

UNIVERSITY OF CALGARY

Searching for genes involved in vitamin B₁₂ metabolism disorders in humans

by

Xiao Li

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOCHEMISTRY & MOLECULAR BIOLOGY

CALGARY, ALBERTA

AUGUST 2007

© Xiao Li 2007

Abstract

Genetic disorders represented by nine complementation groups (*cblA-H* and *mut*) have blocks in intracellular vitamin B₁₂ utilization in humans. The genes responsible for some complementation groups (*cblD*, *cblF* and *cblH*) remain unknown. In this study, two approaches were developed to identify these genes. In one approach, two cDNA libraries were screened in a search for complementing cDNAs, but no positive pools were found. In a second approach, candidate B₁₂ gene lists were generated by comparing genomes of B₁₂ users and non-B₁₂ users. A list of genes accounting for only 3% of the human genome was generated that was highly enriched for known B₁₂ genes.

Acknowledgements

This thesis would not have been possible without the following people, who I thank heartily.

I thank Dr. Roy Gravel for his supervision, guidance, advice and encouragement.

I am grateful to Drs. Jim McGhee, Floyd Snyder and Derrick Rancourt for being a part of my committee for their advice, suggestions, and support.

I express my gratitude to my colleagues, especially those who shared their time, assistance and advice and have become dear friends to me. These include: Sean Froese, Jun Zhang, Shannon Healy, Marko Vujanovic, Megan McDonald, Xuchu Wu, Yongqin Xu, Weija Dong, and Xue Yang.

I thank Drs. David Rosenblatt and John Callahan for cells lines used in this study.

I thank Drs. Shirin Bonni and Jonathan Lytton for accommodating me as a rotation student in their lab.

I thank Judy Gayford for introducing me into the department and keeping me from missing deadlines.

I am indebted to my girlfriend, Xue Han and my parents, Li Xinggui and Wang Suzhen for their support and understanding.

Financial support was provided from the research funds (CIHR and NIH) granted to Dr. Roy Gravel and from scholarships from the Department of Biochemistry & Molecular Biology.

Dedication

This thesis is dedicated in loving memory of my grandma, who passed away in 2006.

Table of Contents

Approval Page.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Dedication.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures and Illustrations.....	ix
List of Symbols, Abbreviations and Nomenclature.....	x
 CHAPTER ONE: OVERVIEW.....	 1
1.1 Vitamin B ₁₂ background.....	1
1.2 Vitamin B ₁₂ biosynthesis and utilization in prokaryotes.....	2
1.3 Vitamin B ₁₂ utilization and genetic diseases of vitamin B ₁₂ metabolism in humans.....	3
1.3.1 B ₁₂ absorption.....	3
1.3.2 The cellular vitamin B ₁₂ pathway and genetic disorders.....	5
1.3.2.1 Complementation groups with a gene identification.....	6
1.3.2.2 Complementation groups without a gene identification.....	13
 CHAPTER TWO: COMPLEMENTATION-BASED CLONING TO IDENTIFY VITAMIN B ₁₂ GENES.....	 16
2.1 Introduction.....	16
2.2 Materials and methods:.....	19
2.2.1 Cell lines.....	19
2.2.2 Tissue culture.....	19
2.2.3 Preparation of polyA ⁺ RNA.....	19
2.2.4 Gene delivery of human fibroblasts.....	19
2.2.4.1 Microinjection.....	19
2.2.4.2 Nucleofection.....	20
2.2.5 [¹⁴ C]--Propionate incorporation in human cells.....	20
2.2.6 Autoradiography.....	21
2.2.7 cDNA library screening.....	22
2.2.7.1 cDNA library from Clontech.....	22
2.2.7.2 cDNA library from OriGene.....	23
2.3 Results.....	24
2.3.1 Selection of culture conditions and tester mutant cell lines.....	24
2.3.2 Complementation by polyA ⁺ RNA from HepG2 cells.....	25
2.3.3 Screening the Clontech library.....	25
2.3.3.1 Test of the system.....	25
2.3.3.2 Library screening.....	26
2.3.4 Screening the Origene Master Plate.....	26
2.3.4.1 Test of the system.....	26
2.3.4.2 Library screening.....	27

