_002859	RAB5B	RAB5B, member RAS oncogene family	5	12	endosome intracellular	protein transport small GTPase mediated signal transduction	GTP binding GTP-dependent protein binding nucleotide binding	
	2	41393614	NP_958842	RAB5C	RAB5C, member RAS oncogene family isoform	5	57	not classified	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
Yes	2	96975097	NP_057661	RAB6B	RAB6B, member RAS oncogene family	1	3	Golgi apparatus	ER to Golgi vesicle-mediated transport protein transport retrograde vesicle-mediated transport, Golgi to ER small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	1,2	34147513	NP_004628	RAB7	RAB7, member RAS oncogene family	9	30	late endosome lysosome	endocytosis protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	2	4506375	NP_003920	RAB7L1	RAB7, member RAS oncogene family-like 1	4	5	membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	1,2	16933567	NP_005361	RAB8A	mel transforming oncogene	5	24	membrane	protein transport small GTPase mediated signal transduction	GTP binding nucleotide binding	
	1,2	7706563	NP_057614	RAB8B	RAB8B, member RAS oncogene family	2	8	intracellular membrane	protein transport regulation of transcription, DNA- dependent small GTPase mediated signal transduction	ATP binding GTP binding GTPase activity nucleotide binding transcription factor binding	
Yes	2	4759012	NP_004242	RAB9A	RAB9A, member RAS oncogene family	1	1	late endosome lysosome membrane	protein transport small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	

Supplementary Table I

Yes	2	9845509	NP_061485	RAC1	ras-related C3 botulinum toxin substrate 1 is	6	15	intracellular	actin filament polymerization lamellipodium biogenesis negative regulation of receptor mediated endocytosis positive regulation of Rho protein signal transduction ruffle organization and biogenesis small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
Yes	2	38505165	NP_942126	RAC1	ras-related C3 botulinum toxin substrate 1 is	6	21	intracellular	actin filament polymerization lamellipodium biogenesis negative regulation of receptor mediated endocytosis positive regulation of Rho protein signal transduction ruffle organization and biogenesis small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
Yes	2	4506381	NP_002863	RAC2	ras-related C3 botulinum toxin substrate 2	1	3	intracellular	signal transduction small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	2	33946329	NP_005393	RALA	ras related v-ral simian leukemia viral oncog	2	4	intracellular membrane	chemotaxis signal transduction small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	
Yes	1,2	4506405	NP_002872	RALB	v-ral simian leukemia viral oncogene homolog	9	21	intracellular membrane	signal transduction small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	
	2	5453555	NP_006316	RAN	ras-related nuclear protein	2	2	chromatin cytoplasm exosome (RNase complex) nuclear pore nucleus	androgen receptor signaling pathway DNA metabolism intracellular protein transport mitosis mitotic spindle organization and biogenesis positive regulation of transcription, DNA-dependent protein export from nucleus regulation of progression through cell c	androgen receptor binding chromatin binding GTP binding GTPase activity nucleotide binding protein binding transcription coactivator activity	
Yes	2	4506413	NP_002875	RAP1A	RAP1A, member of RAS oncogene family	11	37	intracellular membrane	cell cycle negative regulation of progression through cell cycle small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	
Yes	2	58331202	NP_001010933	RAP1A	RAP1A, member of RAS oncogene family	6	22	intracellular membrane	cell cycle negative regulation of progression through cell cycle small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	

Supplementary Table I

Yes	2	58219792	NP_001010942	RAP1B	RAP1B, member of RAS oncogene family	2	3	intracellular membrane	small GTPase mediated signal transduction	GTP binding nucleotide binding	
	2	20070314	NP_066982	RAP1GDS1	RAP1, GTP-GDP dissociation stimulator 1	1	3	not classified	not classified	binding GTPase activator activity	
Yes	2	10518344	NP_066361	RAP2A	RAP2A, member of RAS oncogene family	2	6	intracellular membrane	signal transduction small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	2	38201690	NP_002877	RAP2B	RAP2B, member of RAS oncogene family	1	1	intracellular membrane	small GTPase mediated signal transduction	GTP binding nucleotide binding	
	2	32129209	NP_067006	RAP2C	RAP2C, member of RAS oncogene family	1	1	intracellular membrane	small GTPase mediated signal transduction	nucleotide binding intracellular transporter activity	
	2	20070215	NP_006096	RAPGEF3	RAP guanine-nucleotide-exchange factor 3	2	4	cAMP-dependent protein kinase complex intracellular membrane fraction	cell proliferation protein amino acid phosphorylation regulation of small GTPase mediated signal transduction	cAMP binding cAMP-dependent protein kinase regulator activity guanyl-nucleotide exchange factor activity nucleotide binding	
	2	15149476	NP_002878	RARS	arginyl-tRNA synthetase	1	5	cytoplasm soluble fraction	arginyl-tRNA aminoacylation protein biosynthesis	arginine-tRNA ligase activity ATP binding ligase activity nucleotide binding	
	2	21361292	NP_005602	RBL2	retinoblastoma-like 2 (p130)	1	1	nucleus	cell cycle negative regulation of progression through cell cycle regulation of lipid kinase activity regulation of transcription, DNA-dependent transcription	DNA binding protein binding	
	2	55743122	NP_006735	RBP4	retinol-binding protein 4, plasma precursor	2	3	extracellular region extracellular space	response to stimulus transport visual perception	binding retinal binding retinol binding transporter activity	Retinol-Binding Protein Deficiency [MIM:180250]
	1,2	13899352	NP_113679	RBP5	retinol binding protein 5, cellular	1	8	cytoplasm	transport	intracellular transporter activity lipid binding retinal binding retinol binding	
	1,2	4506467	NP_002897	RDX	radixin	16	23	actin cytoskeleton cytoplasm plasma membrane	barbed-end actin filament capping cytoskeletal anchoring	actin binding binding structural molecule activity	Autosomal Recessive Deafness 24 [MIM:611022]
	2	4506477	NP_002901	RENBP	renin binding protein	1	1	intracellular	blood pressure regulation mannose metabolism	endopeptidase inhibitor activity isomerase activity mannose-6-phosphate isomerase activity N-acetylglucosamine 2-epimerase activity	
	2	32528306	NP_002904	RFC1	replication factor C large subunit	1	1	DNA replication factor C complex intracellular nucleus	DNA-dependent DNA replication regulation of transcription, DNA-dependent telomere maintenance via telomerase transcription	DNA binding enzyme activator activity metal ion binding nucleoside-triphosphatase activity nucleotide binding protein binding zinc ion binding methyltransferase activity	
	2	22748659	NP_689505	RG9MTD2	RNA (guanine-9-) methyltransferase domain	1	1	not classified	tRNA processing	RNA binding transferase activity tRNA methyltransferase activity GTP binding	

Supplementary Table I

	1,2	7706683	NP_057405	RHCG	Rhesus blood group, C glycoprotein	5	8	apical plasma membrane integral to plasma membrane plasma membrane	amine transport ammonium transport cell ion homeostasis epithelial cell differentiation homeostasis	ammonium transporter activity	
	2	5032041	NP_005605	RHEB	Ras homolog enriched in brain	1	2	intracellular membrane	signal transduction small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	1,2	10835049	NP_001655	RHOA	ras homolog gene family, member A	4	13	cytoskeleton intracellular membrane	actin cytoskeleton organization and biogenesis positive regulation of I-kappaB kinase/NF-kappaB cascade positive regulation of NF-kappaB import into nucleus Rho protein signal transduction small GTPase mediated signal transduction	GTP binding GTPase activity magnesium ion binding nucleotide binding protein binding signal transducer activity	
	2	4757764	NP_004031	RHOB	ras homolog gene family, member B	1	1	endosome membrane intracellular nucleus plasma membrane	angiogenesis cell adhesion cell cycle cell differentiation endosome to lysosome transport negative regulation of progression through cell cycle positive regulation of angiogenesis protein transport small GTPase mediated signal transduction transformed cell	GTP binding nucleotide binding protein binding	
	2	28395033	NP_786886	RHOC	ras homolog gene family, member C precursor	1	1	intracellular membrane	positive regulation of I-kappaB kinase/NF-kappaB cascade small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding signal transducer activity	
	2	38016957	NP_061907	RHOF	ras homolog gene family, member F	1	3	cytoskeleton intracellular membrane	actin filament organization small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding	
	1,2	46249393	NP_001656	RHOG	ras homolog gene family, member G	5	11	intracellular membrane	positive regulation of cell proliferation regulation of progression through cell cycle Rho protein signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
Yes	2	50263042	NP_036381	RHOQ	ras-like protein TC10	1	2	intracellular plasma membrane	small GTPase mediated signal transduction	GTP binding GTPase activity nucleotide binding protein binding	
	1,2	7657520	NP_055187	RNF11	ring finger protein 11	1	6	ubiquitin ligase complex	protein ubiquitination during ubiquitination	DNA binding metal ion binding protein binding zinc ion binding	
	2	66571327	NP_065965	RNF213	ring finger protein 213	1	1	not classified	not classified	protein binding zinc ion binding protein binding	

Supplementary Table I

	2	42794608	NP_976318	RNH1	ribonuclease/angiogenin inhibitor	4	6	not classified	mRNA catabolism regulation of angiogenesis	protein binding ribonuclease inhibitor activity transferase activity	
	2	40316915	NP_064601	RNPEP	arginyl aminopeptidase (aminopeptidase B)	4	5	extracellular region Golgi apparatus plasma membrane	protein processing proteolysis	aminopeptidase B activity epoxide hydrolase activity membrane alanyl aminopeptidase activity metal ion binding zinc ion binding	
	2	61888896	NP_002933	ROBO2	roundabout, axon guidance receptor, homolog 2	1	1	cell surface integral to membrane membrane	brain development cell adhesion cell differentiation chemotaxis homophilic cell adhesion positive regulation of axonogenesis ureteric bud development	axon guidance receptor activity identical protein binding	Vesicoureteral Reflux 2 [MIM:610878]
	2	5902060	NP_008846	RP2	XRP2 protein	3	5	not classified	beta-tubulin folding CTP biosynthesis GTP biosynthesis response to stimulus UTP biosynthesis visual perception	ATP binding nucleoside diphosphate kinase activity unfolded protein binding	X-Linked Retinitis Pigmentosa 2 [MIM:312600]
Yes	2	4506597	NP_000967	RPL12	ribosomal protein L12	1	1	cytosolic large ribosomal subunit (sensu Eukaryota) intracellular ribosome	protein biosynthesis	RNA binding structural constituent of ribosome	
Yes	2	4506613	NP_000974	RPL22	ribosomal protein L22 proprotein	1	1	cytosolic large ribosomal subunit (sensu Eukaryota) intracellular ribosome	protein biosynthesis	heparin binding RNA binding structural constituent of ribosome	
Yes	2	4506679	NP_001005	RPS10	ribosomal protein S10	1	1	cytosolic small ribosomal subunit (sensu Eukaryota) ribonucleoprotein complex	protein biosynthesis	RNA binding structural constituent of ribosome	
	2	4506681	NP_001006	RPS11	ribosomal protein S11	1	1	cytosolic small ribosomal subunit (sensu Eukaryota) intracellular ribosome	protein biosynthesis	rRNA binding structural constituent of ribosome	
Yes	2	4506691	NP_001011	RPS16	ribosomal protein S16	1	1	cytosolic small ribosomal subunit (sensu Eukaryota) intracellular ribosome	protein biosynthesis	structural constituent of ribosome	
	2	11968182	NP_072045	RPS18	ribosomal protein S18	2	2	cytosolic small ribosomal subunit (sensu Eukaryota) intracellular	protein biosynthesis	rRNA binding structural constituent of ribosome	
Yes	2	15011936	NP_001020	RPS26	ribosomal protein S26	1	1	cytosolic small ribosomal subunit (sensu Eukaryota) intracellular ribosome	protein biosynthesis	RNA binding structural constituent of ribosome	
	2	5454028	NP_006261	RRAS	related RAS viral (r-ras) oncogene homolog	4	11	intracellular	Ras protein signal transduction	GTP binding GTPase activity nucleotide binding	

Supplementary Table I

	1,2	21361416	NP_036382	RRAS2	related RAS viral (r-ras) oncogene homolog	4	10	intracellular membrane	small GTPase mediated signal transduction	GTP binding nucleotide binding protein binding	
	2	51173737	NP_001003699	RREB1	ras responsive element binding protein 1 isoform 1	1	2	cytoplasm intracellular nuclear body nucleus	development Ras protein signal transduction regulation of transcription, DNA-dependent transcription transcription from RNA polymerase II promoter	DNA binding metal ion binding nucleic acid binding transcriptional activator activity zinc ion binding	
	2	55741719	NP_055621	RUSC2	RUN and SH3 domain containing 2	1	1	not classified	not classified	not classified	
	2	10863871	NP_000531	RYR1	skeletal muscle ryanodine receptor isoform 1	1	1	integral to plasma membrane membrane smooth endoplasmic reticulum	calcium ion homeostasis calcium ion transport ion transport muscle contraction	calcium ion binding calcium-release channel activity ion channel activity receptor activity ryanodine-sensitive calcium-release channel activity	Malignant Hyperthermia [MIM:145600] Central Core Disease [MIM:117000] Minicore Myopathy with External Ophthalmoplegia [MIM:255320]
	2	4506761	NP_002957	S100A10	S100 calcium binding protein A10	1	1	not classified	not classified	calcium ion binding	
Yes	1,2	5032057	NP_005611	S100A11	S100 calcium binding protein A11	4	15	cytoplasm nucleus ruffle	negative regulation of cell proliferation negative regulation of DNA replication signal transduction	calcium ion binding calcium-dependent protein binding protein homodimerization activity S100 beta binding	
	2	10190712	NP_065723	S100A14	S100 calcium binding protein A14	2	3	not classified	not classified	calcium ion binding	
	1,2	17933772	NP_525127	S100A16	S100 calcium binding protein A16	1	5	not classified	not classified	calcium ion binding	
	2	4506765	NP_002952	S100A4	S100 calcium-binding protein A4	1	1	not classified	not classified	calcium ion binding	
	1,2	7657532	NP_055439	S100A6	S100 calcium-binding protein A6	2	8	cytoplasm nuclear envelope nucleus ruffle	axonogenesis cell cycle cell proliferation cell-cell signaling positive regulation of fibroblast proliferation regulation of progression through cell cycle signal transduction	calcium ion binding calcium-dependent protein binding growth factor activity protein homodimerization activity S100 beta binding	
	2	21614544	NP_002955	S100A8	S100 calcium-binding protein A8	6	28	extracellular space	inflammatory response	calcium ion binding	
	2	4506773	NP_002956	S100A9	S100 calcium-binding protein A9	7	60	extracellular space	cell-cell signaling inflammatory response	calcium ion binding signal transducer activity	
	1,2	5174663	NP_005971	S100P	S100 calcium binding protein P	1	11	cytoplasm nucleus	endothelial cell migration	calcium ion binding calcium-dependent protein binding magnesium ion binding	
	1,2	31542301	NP_056195	SAMM50	sorting and assembly machinery component 1	1	1	integral to membrane membrane mitochondrion outer membrane	not classified	not classified	
	2	19424126	NP_597998	SAT2	spermidine/spermine N1-acetyltransferase 2	1	1	not classified	not classified	diamine N-acetyltransferase activity N-acetyltransferase activity transferase activity metal ion binding	
	2	5031631	NP_005497	SCARB2	scavenger receptor class B, member 2	9	18	integral to plasma membrane lysosomal membrane membrane membrane fraction	cell adhesion	receptor activity	
	2	21536308	NP_659001	SCEL	sciellin isoform b	2	3	cornified envelope cytoplasm	embryonic development epidermis development keratinocyte differentiation	protein binding zinc ion binding sodium ion binding	

Supplementary Table I

2	14916473	NP_149119	SCIN	scinderin	3	6	cell cortex cytoplasm cytoskeleton	actin filament severing actin nucleation barbed-end actin filament capping calcium ion-dependent exocytosis negative regulation of cell proliferation positive regulation of apoptosis positive regulation of megakaryocyte differentiation positive regulation	actin filament binding calcium ion binding phosphatidylinositol binding phosphatidylinositol-4,5-bisphosphate binding phosphatidylserine binding
2	5730033	NP_006505	SCN10A	sodium channel, voltage-gated, type X, alpha 1	1	11	membrane membrane fraction voltage-gated sodium channel complex	ion transport sensory perception sodium ion transport synaptic transmission	sodium ion binding voltage-gated calcium channel activity voltage-gated sodium channel activity
2	115583667	NP_054858	SCN11A	sodium channel, voltage-gated, type XI, alpha 1	1	1	membrane membrane fraction voltage-gated sodium channel complex	ion transport response to drug sodium ion transport synaptic transmission	voltage-gated sodium channel activity dipeptidase activity
2	11055992	NP_067639	SCPEP1	serine carboxypeptidase 1 precursor protein	1	1	not classified	proteolysis	peptidase activity serine carboxypeptidase activity
2	24308406	NP_612364	SCRN2	secernin 2	1	1	not classified	proteolysis	cytoskeletal protein binding
2	29568086	NP_002988	SDC1	syndecan 1 precursor	1	2	integral to plasma membrane membrane	not classified	cytoskeletal protein binding
2	55749515	NP_001007070	SDCBP	syntenin isoform 3	11	77	adherens junction cytoskeleton endoplasmic reticulum interleukin-5 receptor complex membrane nucleus	actin cytoskeleton organization and biogenesis intracellular signaling cascade metabolism protein targeting to membrane substrate-bound cell migration, cell extension synaptic transmission	cytoskeletal adaptor activity interleukin-5 receptor binding neurexin binding oxidoreductase activity protein binding protein heterodimerization activity syndecan binding
2	56243522	NP_005616	SDCBP	syntenin isoform 1	16	174	adherens junction cytoskeleton endoplasmic reticulum interleukin-5 receptor complex membrane nucleus	actin cytoskeleton organization and biogenesis intracellular signaling cascade metabolism protein targeting to membrane substrate-bound cell migration, cell extension synaptic transmission	interleukin-5 receptor binding neurexin binding oxidoreductase activity protein binding protein heterodimerization activity syndecan binding cytoskeletal adaptor activity not classified
2	38044114	NP_536737	SDCBP2	syndecan binding protein 2 isoform a	1	1	cytoplasm plasma membrane	intracellular signaling cascade intracellular transport nervous system development	protein C-terminus binding protein heterodimerization activity protein homodimerization activity
2	34335134	NP_899195	SEC13L1	SEC13 protein	1	1	not classified	intracellular protein transport	not classified
2	4506869	NP_002995	SECTM1	secreted and transmembrane 1 precursor	2	5	extracellular space Golgi apparatus integral to membrane membrane	immune response mesoderm development	cytokine activity signal transducer activity
2	16306550	NP_003935	SELENBP1	selenium binding protein 1	5	7	not classified	not classified	selenium binding

Supplementary Table 1

	2	71834872	NP_849196	SERINC2	tumor differentially expressed 2-like	1	1	membrane	phosphatidylserine metabolism positive regulation of transferase activity	not classified	
	2	30172548	NP_840060	SERINC5	developmentally regulated protein TPO1	1	1	membrane	phosphatidylserine metabolism positive regulation of transferase activity	protein binding	
	2	50363219	NP_001002236	SERPINA1	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	4	15	extracellular region extracellular space	acute-phase response	protein binding serine-type endopeptidase inhibitor activity	
	1,2	50659080	NP_001076	SERPINA3	serpin peptidase inhibitor, clade A, member 3 precursor	3	6	extracellular region intracellular	acute-phase response inflammatory response regulation of lipid metabolism	chymotrypsin inhibitor activity DNA binding protein binding	
	2	21361302	NP_006206	SERPINA4	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 4	2	3	not classified	not classified	protein binding serine-type endopeptidase inhibitor activity	
	1,2	21361195	NP_000615	SERPINA5	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 5	12	74	extracellular region membrane protein complex	fusion of sperm to egg plasma membrane spermatogenesis	acrosin binding heparin binding serine-type endopeptidase inhibitor activity	
	2	4507377	NP_000345	SERPINA7	serine (or cysteine) proteinase inhibitor, clade A	1	1	extracellular region extracellular space	not classified	binding serine-type endopeptidase inhibitor activity	
	2	13489087	NP_109591	SERPINB1	serine (or cysteine) proteinase inhibitor, clade B	7	9	cytoplasm	not classified	serine-type endopeptidase inhibitor activity	
	2	8393956	NP_036529	SERPINB13	serine (or cysteine) proteinase inhibitor, clade B	1	3	not classified	regulation of proteolysis response to UV	serine-type endopeptidase inhibitor activity	
	2	5902072	NP_008850	SERPINB3	serine (or cysteine) proteinase inhibitor, clade B	3	7	not classified	not classified	serine-type endopeptidase inhibitor activity	
Yes	2	28076869	NP_002965	SERPINB4	serine (or cysteine) proteinase inhibitor, clade B	2	12	cytoplasm	immune response regulation of proteolysis	serine-type endopeptidase inhibitor activity	
	2	41152086	NP_004559	SERPINB6	serine (or cysteine) proteinase inhibitor, clade B	2	6	cytosol	not classified	protein binding serine-type endopeptidase inhibitor activity	
	2	39725934	NP_002606	SERPINF1	serine (or cysteine) proteinase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1	2	7	extracellular region	cell proliferation development negative regulation of angiogenesis positive regulation of neurogenesis	serine-type endopeptidase inhibitor activity	
	2	73858568	NP_000053	SERPING1	complement component 1 inhibitor precursor	7	14	extracellular region	blood coagulation circulation complement activation, classical pathway innate immune response	serine-type endopeptidase inhibitor activity	Hereditary Angioedema Type I [MIM:106100]
Yes	2	5453886	NP_006208	SERPINI2	serpin peptidase inhibitor, clade I (pancrein), member 2	1	2	not classified	cell motility	serine-type endopeptidase inhibitor activity	
	1,2	5454052	NP_006133	SFN	stratifin	4	10	cytoplasm extracellular space	cell proliferation negative regulation of protein kinase activity regulation of progression through cell cycle signal transduction	protein domain specific binding	
	2	40786408	NP_955376	SFT2D2	SFT2 domain containing 2	1	1	integral to membrane membrane	protein transport	SH3/SH2 adaptor activity	