2.4 Discussion	27
2.5 Future directions	32
CHAPTER THREE: IDENTIFICATION OF GENE CANDIDATES BY A	
COMPARATIVE GENOMICS APPROACH	39
3.1 Introduction.....	39
3.2 Methods	42
3.2.1 BLASTP	42
3.2.2 Comparison of selected eukaryotic genomes	42
3.2.2.1 Generate candidate gene list using HomoloGene	42
3.2.2.2 Generate candidate gene list using InParanoid	43
3.3 Results and discussion	46
3.3.1 Survey of vitamin B ₁₂ utilization in eukaryotic genomes	46
3.3.2 Comparison of selected eukaryotic genomes	57
3.4 Future directions	61
3.5 Appendix.....	70
3.5.1 Table for gene list generated using HomoloGene	70
3.5.2 Table for gene list generated using InParanoid	147
REFERENCES	176

List of Tables

Table 2.1 Puck's saline	21
Box 3.1 The Boolean query used in HomoloGene WWW site	43
Table 3.1 Survey of B ₁₂ genes, <i>TC1</i> , <i>TC2</i> and <i>IF</i> across eukaryotic genomes	55
Table 3.2 The 50 genes shared by the two lists	64
Table A.1 Gene list generated using HomoloGene	70
Table A.2. Gene list generated using InParanoid	147

List of Figures and Illustrations

Figure 1.1 Structure of cyanocobalmin.....	2
Figure 1.2 Vitamin B ₁₂ Absorption (Adapted from (18)).....	4
Figure 1.3 The intracellular vitamin B ₁₂ pathway	6
Figure 2.1 Principle of [¹⁴ C]-propionate incorporation to monitor B ₁₂ metabolism.....	18
Figure 2.2 Autoradiography of TCA-precipitable protein after [¹⁴ C]-propionate incorporation	34
Figure 2.3 Microinjection of HepG2 mRNA.....	35
Figure 2.4 A cblB clone isolated from the library: pEXp1_MMAB	36
Figure 2.5 Screening the Clontech library	37
Figure 2.6 Test of the cDNA clones from OriGene.....	38
Figure 3.1 Example clusters of the data set for Homo sapiens (Ensembl) and Caenorhabditis elegans (Wormbase)	45
Figure 3.2 Synthesis and utilization of Vitamin B ₁₂	55
Figure 3.3 Comparison based on HomoloGene.....	58
Figure 3.4 Comparison based on InParanoid.....	60

List of Symbols, Abbreviations and Nomenclature

Symbol/ Abbreviations/ Nomenclature	Definition
B ₁₂	Vitamin B ₁₂
Cbl	Any form of cellular cobalamin
AdoCbl	5'-deoxyadenosylcobalamin
MeCbl	methylcobalamin
MS	Methionine synthase enzyme (5-Methyltetrahydrofolate-homocysteine methyltransferase)
MCM	Methylmalonyl-CoA mutase enzyme
<i>MTR</i>	Methionine synthase gene
<i>MTRR</i>	Methionine synthase reductase gene
<i>MUT</i>	Methylmalonyl-CoA mutase gene
<i>MMAA</i>	gene for complementation group <i>cbIA</i>
<i>MMAB</i>	gene for complementation group <i>cbIB</i> , encodes the ATP:cobalamin adenosyltransferase
<i>MMACHC</i>	gene for complementation group <i>cbIC</i>
<i>cbIA</i>	complementation group <i>cbIA</i> , methylmalonic aciduria type A, defective in <i>MMAA</i>
<i>cbIB</i>	complementation group <i>cbIB</i> , methylmalonic aciduria type B, defective in <i>MMAB</i>
<i>cbIC</i>	complementation group <i>cbIC</i> , methylmalonic aciduria and homocystinuria, <i>cbIC</i> type, defective in <i>MMACHC</i>
<i>cbID</i>	complementation group <i>cbID</i> , methylmalonic aciduria and homocystinuria, <i>cbID</i> type
<i>cbIE</i>	complementation group <i>cbIE</i> ,
<i>cbIG</i>	complementation group <i>cbIG</i> , Homocystinuria-megaloblastic anemia, defective in methionine synthase
<i>cbIF</i>	complementation group <i>cbIF</i> , methylmalonic aciduria and homocystinuria, <i>cbIF</i> type
<i>cbIH</i>	complementation group <i>cbIH</i> , methylmalonic aciduria, <i>cbIH</i> type
<i>mut</i>	complementation group <i>mut</i> , methylmalonic aciduria, defective in methylmalonyl-CoA mutase
TCI	Transcobalamin I, a.k.a. haptocorrin or R-binder
TCII	Transcobalamin II, a.k.a. transcobalamin
HC	Haptocorrin, a.k.a. transcobalamin I

Chapter One: Overview

1.1 Vitamin B₁₂ background

Vitamin B₁₂ (cobalamin, cbl) is an essential human nutrient. It is required in humans as the precursor for cofactors of two enzymes: methionine synthase (MS) and methylmalonyl-CoA mutase (MCM). The active cofactor form for MS is methylcobalamin (MeCbl) and is 5'-deoxyadenosylcobalamin (AdoCbl) for MCM. The deficiency in humans can be caused by insufficient supply from food (e.g. as for strict vegetarians), malabsorption (e.g. damage of intestine), or blocks in cellular modification and utilization.

B₁₂ was first recognized as a factor in beef liver as treatment for pernicious anemia(1). The purification was achieved by two groups in 1948 (Lester Smith, Glaxo Laboratories, UK and Karl Folkers, Merck Laboratories) (2,3). The term “vitamin B₁₂” refers to a group of cobalamins, such as cyanocobalamin, adenosylcobalamin, methylcobalamin, etc. Cyanocobalamin is the most stable and commonly available form. The crystal structures of cyanocobalamin and adenosylcobalamin were determined (4,5) (figure 1 Cyanocobalamin). Vitamin B₁₂ has a planar corrin ring, within which a cobalt atom is chelated via four of its six covalent bonds. A 5, 6-dimethylbenzimidazole phosphoribosyl moiety extends down from the D-ring of the corrin ring and loops back to the planar ring via the α -axial position to form a nitrogen-cobalt bond. The upper β -axial ligand can be a cyano group, an adenosyl group, a methyl group or a hydroxyl group, etc.