Supplementary Table 1

	2	4506925	NP_003013	SH3BGR1	SH3 domain binding glutamic acid-rich prote	1	3	nucleus	not classified	glycine hydroxymethyltransferase activity	
	2	13775198	NP_112576	SH3BGR3	SH3 domain binding glutamic acid-rich prote	1	2	nucleus	not classified	not classified	
	2	7657562	NP_055336	SH3BP4	SH3-domain binding protein 4	1	1	nucleus	cell cycle endocytosis	signal transducer activity	
Yes	2	32261324	NP_003020	SHC1	SHC (Src homology 2 domain containing) transforming protein 1 isoform p52Shc	1	1	plasma membrane	activation of MAPK activity intracellular signaling cascade positive regulation of cell proliferation positive regulation of mitosis regulation of cell growth regulation of epidermal growth factor receptor activity	insulin receptor binding insulin-like growth factor receptor binding phospholipid binding protein binding transmembrane receptor protein tyrosine kinase adaptor protein activity	
	2	22547189	NP_683718	SHMT1	serine hydroxymethyltransferase 1 (soluble)	8	20	cytosol	glycine metabolism L-serine catabolism L-serine metabolism one-carbon compound metabolism	glycine hydroxymethyltransferase activity transferase activity	
	2	19923315	NP_005403	SHMT2	serine hydroxymethyltransferase 2 (mitocho	1	4	mitochondrion	glycine metabolism L-serine metabolism one-carbon compound metabolism	transferase activity hydrolase activity	
	2	4502175	NP_001640	SHROOM2	apical protein of Xenopus-like	1	1	not classified	not classified	amiloride-sensitive sodium channel activity protein binding	
Yes	2	157364974	NP_001032	SI	sucrase-isomaltase (alpha-glucosidase)	1	11	brush border Golgi apparatus integral to membrane membrane	carbohydrate metabolism metabolism	hydrolase activity, hydrolyzing O-glycosyl compounds oligo-1,6-glucosidase activity sucrose alpha-glucosidase activity	
	2	24850115	NP_733746	SIAE	cytosolic sialic acid 9-O-acetylerase hom	1	1	lysosome	not classified	serine esterase activity sialate O-acetylerase activity not classified	
	2	91105767	NP_00103511	SIRPA	signal-regulatory protein alpha precursor	1	1	integral to membrane plasma membrane	cell adhesion	sodium ion binding	
	1,2	4557849	NP_000329	SLC12A1	sodium potassium chloride cotransporter 2	25	94	integral to membrane membrane membrane fraction plasma membrane	chloride transport ion transport potassium ion transport sodium ion transport	potassium ion binding sodium ion binding sodium:potassium:chloride symporter activity symporter activity transporter activity	Barter Syndrome, Antenatal, Type 1 [MIM:601678]
	1,2	4506977	NP_000330	SLC12A3	solute carrier family 12 (sodium/chloride transp	28	102	integral to plasma membrane membrane membrane fraction	chloride transport ion transport sodium ion transport	sodium ion binding sodium:chloride symporter activity symporter activity transporter activity	Gitelman Syndrome [MIM:263800]
	2	31881740	NP_064631	SLC12A9	solute carrier family 12 (potassium/chloride	1	1	integral to membrane membrane	transport	cation:chloride symporter activity protein binding	
	2	4506979	NP_003975	SLC13A2	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	5	10	integral to plasma membrane membrane membrane fraction	ion transport sodium ion transport	low affinity sodium:dicarboxylate symporter activity sodium ion binding symporter activity transporter activity	
	2	31377715	NP_073740	SLC13A3	solute carrier family 13 member 3 isoform a	2	2	integral to membrane membrane	ion transport	sodium ion binding symporter activity transporter activity	
	2	58761541	NP_00101155	SLC13A3	solute carrier family 13 member 3 isoform b	1	3	integral to membrane membrane	ion transport sodium ion transport	symporter activity transporter activity oligopeptide transporter activity	

Supplementary Table I

	2	31543624	NP_066568	SLC15A2	solute carrier family 15 (H+/peptide transpor	1	1	integral to plasma membrane	drug transport oligopeptide transport protein transport	peptide:hydrogen symporter activity protein binding symporter activity transporter activity organic cation transporter activity	
	1,2	66773030	NP_004161	SLC1A1	solute carrier family 1, member 1	3	6	integral to plasma membrane membrane fraction	dicarboxylic acid transport glutamate transport synaptic transmission transport	L-glutamate transporter activity sodium:dicarboxylate symporter activity symporter activity	
	2	8923870	NP_060954	SLC22A11	solute carrier family 22 member 11	2	9	integral to plasma membrane	ion transport	sodium-independent organic anion transporter activity transporter activity	
	2	24497485	NP_653186	SLC22A12	urate anion exchanger 1 isoform a	1	2	apical plasma membrane brush border membrane integral to membrane plasma membrane	cell homeostasis transport urate transport	PDZ domain binding protein binding transporter activity urate transporter activity	Renal Hypouricemia [MIM:220150]
	2	4758852	NP_004247	SLC22A13	organic cation transporter like 3	1	1	integral to plasma membrane membrane fraction	transport	transporter activity ATP binding	
	2	4507001	NP_003049	SLC22A2	solute carrier family 22 member 2 isoform a	2	3	integral to membrane integral to plasma membrane membrane fraction	fluid secretion ion transport organic cation transport sodium ion transport transport	ATP binding ion transporter activity nucleotide binding organic cation transporter activity sodium ion binding symporter activity transporter activity	
	2	4507005	NP_003051	SLC22A5	solute carrier family 22 member 5	1	1	integral to membrane	carnitine transport ion transport sodium ion transport transport	carnitine transporter activity ion transporter activity nucleotide binding sodium ion binding symporter activity transporter activity ATP binding	
	2	20070188	NP_004781	SLC22A6	solute carrier family 22 member 6 isoform a	1	3	integral to membrane integral to plasma membrane membrane fraction	alpha-ketoglutarate transport ion transport organic anion transport sodium ion transport	ATP binding ion transporter activity nucleotide binding organic anion transporter activity sodium ion binding symporter activity	
	2	24497480	NP_695010	SLC22A6	solute carrier family 22 member 6 isoform d	1	1	integral to membrane integral to plasma membrane membrane fraction	alpha-ketoglutarate transport ion transport organic anion transport sodium ion transport	ion transporter activity nucleotide binding organic anion transporter activity sodium ion binding symporter activity nucleobase transporter activity	
	2	24497499	NP_004245	SLC22A8	solute carrier family 22 member 8	1	2	integral to membrane integral to plasma membrane membrane fraction	ion transport organic anion transport sodium ion transport transport	nucleotide binding organic anion transporter activity sodium ion binding symporter activity transporter activity	
	2	44680143	NP_689898	SLC23A1	solute carrier family 23 (nucleobase transpo	6	10	integral to plasma membrane membrane fraction	ion transport nucleobase transport nucleobase, nucleoside, nucleotide and nucleic acid metabolism sodium ion transport	nucleobase transporter activity sodium ion binding symporter activity transporter activity	

Supplementary Table I

	2	44680145	NP_005838	SLC23A1	solute carrier family 23 (nucleobase transp	4	6	integral to plasma membrane membrane fraction	ion transport nucleobase transport nucleobase, nucleoside, nucleotide and nucleic acid metabolism sodium ion transport	sodium ion binding symporter activity transporter activity binding	
	2	21389315	NP_005975	SLC25A1	solute carrier family 25 (mitochondrial carrie	1	1	integral to membrane mitochondrial inner membrane mitochondrion	mitochondrial citrate transport transport	citrate transporter activity transporter activity adenine transporter activity	
	2	47132595	NP_998776	SLC25A3	solute carrier family 25 member 3 isoform b pre	1	5	integral to plasma membrane mitochondrial inner membrane mitochondrion	generation of precursor metabolites and energy transport	binding phosphate carrier activity symporter activity	Mitochondrial Phosphate Carrier Deficiency [MIM:610773]
Yes	2	55749577	NP_001142	SLC25A4	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	1	1	integral to membrane integral to plasma membrane mitochondrial inner membrane mitochondrion	generation of precursor metabolites and energy mitochondrial genome maintenance mitochondrial transport transport	adenine transporter activity binding transporter activity	
Yes	2	156071459	NP_001143	SLC25A5	solute carrier family 25, member 5	2	1	integral to plasma membrane mitochondrial inner membrane mitochondrion	mitochondrial transport transport	adenine transporter activity binding transporter activity	
Yes	2	156071462	NP_001627	SLC25A6	solute carrier family 25, member A6	2	3	integral to membrane mitochondrial inner membrane mitochondrial inner membrane presequence translocase complex mitochondrion	apoptosis mitochondrial transport transport	ATP:ADP antiporter activity binding transporter activity	
	2	85540447	NP_775897	SLC26A11	solute carrier family 26, member 11	1	1	integral to membrane	transport	transporter activity	
	2	4505697	NP_000432	SLC26A4	pendrin	2	4	integral to membrane membrane	sensory perception of sound sulfate transport transport	chloride ion binding chloride transporter activity iodide transporter activity sulfate porter activity transporter activity	Pendred Syndrome [MIM:274600] Deafness, Autosomal Recessive 4 [MIM:600791]
	2	4503653	NP_003636	SLC27A2	solute carrier family 27 (fatty acid transporte	1	1	endoplasmic reticulum integral to membrane membrane peroxisomal matrix peroxisome	lipid metabolism metabolism very-long-chain fatty acid metabolism	catalytic activity ligase activity long-chain-fatty-acid-CoA ligase activity nucleotide binding	
	1,2	4507013	NP_003030	SLC2A5	solute carrier family 2 (facilitated glucose/fru	12	22	integral to membrane membrane plasma membrane	carbohydrate metabolism carbohydrate transport fructose transport glucose transport	fructose transporter activity glucose transporter activity sugar porter activity transporter activity	
	2	32401453	NP_861441	SLC36A2	solute carrier family 36 (proton/amino acid s	7	15	not classified	not classified	not classified	
	2	27734777	NP_775867	SLC39A5	solute carrier family 39 (metal ion transporte	1	1	integral to membrane membrane	ion transport metal ion transport zinc ion transport	zinc ion binding calcium:sodium antiporter activity	

Supplementary Table I

	2	4507019	NP_000332	SLC3A1	solute carrier family 3, member 1	14	25	integral to plasma membrane membrane fraction	amino acid metabolism basic amino acid transport carbohydrate metabolism L-cystine transport transport	catalytic activity cation binding L-cystine transporter activity	Cystinuria [MIM:220100]
	2	61744475	NP_001012679	SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a	10	18	cell surface integral to membrane membrane	amino acid transport calcium ion transport carbohydrate metabolism cell growth tryptophan transport	calcium:sodium antiporter activity catalytic activity cation binding protein binding	
	2	65506891	NP_002385	SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a	3	5	cell surface integral to membrane membrane	amino acid transport calcium ion transport carbohydrate metabolism cell growth tryptophan transport	catalytic activity cation binding protein binding not classified	
	2	18034692	NP_536856	SLC44A1	CDW92 antigen	1	1	integral to membrane membrane	transport	transporter activity	
	1,2	31377727	NP_065161	SLC44A2	CTL2 protein	18	70	integral to membrane membrane	positive regulation of I-kappaB kinase/NF-kappaB cascade transport	signal transducer activity	
	1,2	13376868	NP_079533	SLC44A4	NG22 protein isoform 1	6	59	integral to membrane membrane	not classified	not classified	Sialidosis 1 [MIM:606107]
	2	4507021	NP_000333	SLC4A1	solute carrier family 4, anion exchanger, member 1	2	2	cortical cytoskeleton integral to plasma membrane membrane	anion transport cell ion homeostasis	anion transporter activity inorganic anion exchanger activity protein binding	Defective kidney acid secretion leading to distal renal tubular acidosis [MIM:179800]
	2	4507025	NP_003750	SLC4A4	solute carrier family 4, sodium bicarbonate cotransporter 1	2	3	integral to membrane membrane	anion transport sodium ion transport	inorganic anion exchanger activity sodium ion binding symporter activity	Renal Tubular Acidosis, Proximal, With Ocular Abnormalities [MIM:604278]
Yes	2	90568034	NP_004849	SLC4A8	solute carrier family 4, sodium bicarbonate cotransporter 2	1	1	integral to membrane membrane	anion transport	inorganic anion exchanger activity	
	2	4507031	NP_000334	SLC5A1	solute carrier family 5 (sodium/glucose cotransporter 1)	2	3	integral to plasma membrane membrane	glucose transport ion transport sodium ion transport	glucose:sodium symporter activity sodium ion binding sugar porter activity symporter activity transporter activity	Glucose/Galactose Malabsorption [MIM:606824]
	2	31982945	NP_689564	SLC5A10	solute carrier family 5 (sodium/glucose cotransporter 10)	2	2	integral to membrane membrane	ion transport sodium ion transport	transporter activity	
	2	109659836	NP_001035913	SLC5A10	solute carrier family 5 (sodium/glucose cotransporter 10)	1	3	integral to membrane membrane	ion transport sodium ion transport	transporter activity	
	2	30425416	NP_848593	SLC5A12	solute carrier family 5 (sodium/glucose cotransporter 12)	2	4	integral to membrane membrane	ion transport sodium ion transport transport	transporter activity	
	2	4507033	NP_003032	SLC5A2	solute carrier family 5 (sodium/glucose cotransporter 2)	4	9	integral to membrane membrane	carbohydrate metabolism ion transport sodium ion transport	low-affinity glucose:sodium symporter activity sodium ion binding sugar porter activity symporter activity transporter activity	Renal Glucosuria [MIM:233100]
	2	33942076	NP_666018	SLC5A8	solute carrier family 5 (iodide transporter), member 1	1	1	integral to membrane membrane	ion transport sodium ion transport	transporter activity	
	2	58531219	NP_001011547	SLC5A9	solute carrier family 5 (sodium/glucose cotransporter 9)	1	3	integral to membrane membrane	ion transport sodium ion transport	protein binding	

Supplementary Table 1

	2	21361581	NP_057699	SLC6A13	solute carrier family 6 (neurotransmitter tran	1	1	integral to plasma membrane	neurotransmitter transport	gamma-aminobutyric acid:sodium symporter activity neurotransmitter:sodium symporter activity symporter activity	
	2	51468073	NP_001003841	SLC6A19	solute carrier family 6, member 19	4	8	integral to plasma membrane	amino acid transport neurotransmitter transport	amino acid transporter activity neurotransmitter:sodium symporter activity symporter activity	Hartnup Disorder [MIM:234500]
	2	6806921	NP_004165	SLC9A3	solute carrier family 9 (sodium/hydrogen ex	2	3	integral to membrane	ion transport regulation of pH sodium ion transport	antiporter activity sodium ion binding sodium:hydrogen antiporter activity solute:hydrogen antiporter activity	
	1,2	4759140	NP_004243	SLC9A3R1	solute carrier family 9 (sodium/hydrogen ex	4	9	actin cytoskeleton	protein complex assembly Wnt receptor signaling pathway	beta-2 adrenergic receptor binding protein binding sodium:hydrogen antiporter regulator activity	
	2	71897229	NP_004776	SLC9A3R2	solute carrier family 9 isoform 3 regulator 2	1	1	nucleus plasma membrane	intracellular signaling cascade protein complex assembly	ATP binding	
	2	38679890	NP_851322	SLCO4C1	solute carrier organic anion transporter fam	2	2	integral to membrane	transport	transporter activity	
	2	4759146	NP_004778	SLIT2	slit homolog 2	1	1	extracellular region extracellular space integral to membrane	cell differentiation chemotaxis G-protein coupled receptor protein signaling pathway glial cell migration induction of negative chemotaxis mesoderm migration motor axon guidance nervous system development neuron recognition positive regulation of axonogen	calcium ion binding follicle-stimulating hormone receptor activity receptor binding transferase activity	
	2	41281453	NP_055535	SLK	serine/threonine kinase 2	9	19	not classified	apoptosis nucleotide-excision repair protein amino acid phosphorylation	ATP binding DNA binding nuclease activity nucleotide binding protein serine/threonine kinase activity transferase activity	
	2	110347425	NP_001036016	SMC2	structural maintenance of chromosomes 2-li	1	3	chromosome condensin complex cytoplasm nuclear chromosome nucleus	cell cycle cell division chromosome organization and biogenesis mitosis mitotic chromosome condensation	nucleotide binding protein binding protein heterodimerization activity protein binding	
	2	5032099	NP_005622	SMO	smoothened	1	1	integral to membrane membrane membrane fraction	development G-protein coupled receptor protein signaling pathway	G-protein coupled receptor activity non-G-protein coupled 7TM receptor activity protein binding receptor activity	
	1,2	57242798	NP_055289	SMPDL3B	acid sphingomyelinase-like phosphodiester	1	2	not classified	metabolism	hydrolase activity, acting on glycosyl bonds	

Supplementary Table I

	2	31317290	NP_851994	SMURF1	Smad ubiquitination regulatory factor 1 isofo	1	1	intracellular	cell differentiation ectoderm development negative regulation of BMP signaling pathway protein ubiquitination during ubiquitin-dependent protein catabolism ubiquitin cycle	ligase activity ubiquitin-protein ligase activity	
	1,2	16418371	NP_443102	SNAG1	sorting nexin associated golgi protein 1	4	7	not classified	intracellular signaling cascade protein transport	phosphoinositide binding protein binding	
	2	18765731	NP_570710	SNAP23	synaptosomal-associated protein 23 isoform	1	1	plasma membrane synapse synaptosome	membrane fusion post-Golgi-vesicle-mediated transport protein transport vesicle targeting	t-SNARE activity phosphoinositide binding	
	1,2	4507113	NP_003078	SNCG	synuclein, gamma (breast cancer-specific p	9	20	cytoplasm	not classified	not classified	
	2	21361380	NP_009172	SNF8	EAP30 subunit of ELL complex	1	1	nucleus	protein transport regulation of transcription, DNA- dependent transcription	not classified	
	2	23111030	NP_037478	SNX12	sorting nexin 12	1	3	not classified	intracellular signaling cascade protein transport	protein binding phosphoinositide binding	
	2	4507143	NP_003786	SNX3	sorting nexin 3	1	1	not classified	endocytosis intracellular signaling cascade protein transport	phosphoinositide binding protein binding	
	2	7706706	NP_057308	SNX9	sorting nexin 9	2	3	not classified	intracellular signaling cascade protein localization protein transport	protein binding SH3/SH2 adaptor activity lipid transporter activity	
	1,2	4507149	NP_000445	SOD1	superoxide dismutase 1, soluble	2	4	cytoplasm	nervous system development response to oxidative stress superoxide metabolism	antioxidant activity copper ion binding copper, zinc superoxide dismutase activity metal ion binding oxidoreductase activity zinc ion binding	
	2	4507155	NP_003095	SORD	sorbitol dehydrogenase	3	11	not classified	sorbitol metabolism visual perception	L-iditol 2-dehydrogenase activity metal ion binding oxidoreductase activity zinc ion binding	
	2	4507157	NP_003096	SORL1	sortilin-related receptor containing LDLR cla	1	1	integral to plasma membrane membrane	cholesterol metabolism intracellular protein transport lipid metabolism lipid transport receptor mediated endocytosis steroid metabolism	protein transporter activity transmembrane receptor activity not classified	
	2	27436920	NP_003962	SPAG9	sperm associated antigen 9 isoform 1	2	2	not classified	not classified	not classified	
	2	40806170	NP_955468	SPAST	spastin isoform 2	3	4	nucleus	not classified	ATP binding nucleoside-triphosphatase activity nucleotide binding	
	2	14790190	NP_055816	SPEN	spen homolog, transcriptional regulator	2	2	nucleus	Notch signaling pathway regulation of transcription, DNA- dependent	DNA binding nucleic acid binding nucleotide binding RNA binding	

Supplementary Table I

	2	6912682	NP_036577	SPON2	spondin 2, extracellular matrix protein	1	1	extracellular matrix (sensu Metazoa) extraellular space	axon guidance cell adhesion development immune response	actin binding	
	2	4507185	NP_003115	SPR	sepiapterin reductase (7,8-dihydrobiopterin:NADH)	1	2	not classified	electron transport metabolism nitric oxide biosynthesis tetrahydrobiopterin biosynthesis	aldo-keto reductase activity NADP binding oxidoreductase activity sepiapterin reductase activity	Dystonia, Dopa-Responsive, Due To Sepiapterin Reductase Deficiency [MIM:251120]
Yes	2	45827734	NP_005978	SPRR1A	small proline-rich protein 1A	2	7	cornified envelope	epidermis development keratinization keratinocyte differentiation peptide cross-linking	protein binding, bridging structural molecule activity	
Yes	2	62955831	NP_001017418	SPRR2B	small proline-rich protein 2B	1	1	cornified envelope cytoplasm	embryo implantation epidermis development keratinization keratinocyte differentiation menstrual cycle	protein binding structural molecule activity	
Yes	2	83582817	NP_001019380	SPRR2E	small proline-rich protein 2E	1	4	not classified	not classified	not classified	
	2	4885607	NP_005407	SPRR3	small proline-rich protein 3	5	6	cornified envelope	epidermis development keratinization keratinocyte differentiation wound healing	protein binding structural molecule activity	
	2	4507195	NP_003119	SPTBN1	spectrin, beta, non-erythrocytic 1 isoform 1	1	1	cytoskeleton membrane spectrin	barbed-end actin filament capping	actin binding calmodulin binding structural constituent of cytoskeleton	
	2	112382252	NP_842565	SPTBN1	spectrin, beta, non-erythrocytic 1 isoform 2	1	1	cytoskeleton membrane spectrin	barbed-end actin filament capping	calmodulin binding structural constituent of cytoskeleton calcium channel regulator activity	
	2	4505571	NP_003891	SQSTM1	sequestosome 1	1	2	cytosol nucleus	apoptosis cell differentiation endosome transport immune response intracellular signaling cascade positive regulation of transcription from RNA polymerase II promoter protein localization regulation of I-kappaB kinase/NF-kappaB cascade response to stress	metal ion binding protein binding protein kinase binding SH2 domain binding ubiquitin binding zinc ion binding	Paget Disease of Bone [MIM:602080]
	2	4885609	NP_005408	SRC	proto-oncogene tyrosine-protein kinase SRC	1	1	integral to membrane membrane	protein amino acid phosphorylation protein kinase cascade signal complex formation	ATP binding nucleotide binding protein binding protein-tyrosine kinase activity SH2 domain binding SH3/SH2 adaptor activity transferase activity	

Supplementary Table I

	1,2	4507207	NP_003121	SRI	sorcini isoform a	4	10	cytoplasm	heart development intracellular sequestering of iron ion muscle development regulation of action potential regulation of heart contraction regulation of striated muscle contraction signal transduction transport	calcium channel regulator activity calcium ion binding receptor binding	
	2	38679884	NP_944490	SRI	sorcini isoform b	3	8	cytoplasm	heart development intracellular sequestering of iron ion muscle development regulation of action potential regulation of heart contraction regulation of striated muscle contraction signal transduction transport	calcium ion binding receptor binding binding	
	2	19923193	NP_003923	ST13	heat shock 70kD protein binding protein	2	3	cytoplasm	protein folding	protein binding, bridging metal ion binding	
	2	47132534	NP_998787	STAMPB	STAM binding protein	2	2	membrane nucleus	JAK-STAT cascade positive regulation of cell proliferation ubiquitin cycle	metallopeptidase activity protein binding ubiquitin thiolesterase activity zinc ion binding not classified	
	2	4507271	NP_000446	STK11	serine/threonine protein kinase 11	1	1	not classified	protein amino acid phosphorylation	ATP binding nucleotide binding protein binding protein serine/threonine kinase activity transferase activity	
	2	73808092	NP_001027467	STK24	serine/threonine kinase 24 (STE20 homolog)	5	12	not classified	protein amino acid phosphorylation signal transduction	ATP binding kinase activity nucleotide binding protein serine/threonine kinase activity protein-tyrosine kinase activity transferase activity	
Yes	2	21361358	NP_006365	STK25	serine/threonine kinase 25	1	2	not classified	protein amino acid phosphorylation response to oxidative stress signal transduction	ATP binding identical protein binding magnesium ion binding nucleotide binding protein serine/threonine kinase activity transferase activity	
	1,2	38016911	NP_004090	STOM	stomatin isoform a	13	76	cytoskeleton integral to plasma membrane membrane	not classified	not classified	
	2	4759184	NP_004168	STX3	syntaxin 3A	2	5	integral to membrane	intracellular protein transport neurotransmitter transport	protein transporter activity	
	2	73760415	NP_001027392	STXBP1	syntaxin binding protein 1 isoform b	1	1	not classified	protein transport vesicle docking during exocytosis vesicle-mediated transport	calcium ion binding	

Supplementary Table 1

	2	5902128	NP_008880	STXBP2	syntaxin binding protein 2	10	22	not classified	intracellular protein transport vesicle docking during exocytosis vesicle-mediated transport	syntaxin-2 binding syntaxin-3 binding	
	2	63999048	NP_848604	STXBP4	syntaxin binding protein 4	1	1	not classified	not classified	protein binding ATP binding	
	2	11321583	NP_003841	SUCLA2	succinate-CoA ligase, ADP-forming, beta subun	1	1	mitochondrion	metabolism succinyl-CoA metabolism succinyl-CoA pathway tricarboxylic acid cycle	ligase activity succinate-CoA ligase (ADP-forming) activity G-protein coupled receptor activity, unknown ligand	Mitochondrial DNA Depletion Syndrome [MIM: 609560]
	2	47271392	NP_149039	SUCNR1	succinate receptor 1	1	1	integral to membrane membrane	G-protein coupled receptor protein signaling pathway signal transduction	purinergic nucleotide receptor activity, G- protein coupled receptor activity rhodopsin-like receptor activity not classified	
	2	10092665	NP_062547	SUSD2	sushi domain containing 2	5	10	not classified	not classified	not classified	
	2	33239443	NP_874384	SYPL1	synaptophysin-like 1 isoform b	1	12	integral to plasma membrane membrane synaptic vesicle	synaptic transmission transport	transporter activity	
	2	4505059	NP_002345	TACSTD1	tumor-associated calcium signal transducer	1	1	integral to membrane plasma membrane	not classified	ATP binding	
	2	4507357	NP_003555	TAGLN2	transgelin 2	4	6	nuclear membrane plasma membrane	muscle development	not classified	
	2	45439370	NP_065842	TAOK1	TAO kinase 1	1	1	not classified	protein amino acid phosphorylat	ATP binding nucleotide binding protein serine/threonine kinase activity transferase activity	
Yes	2	19923464	NP_057365	TAOK3	TAO kinase 3	1	1	not classified	MAPKKK cascade negative regulation of JNK cascade positive regulation of JNK cascade protein amino acid autophosphorylation	ATP binding nucleotide binding protein serine/threonine kinase activity transferase activity	
	2	11993943	NP_055419	TAX1BP3	Tax1 (human T-cell leukemia virus type I) bi	1	1	cytoplasm integral to membrane membrane nucleus	intracellular signaling cascade ion transport negative regulation of Wnt receptor signaling pathway Rho protein signal transduction Wnt receptor signaling pathway	ATP binding ion channel activity protein binding receptor activity	
	2	13994322	NP_114143	TBC1D10A	TBC1 domain family, member 10A	4	5	microvillus	not classified	PDZ domain binding protein binding	
	2	23503283	NP_699187	TBC1D21	TBC1 domain family, member 21	1	1	not classified	not classified	GTPase activator activity	
	2	57863257	NP_110379	TCP1	T-complex protein 1 isoform a	1	1	cytosol	protein folding tubulin folding	ATP binding nucleotide binding unfolded protein binding	
	2	57863259	NP_00100889	TCP1	T-complex protein 1 isoform b	2	4	cytosol	protein folding tubulin folding	nucleotide binding unfolded protein binding structural molecule activity	
	2	4885627	NP_005413	TECTA	tectorin alpha precursor	1	1	extracellular matrix (sensu Metazoa) extracellular space membrane	cell-matrix adhesion sensory perception of sound	extracellular matrix structural constituent GPI anchor binding	Autosomal Dominant nonsyndromic sensorineural hearing loss [MIM:601842]

Supplementary Table I

	2	13994250	NP_114104	TEKT3	tektin 3	1	1	microtubule	microtubule cytoskeleton organization and biogenesis	glycosaminoglycan binding	
	1,2	4557871	NP_001054	TF	transferrin	12	20	endocytic vesicle extracellular region	ion transport iron ion homeostasis iron ion transport	ferric iron binding metal ion binding	Alzheimer Disease [MIM:104300]
	2	56682966	NP_003234	TGFB3	transforming growth factor, beta receptor III	1	1	integral to membrane membrane	development signal transduction transforming growth factor beta receptor signaling pathway	receptor activity acyltransferase activity	
	2	39777599	NP_945189	TGM2	transglutaminase 2 isoform b	1	1	cytosol extracellular matrix (sensu Metazoa) membrane	G-protein coupled receptor protein signaling pathway peptide cross-linking positive regulation of cell adhesion	calcium ion binding GTP binding protein-glutamine gamma-glutamyltransferase activity transferase activity nucleotide binding	
	2	4507479	NP_003232	TGM4	transglutaminase 4 (prostate)	5	6	not classified	peptide cross-linking protein amino acid polyamination	acyltransferase activity calcium ion binding protein-glutamine gamma-glutamyltransferase activity transferase activity	
	1,2	40317626	NP_003237	THBS1	thrombospondin 1 precursor	9	14	extracellular region	blood coagulation cell adhesion cell motility development nervous system development	calcium ion binding endopeptidase inhibitor activity heparin binding protein binding signal transducer activity structural molecule activity	
	2	55770864	NP_005773	THOC4	THO complex 4	1	1	nucleus spliceosome complex	mRNA export from nucleus nuclear mRNA splicing, via spliceosome protein folding transport	protein binding RNA binding not classified	
	1,2	19923362	NP_006279	THY1	Thy-1 cell surface antigen	3	20	cytosol endoplasmic reticulum external side of plasma membrane growth cone integral to plasma membrane lipid raft plasma membrane	angiogenesis cell-cell adhesion cytoskeleton organization and biogenesis focal adhesion formation mast cell activation negative regulation of apoptosis negative regulation of axonogenesis negative regulation of cell migration negative regulation of fibroblast	GPI anchor binding integrin binding protein binding Rho GTPase activator activity	
	2	58331189	NP_00101092	TIAM2	T-cell lymphoma invasion and metastasis 2	1	2	intracellular	intracellular signaling cascade regulation of Rho protein signal transduction	protein binding receptor signaling protein activity Rho guanyl-nucleotide exchange factor activity	
	2	4507521	NP_001055	TKT	transketolase	1	1	not classified	not classified	calcium ion binding transferase activity transketolase activity	
	2	16753233	NP_006280	TLN1	talín 1	4	6	cytoskeleton cytosol focal adhesion intercellular junction ruffle	cell motility cytoskeletal anchoring intercellular junction assembly	actin binding binding LIM domain binding protein binding structural constituent of cytoskeleton vinculin binding	

Supplementary Table I

	1,2	7706575	NP_057635	TM7SF3	transmembrane 7 superfamily member 3	3	8	integral to membrane membrane	not classified	not classified	
	1,2	50593008	NP_071435	TMBIM1	transmembrane BAX inhibitor motif containi	1	7	integral to membrane membrane	not classified	not classified	
	2	22748685	NP_689524	TMEM12	transmembrane protein 12	1	1	integral to membrane	not classified	not classified	
	2	70778895	NP_00102052	TMEM16F	transmembrane protein 16F	6	10	integral to membrane membrane	not classified	not classified	
	2	7019555	NP_037522	TMEM2	transmembrane protein 2	1	1	integral to membrane	not classified	not classified	
	2	15042969	NP_065716	TMEM27	transmembrane protein 27	2	3	integral to membrane membrane	not classified	not classified	
	2	13129060	NP_076979	TMEM38A	transmembrane protein 38A	1	1	integral to membrane	not classified	serine-type endopeptidase inhibitor activity	
	2	20070191	NP_004863	TMEM59	thymic dendritic cell-derived factor 1	1	1	integral to membrane	not classified	peptidase activity	
	2	7662308	NP_055513	TMEM63A	transmembrane protein 63A	2	3	integral to membrane membrane	not classified	nucleotide binding	
	2	4758508	NP_004253	TMPRSS11D	transmembrane protease, serine 11D	1	1	extracellular region integral to plasma membrane membrane	proteolysis respiratory gaseous exchange	serine-type endopeptidase activity not classified	
	1,2	14602459	NP_005647	TMPRSS2	transmembrane protease, serine 2	2	2	integral to plasma membrane membrane	proteolysis	scavenger receptor activity serine-type endopeptidase activity	
	2	21361619	NP_061882	TOLLIP	toll interacting protein	1	1	cytoplasm interleukin-1 receptor complex interleukin-18 receptor complex	cell-cell signaling inflammatory response intracellular signaling cascade leukocyte activation phosphorylation	protein binding signal transducer activity Toll binding	
	2	4885637	NP_005479	TOM1	target of myb1	1	1	cytosol early endosome endosome intracellular lysosome membrane	endocytosis endosome transport intracellular protein transport	protein binding	
	2	75677323	NP_653279	TOM1L2	target of myb1-like 2 isoform 2	1	1	intracellular	intracellular protein transport	not classified	
	2	75677327	NP_00102872	TOM1L2	target of myb1-like 2 isoform 1	1	3	intracellular	intracellular protein transport	protein binding	
	1,2	4507645	NP_000356	TP11	triosephosphate isomerase 1	7	16	not classified	fatty acid biosynthesis gluconeogenesis glycolysis metabolism pentose-phosphate shunt	isomerase activity triose-phosphate isomerase activity	
	1,2	5729770	NP_000382	TPP1	tripeptidyl-peptidase I preproprotein	8	42	lysosome	lipid metabolism nervous system development proteolysis	peptidase activity serine-type endopeptidase activity tripeptidyl-peptidase I activity	Ceroid Lipofuscinosis Neuronal 2 [MIM:204500]
	2	6005914	NP_009111	TREH	trehalase	4	7	anchored to plasma membrane plasma membrane	metabolism trehalose catabolism trehalose metabolism	alpha, alpha-trehalase activity GPI anchor binding hydrolase activity, acting on glycosyl bonds	
	2	11342676	NP_004231	TRIP10	thyroid hormone receptor interactor 10	1	1	cytoplasm	actin cytoskeleton organization and biogenesis signal transduction	alpha-actinin binding	
Yes	2	17505200	NP_062815	TRPV5	transient receptor potential cation channel, s	1	1	integral to plasma membrane membrane	calcium ion homeostasis calcium ion transport ion transport protein tetramerization	calcium channel activity calcium ion binding ion channel activity protein binding	

Supplementary Table I

Yes	2	21314682	NP_061116	TRPV6	transient receptor potential cation channel, s	1	3	integral to plasma membrane	calcium ion transport ion transport regulation of calcium ion-dependent exocytosis	calcium channel activity calcium ion binding calmodulin binding DNA binding ion channel activity	
	1,2	5454140	NP_006283	TSG101	tumor susceptibility gene 101	17	66	membrane	protein transport regulation of cell growth ubiquitin cycle	DNA binding protein binding transcription corepressor activity ubiquitin-protein ligase activity	Breast Cancer [MIM:176960]
	1,2	21264578	NP_005718	TSPAN1	tetraspan 1	1	36	integral to membrane membrane	cell adhesion cell motility cell proliferation	not classified	
	2	38787991	NP_944492	TSPAN3	transmembrane 4 superfamily member 8 iso	1	1	integral to membrane membrane	cell motility cell proliferation	not classified	
	2	4507541	NP_003261	TSPAN6	transmembrane 4 superfamily member 6	1	1	integral to membrane membrane	cell motility positive regulation of I-kappaB kinase/NF-kappaB cascade	signal transducer activity	
	2	74275350	NP_000705	TSPO	translocator protein (18kDa) isoform PBR	1	2	integral to membrane membrane mitochondrial outer membrane mitochondrion	protein targeting to mitochondrion signal transduction	benzodiazepine receptor activity receptor activity	
	2	4507709	NP_003304	TSTA3	tissue specific transplantation antigen P35B	1	1	cytoplasm	de novo GDP-L-fucose biosynthesis cellular metabolism leukocyte adhesion nucleotide-sugar metabolism protein amino acid glycosylation	coenzyme binding electron carrier activity GDP-4-dehydro-D-rhamnose reductase activity GDP-L-fucose synthase activity isomerase activity	
	2	40255087	NP_660153	TTC18	tetratricopeptide repeat domain 18	1	1	not classified	not classified	binding	
	2	20143918	NP_597676	TTN	titin isoform novex-1	4	5	condensed nuclear chromosome integral to membrane membrane myofibril nucleus Z disc	carbohydrate metabolism mitosis muscle contraction myofibril assembly protein amino acid autophosphorylation protein amino acid phosphorylation response to oxidative stress	alpha-actinin binding ATP binding calmodulin binding identical protein binding myosin binding nucleotide binding peroxidase activity protein serine/threonine kinase activity protein-tyrosine kinase activity structural constituent of cell wall transferase	Cardiomyopathy [MIM:188840]
	2	110349719	NP_596869	TTN	titin isoform N2-A	1	1	condensed nuclear chromosome integral to membrane membrane myofibril nucleus Z disc	carbohydrate metabolism mitosis muscle contraction myofibril assembly protein amino acid autophosphorylation protein amino acid phosphorylation response to oxidative stress	ATP binding calmodulin binding identical protein binding myosin binding nucleotide binding peroxidase activity protein serine/threonine kinase activity protein-tyrosine kinase activity structural constituent of cell wall transferase activity GTP binding	
	1,2	51100978	NP_079526	TTYH3	tweety 3	1	7	not classified	not classified	not classified	
Yes	2	17921989	NP_005991	TUBA1	tubulin, alpha 1	2	7	microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding protein binding structural molecule activity	

Supplementary Table I

Yes	2	57013276	NP_006073	TUBA1B	tubulin, alpha, ubiquitous	3	19	microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural molecule activity
	2	17986283	NP_006000	TUBA3	tubulin, alpha 3	15	41	microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural molecule activity
Yes	2	156564363	NP_525125	TUBA3D	tubulin, alpha 3d	13	35	microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity structural molecule activity
Yes	2	14389309	NP_116093	TUBA6	tubulin alpha 6	2	2	microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural molecule activity
Yes	2	9507215	NP_061816	TUBA8	tubulin, alpha 8	2	10	cytoskeleton microtubule protein complex	microtubule-based movement	GTP binding GTPase activity nucleotide binding structural molecule activity
Yes	2	29788785	NP_821133	TUBB	tubulin, beta polypeptide	3	12	cytoplasm microtubule protein complex	microtubule-based movement natural killer cell mediated cytotoxicity protein polymerization	GTP binding GTPase activity MHC class I protein binding nucleotide binding structural molecule activity
	1,2	4507729	NP_001060	TUBB2A	tubulin, beta 2	7	25	cytoplasm cytoskeleton microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural constituent of cytoskeleton
Yes	2	29788768	NP_821080	TUBB2B	tubulin, beta 2B	1	3	cytoplasm microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural molecule activity
Yes	2	5174735	NP_006079	TUBB2C	tubulin, beta, 2	3	9	cytoplasm microtubule protein complex	microtubule-based movement natural killer cell mediated cytotoxicity protein polymerization	GTP binding GTPase activity MHC class I protein binding nucleotide binding structural molecule activity unfolded protein binding
Yes	2	50592996	NP_006077	TUBB3	tubulin, beta, 4	3	7	cytoplasm cytoskeleton integral to membrane membrane microtubule protein complex	G-protein coupled receptor protein signaling pathway microtubule-based movement protein polymerization signal transduction	GTP binding GTPase activity melanocyte stimulating hormone receptor activity nucleotide binding receptor activity rhodopsin-like receptor activity structural constituent of cytoskeleton structural molecule activity
Yes	2	55770868	NP_064424	TUBB4Q	tubulin, beta polypeptide 4, member Q	1	4	cytoskeleton microtubule protein complex	microtubule-based movement protein polymerization	GTP binding GTPase activity nucleotide binding structural molecule activity
	2	14210536	NP_115914	TUBB6	tubulin, beta 6	1	1	cytoplasm microtubule protein complex	microtubule-based movement protein polymerization	GTPase activity nucleotide binding structural molecule activity GTP binding

Supplementary Table I

	2	42558279	NP_817124	TUBB8	tubulin, beta 8	1	1	cytoplasm microtubule protein complex	microtubule-based movement protein polymerization	GTPase activity nucleotide binding structural molecule activity GTP binding not classified	
	2	50592994	NP_003320	TXN	thioredoxin	2	8	not classified	cell motility cell proliferation cell-cell signaling electron transport signal transduction transport	electron carrier activity protein binding protein disulfide oxidoreductase activity thiol-disulfide exchange intermediate activity	
	2	51510891	NP_001003936	TXNDC8	thioredoxin domain containing 8	1	1	not classified	cell differentiation spermatogenesis	electron carrier activity protein disulfide oxidoreductase activity	
	2	14249348	NP_116120	TXNL5	thioredoxin-like 5	2	2	not classified	not classified	electron carrier activity protein disulfide oxidoreductase activity	
	2	59850762	NP_060473	UACA	uveal autoantigen with coiled-coil domains 2	1	1	cytoskeleton extracellular region nucleus	viral life cycle	ATP binding	
	2	23510340	NP_695012	UBA1	ubiquitin-activating enzyme E1	4	9	not classified	DNA replication ubiquitin cycle	catalytic activity ligase activity nucleotide binding protein binding ubiquitin activating enzyme activity ligase activity	
Yes	2	77539055	NP_001029102	UBA52	ubiquitin and ribosomal protein L40 precursor	5	48	intracellular nucleus ribosome	axon guidance cell cycle ER-associated protein catabolism long-term strengthening of neuromuscular junction positive regulation of transcription protein modification protein ubiquitination regulation of synaptic plasticity	protein binding structural constituent of ribosome transcription regulator activity	
	2	55770884	NP_057256	UBADC1	ubiquitin associated domain containing 1	1	1	not classified	not classified	not classified	
Yes	1,2	67191208	NP_066289	UBC	ubiquitin C	7	198	not classified	protein modification	not classified	
	2	23510338	NP_003325	UBE1	ubiquitin-activating enzyme E1	2	4	not classified	DNA replication ubiquitin cycle	ATP binding catalytic activity ligase activity nucleotide binding protein binding ubiquitin activating enzyme activity	
Yes	2	33188456	NP_862821	UBE2D2	ubiquitin-conjugating enzyme E2D 2 isoform	1	1	not classified	ubiquitin cycle	ligase activity protein binding ubiquitin-protein ligase activity	
	2	13489085	NP_003333	UBE2G1	ubiquitin-conjugating enzyme E2G 1	1	1	not classified	ubiquitin cycle	ubiquitin-protein ligase activity not classified	
Yes	1,2	4507793	NP_003339	UBE2N	ubiquitin-conjugating enzyme E2N	1	4	cytoplasm nucleus UBC13-MMS2 complex	activation of NF-kappaB transcription factor positive regulation of DNA repair positive regulation of I-kappaB kinase/NF-kappaB cascade protein ubiquitination	ligase activity protein binding ubiquitin binding ubiquitin conjugating enzyme activity ubiquitin-protein ligase activity	

Supplementary Table I

									cytoplasm nucleus ubiquitin conjugating enzyme complex	activation of NF-kappaB transcription factor cell differentiation positive regulation of I-kappaB kinase/NF-kappaB cascade protein polyubiquitination regulation of DNA repair regulation of progression through cell cycle regulation of transcription, DNA- de	protein binding transcriptional activator activity ubiquitin conjugating enzyme activity ubiquitin-protein ligase activity	
Yes	2	40806167	NP_954595	UBE2V1	ubiquitin-conjugating enzyme E2 variant 1 is	2	4					
Yes	2	73765546	NP_001027459	UBE2V1	ubiquitin-conjugating enzyme E2 variant 1 is	2	5	cytoplasm nucleus ubiquitin conjugation enzyme complex	activation of NF-kappaB transcription factor cell differentiation positive regulation of I-kappaB kinase/NF-kappaB cascade protein polyubiquitination regulation of DNA repair regulation of progression through cell cycle regulation of transcription, DNA- dependent ubiquitin cycle	protein binding transcriptional activator activity ubiquitin conjugating enzyme activity ubiquitin-protein ligase activity		
	2	6005928	NP_009037	UBL3	ubiquitin-like 3	2	8	intracellular membrane	protein modification	L-lactate dehydrogenase activity		
Yes	2	109659841	NP_689490	UBTD2	dendritic cell-derived ubiquitin-like protein	1	1	not classified	protein modification	not classified		
	2	13376854	NP_079517	UBXD1	UBX domain containing 1	1	1	not classified	not classified	not classified		
	1,2	23943814	NP_060784	UEVLD	ubiquitin-conjugating enzyme E2-like isoform	11	42	not classified	glycolysis protein transport tricarboxylic acid cycle intermediate metabolism ubiquitin cycle	L-lactate dehydrogenase activity oxidoreductase activity ubiquitin-protein ligase activity		
	2	103472011	NP_00103578	UEVLD	ubiquitin-conjugating enzyme E2-like isoform	4	10	not classified	glycolysis protein transport tricarboxylic acid cycle intermediate metabolism ubiquitin cycle	oxidoreductase activity ubiquitin-protein ligase activity not classified		
	2	156416000	NP_057490	UFC1	ubiquitin-fold modifier conjugating enzyme 1	1	1	not classified	ubiquitin cycle	kinase activity		
	2	48255966	NP_006750	UGP2	UDP-glucose pyrophosphorylase 2 isoform	2	2	not classified	metabolism UDP-glucose metabolism	kinase activity nucleotidyltransferase activity transferase activity UTP:glucose-1-phosphate uridylyltransferase activity		
	2	48255968	NP_00100152	UGP2	UDP-glucose pyrophosphorylase 2 isoform	3	4	not classified	metabolism UDP-glucose metabolism	nucleotidyltransferase activity transferase activity UTP: glucose-1-phosphate uridylyltransfease activity ATP binding		
	1,2	39930361	NP_056333	ULK3	unc-51-like kinase 3	4	12	not classified	protein amino acid phosphorylat	ATP binding nucleotide binding protein serine/threonine kinase activity protein-tyrosine kinase activity transferase activity		

Supplementary Table I

	1,2	59850812	NP_003352	UMOD	uromodulin precursor	35	1278	extracellular space extrinsic to membrane membrane	cellular defense response negative regulation of cell proliferation	calcium ion binding GPI anchor binding	Medullary cystic kidney disease-2 (MCKD2) [MIM:603860] Familial juvenile hyperuricemic nephropathy (FJHN) [MIM:16200]
	2	7706509	NP_057411	UPB1	beta-ureidopropionase	1	1	not classified	nitrogen compound metabolism	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds metal ion binding zinc ion binding protein binding	
	1,2	5902148	NP_008931	UPK1A	uroplakin 1A	5	8	endoplasmic reticulum integral to membrane plasma membrane	endocytosis epithelial cell differentiation protein oligomerization signal transduction	monosaccharide binding protein binding protein homodimerization activity	
	2	49619237	NP_008883	UPK1B	uroplakin 1B	1	2	integral to membrane membrane	epithelial cell differentiation physiological process	structural molecule activity	
	1,2	5803215	NP_006751	UPK2	uroplakin 2	4	27	endoplasmic reticulum integral to endoplasmic reticulum membrane integral to plasma membrane membrane	development epithelial cell differentiation membrane organization and biogenesis	not classified	
	2	5902152	NP_008884	UPK3A	uroplakin 3A precursor	1	1	endoplasmic reticulum integral to membrane membrane membrane fraction	epithelial cell differentiation	not classified	
Yes	2	32698815	NP_872294	USP12	ubiquitin-specific protease 12-like 1	1	1	not classified	ubiquitin-dependent protein cata	cysteine-type endopeptidase activity ubiquitin thiolesterase activity	
Yes	2	148727247	NP_003472	USP5	ubiquitin specific peptidase 5 isoform 2	1	1	lysosome proteasome complex (sensu Eukaryota)	ubiquitin-dependent protein cata	cysteine-type endopeptidase activity metal ion binding protein binding ubiquitin-specific protease 5 activity zinc ion binding	
	2	6005938	NP_009055	UTRN	utrophin	1	1	cytoskeleton membrane fraction plasma membrane	muscle contraction muscle development	actin binding calcium ion binding protein binding structural molecule activity zinc ion binding	
Yes	2	7657675	NP_055047	VAMP2	vesicle-associated membrane protein 2 (syn	1	1	integral to membrane membrane synapse synaptosome	vesicle-mediated transport	not classified	
	2	5032137	NP_005629	VAMP7	vesicle-associated membrane protein 7	1	1	endoplasmic reticulum integral to membrane membrane	protein transport vesicle docking during exocytosis vesicle-mediated transport	actin binding	Beta-Ureidopropionase Deficiency [MIM: 606673]
	2	14043026	NP_003752	VAMP8	vesicle-associated membrane protein 8	1	1	early endosome integral to membrane membrane membrane fraction synaptic vesicle	protein complex assembly vesicle docking during exocytosis vesicle-mediated transport	not classified	
	2	39930521	NP_612449	VASN	slit-like 2	1	2	integral to membrane membrane	not classified	transferase activity	
	2	4507869	NP_003361	VASP	vasodilator-stimulated phosphoprotein	1	1	actin cytoskeleton	cell motility	actin binding	
	2	18379349	NP_006364	VAT1	vesicle amine transport protein 1	7	11	integral to membrane nucleus synaptic vesicle	cell growth	DNA binding oxidoreductase activity zinc ion binding	

Supplementary Table 1

	2	4507877	NP_003364	VCL	vinculin isoform VCL	4	6	actin cytoskeleton adherens junction cell-matrix junction costamere focal adhesion intercellular junction protein complex	apical junction assembly cell adhesion cell motility lamellipodium biogenesis negative regulation of cell migration	actin binding alpha-catenin binding oxidoreductase activity structural molecule activity	
	2	7669550	NP_054706	VCL	vinculin isoform meta-VCL	3	5	actin cytoskeleton adherens junction cell-matrix junction costamere focal adhesion intercellular junction protein complex	apical junction assembly cell adhesion cell motility lamellipodium biogenesis negative regulation of cell migration	alpha-catenin binding oxidoreductase activity structural molecule activity not classified	Cardiomyopathy, Dilated [MIM: 611407]
	2	6005942	NP_009057	VCP	valosin-containing protein	2	2	cytosol endoplasmic reticulum microsome nucleus	caspase activation double-strand break repair ER-associated protein catabolism establishment of protein localization protein ubiquitination regulation of apoptosis retrograde protein transport, ER to cytosol transport unfolded protein response	ATP binding ATPase activity hydrolase activity lipid binding nucleotide binding protein binding	Inclusion Body Myopathy with Early Onset Paget Disease and Frontotemporal Dementia [MIM:167320]
	2	4507879	NP_003365	VDAC1	voltage-dependent anion channel 1	6	43	integral to membrane membrane mitochondrial outer membrane mitochondrion outer membrane	anion transport apoptotic program	apoptogenic cytochrome c release channel activity voltage-gated anion channel porin activity voltage-gated ion-selective channel activity	
Yes	2	42476281	NP_003366	VDAC2	voltage-dependent anion channel 2	3	4	integral to membrane membrane mitochondrial outer membrane mitochondrion outer membrane	anion transport	voltage-gated anion channel porin activity voltage-gated ion-selective channel activity	
	2	25188179	NP_005653	VDAC3	voltage-dependent anion channel 3	1	1	integral to plasma membrane membrane mitochondrial outer membrane mitochondrion outer membrane	adenine transport anion transport	voltage-gated anion channel porin activity voltage-gated ion-selective channel activity	
	2	6005944	NP_009058	VIL1	villin 1	2	3	cytoskeleton F-actin capping protein complex	actin filament bundle formation actin filament severing barbed-end actin filament capping cytoskeleton organization and biogenesis protein complex assembly	actin binding calcium ion binding	
Yes	1,2	21614499	NP_003370	VIL2	villin 2	43	257	actin filament cortical cytoskeleton cytoplasm cytoskeleton membrane microvillus	actin filament bundle formation cytoskeletal anchoring regulation of cell shape	actin filament binding binding cytoskeletal protein binding structural molecule activity	

Supplementary Table 1

	2	32698964	NP_872372	VMO1	secretory protein LOC284013	4	7	not classified	not classified	not classified	
	2	66348091	NP_060550	VPS13C	vacuolar protein sorting 13C protein isoform	1	1	not classified	protein localization	not classified	
	2	54607139	NP_056193	VPS13D	vacuolar protein sorting 13D isoform 1	1	1	not classified	protein localization	not classified	
	2	7706353	NP_057163	VPS24	vacuolar protein sorting 24 isoform 1	1	3	membrane	protein transport	not classified	
	2	54144646	NP_00100575	VPS24	vacuolar protein sorting 24 isoform 2	1	4	membrane	protein transport	protein binding	
	2	14150155	NP_115729	VPS25	vacuolar protein sorting 25	6	20	nucleus	protein transport regulation of transcription, DNA-dependent transcription	not classified	
	1,2	7705885	NP_057292	VPS28	vacuolar protein sorting 28 isoform 1	7	13	cytosol endosome	protein transport	protein binding	
	2	34452693	NP_898880	VPS28	vacuolar protein sorting 28 isoform 2	2	8	cytosol endosome	protein transport	protease binding	
	2	17999541	NP_060676	VPS35	vacuolar protein sorting 35	3	6	cytosol endosome membrane	protein transport retrograde transport, endosome to Golgi	protein binding	
	2	71051598	NP_057159	VPS36	vacuolar protein sorting 36	5	8	nucleus	protein transport regulation of transcription, DNA-dependent transcription	lipid binding	
	2	13375926	NP_078943	VPS37B	vacuolar protein sorting 37B	5	11	not classified	not classified	not classified	
	2	57863314	NP_060436	VPS37C	vacuolar protein sorting 37C	1	1	not classified	not classified	not classified	
	2	51466316	NP_00107108	VPS37D	vacuolar protein sorting 37D	5	19	not classified	not classified	not classified	
Yes	1,2	7019569	NP_037377	VPS4A	vacuolar protein sorting factor 4A	17	44	cytoplasm membrane perinuclear region vacuole	protein transport vesicle-mediated transport	ATP binding nucleoside-triphosphatase activity nucleotide binding protein binding	
	1,2	17865802	NP_004860	VPS4B	vacuolar protein sorting factor 4B	15	47	endosome membrane vacuole	endosome transport membrane fusion peroxisome organization and biogenesis regulation of transcription, DNA-dependent	ATP binding ATPase activity, coupled nucleotide binding protein binding	
Yes	1,2	18201911	NP_000629	VTN	vitronectin precursor	4	12	extracellular region extracellular space	cell adhesion immune response	heparin binding protein binding	
	2	89191868	NP_000543	VWF	von Willebrand factor preproprotein	1	4	extracellular matrix (sensu Metazoa) extracellular region	cell adhesion platelet activation response to wounding	protein binding actin binding	Von Willebrand Disease [MIM: 193400]
	2	51702526	NP_003932	WASL	Wiskott-Aldrich syndrome gene-like protein	1	1	actin cytoskeleton nucleus	actin polymerization and/or depolymerization cell motility protein complex assembly regulation of transcription, DNA-dependent transcription	actin binding small GTPase regulator activity	
	2	9257257	NP_059830	WDR1	WD repeat-containing protein 1 isoform 1	10	18	cytoskeleton	sensory perception of sound	actin binding protein binding	
	2	53729352	NP_005103	WDR1	WD repeat-containing protein 1 isoform 2	4	9	cytoskeleton	sensory perception of sound	protein binding metal ion binding	
	2	151301215	NP_067064	WIZ	widely-interspaced zinc finger motifs	1	1	intracellular nucleus	not classified	nucleic acid binding zinc ion binding metal ion binding	
Yes	2	12711660	NP_061852	WNK1	WNK lysine deficient protein kinase 1	11	67	cytoplasm	ion transport protein amino acid phosphorylation protein kinase cascade regulation of cellular process	ATP binding nucleotide binding protein binding protein kinase inhibitor activity protein serine/threonine kinase activity transferase activity	

Supplementary Table I

Yes	2	50845416	NP_001002838	WNK3	WNK lysine deficient protein kinase 3 isoform 8	8	17	not classified	protein amino acid phosphorylation protein kinase cascade	ATP binding kinase activity nucleotide binding protein serine/threonine kinase activity transferase activity	
Yes	2	50845418	NP_065973	WNK3	WNK lysine deficient protein kinase 3 isoform 5	5	12	not classified	protein amino acid phosphorylation protein kinase cascade	ATP binding kinase activity nucleotide binding protein serine/threonine kinase activity transferase activity	
Yes	2	34365793	NP_115763	WNK4	WNK lysine deficient protein kinase 4	4	11	tight junction	ion transport protein amino acid phosphorylation protein kinase cascade regulation of cellular process	ATP binding nucleotide binding protein serine/threonine kinase activity transferase activity	
	2	14249180	NP_116031	WNT5B	wingless-type MMTV integration site family, 2	2	2	extracellular region	development frizzled-2 signaling pathway	signal transducer activity	
	1,2	62865899	NP_003390	XPNPEP2	X-prolyl aminopeptidase 2, membrane-bound	33	92	membrane	proteolysis	GPI anchor binding manganese ion binding metal ion binding metalloexopeptidase activity X-Pro aminopeptidase activity	
Yes	2	4885661	NP_005424	YES1	viral oncogene yes-1 homolog 1	6	13	not classified	intracellular signaling cascade protein amino acid phosphorylation	ATP binding nucleotide binding protein binding protein-tyrosine kinase activity transferase activity	
Yes	2	21328448	NP_647539	YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	7	22	cytoplasm perinuclear region	not classified	monooxygenase activity protein domain specific binding	
	1,2	5803225	NP_006752	YWHAE	tyrosine 3/tryptophan 5 -monooxygenase activation protein, epsilon polypeptide	4	13	not classified	not classified	enzyme binding protein domain specific binding	
	1,2	21464101	NP_036611	YWHAG	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide	2	4	cytoplasm	cytoskeleton organization and biogenesis negative regulation of apoptosis negative regulation of protein kinase activity regulation of mitosis regulation of neuron differentiation regulation of signal transduction regulation of synaptic plasticity	actin binding insulin-like growth factor receptor binding protein domain specific binding protein homodimerization activity protein kinase C binding protein kinase C inhibitor activity	

Supplementary Table I

	2	4507951	NP_003396	YWHAH	tyrosine 3/tryptophan 5 -monooxygenase ac	1	1	cytoplasm	cytoskeleton organization and biogenesis glucocorticoid catabolism glucocorticoid receptor signaling pathway intracellular protein transport negative regulation of apoptosis negative regulation of dendrite morphogenesis negative regulation of protein kina	actin binding glucocorticoid receptor binding insulin-like growth factor receptor binding protein domain specific binding protein kinase C inhibitor activity transcriptional activator activity	
	1,2	5803227	NP_006817	YWHAQ	tyrosine 3/tryptophan 5 -monooxygenase activation protein, theta polypeptide	1	2	not classified	exocytosis regulation of progression through cell cycle	protein domain specific binding protein kinase C inhibitor activity	
	1,2	4507953	NP_003397	YWHAZ	tyrosine 3/tryptophan 5 -monooxygenase activation protein, zeta polypeptide	8	28	not classified	not classified	protein domain specific binding transcription factor binding	
Yes	2	21735625	NP_663723	YWHAZ	tyrosine 3/tryptophan 5 -monooxygenase ac	7	29	not classified	not classified	protein domain specific binding transcription factor binding	
	2	57863248	NP_00100988	ZCCHC11	zinc finger, CCHC domain containing 11 iso	1	5	not classified	not classified	metal ion binding nucleic acid binding nucleotidyltransferase activity zinc ion binding	
	2	24307963	NP_037436	ZDHHC1	zinc finger, DHHC domain containing 1	1	2	integral to membrane membrane	not classified	acyltransferase activity DNA binding metal ion binding protein binding transferase activity zinc ion binding	
	2	18379366	NP_005848	ZMPSTE24	zinc metalloproteinase STE24	1	1	endoplasmic reticulum integral to membrane membrane	proteolysis	metalloendopeptidase activity metalloexopeptidase activity zinc ion binding DNA binding	Madibuloacral Dysplasia [MIM: 608612]
	2	23957700	NP_705836	ZNF114	zinc finger protein 114	1	1	intracellular nucleus	regulation of transcription, DNA-dependent	metal ion binding nucleic acid binding zinc ion binding	
	2	150170665	NP_443084	ZNF486	zinc finger protein 486	1	1	intracellular nucleus	regulation of transcription, DNA-dependent transcription	metal ion binding zinc ion binding	