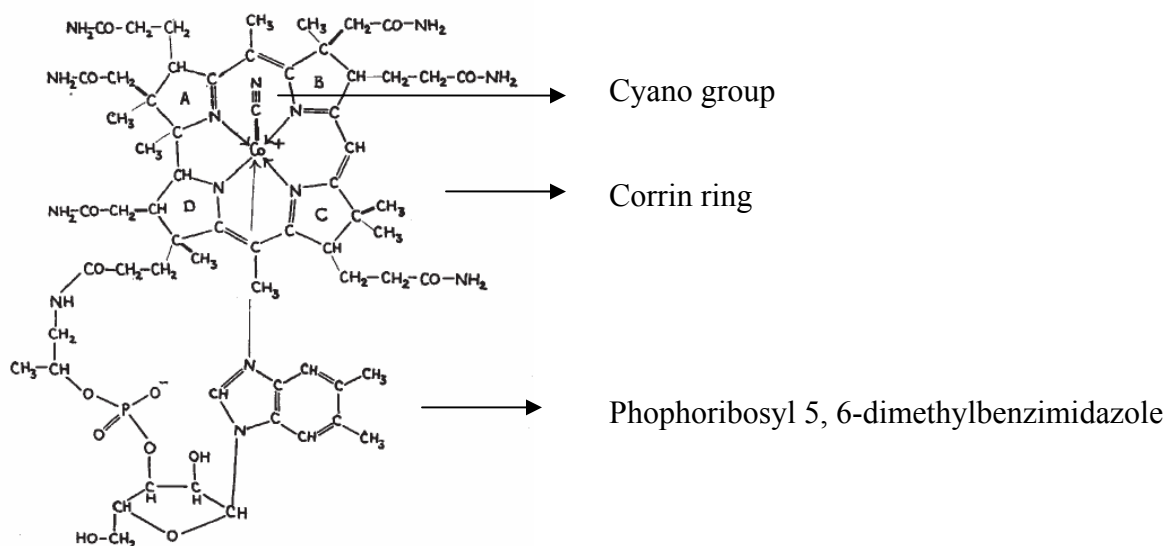


Figure 1.1 Structure of cyanocobalamin

(Adapted from Hodgkin *et al* 1954) The cyano group can be replaced by an adenosyl group, a methyl group or a hydroxyl group, etc.

1.2 Vitamin B₁₂ biosynthesis and utilization in prokaryotes

The biosynthesis (6) and utilization (7) of vitamin B₁₂ in prokaryotes were immediately recognized as it was isolated. In fact, the requirement of vitamin B₁₂ for bacterial growth as the basis for the assay for the vitamin during fractionation speeded up its isolation. It is now known that vitamin B₁₂ production and utilization are present in select archae and eubacteria. It appears that B₁₂ synthesis did not reach eukaryotes in evolution but select groups, such as vertebrates, nematodes, some protists and some algae, use vitamin B₁₂ as a growth factor. It is generally believed that fungi, plants and insects don not use B₁₂. However, recent genomic evidence showed that an ancient fungus may use vitamin B₁₂ (8).

In prokaryotes, the components of adenosylcobalamin are synthesized individually and then assembled. The corrinoid ring is made from uroporphyrinogen III, which is a common precursor of heme, siroheme, F430 chlorophyll, and cobalamins. The 5,6-dimethylbenzimidazole (DMB) is made from riboflavin (9,10).

In prokaryotes, AdoCbl serves as cofactors for enzymes such as MCM, glutamate mutase, glycerol dehydratase, ribonucleotide reductase, ethanolamine ammonia-lyase and diol-dehydratase(10,11). MeCbl is required for the vitamin B₁₂ dependent MS encoded by *metH* (12). Alternative forms of cobalamins (corrinoids) are also involved in many other methyl transfer reactions (13).

1.3 Vitamin B₁₂ utilization and genetic diseases of vitamin B₁₂ metabolism in humans

Humans acquire vitamin B₁₂ from meat and dairy products. Daily intake is about 1-5 ug. Inherited vitamin B₁₂ deficiency is rare. Strict vegetarians may have problems of B₁₂ deficiency due to the limitation of B₁₂ source. The elderly may experience B₁₂ deficiency due to decreased absorption with age (14).

1.3.1 B₁₂ absorption

B₁₂ is a water soluble vitamin. Several B₁₂ (Cbl) binding proteins and cell membrane receptors are involved in the absorption and transport of B₁₂ (figure 2). They are transcobalmin I (TCI, a.k.a. R-binder, Haptocorrin), a glycoprotein secreted by salivary glands; intrinsic factor (IF), a glycoprotein secreted by gastric parietal cells; cubilin-amnionless, a multiple function ileal receptor complex (cubilin–amnionless complex, designated *cubam*) (15); and transcobalamin II receptor(16,17). Dietary B₁₂ is released in the stomach and bound to TCI. It travels down to the intestine, where TCI is

digested by pancreatic proteases and B₁₂ is released and captured by IF. IF-Cbl complex interacts with *cubam* in the ileum and is endocytosed into the enterocyte, where Cbl dissociates and enters circulation. In the circulation, cbl is bound to transcobalamin II (TCII, a.k.a. transcobalamin) or TCI. The function of plasma TCI-Cbl is unknown while plasma TCII-Cbl is thought to enter the target cell by transcobalamin II receptor mediated endocytosis mechanism (18). After intracellular process of cbl, it is converted to AdoCbl and MeCbl, which are cofactors of MCM and MS, respectively.

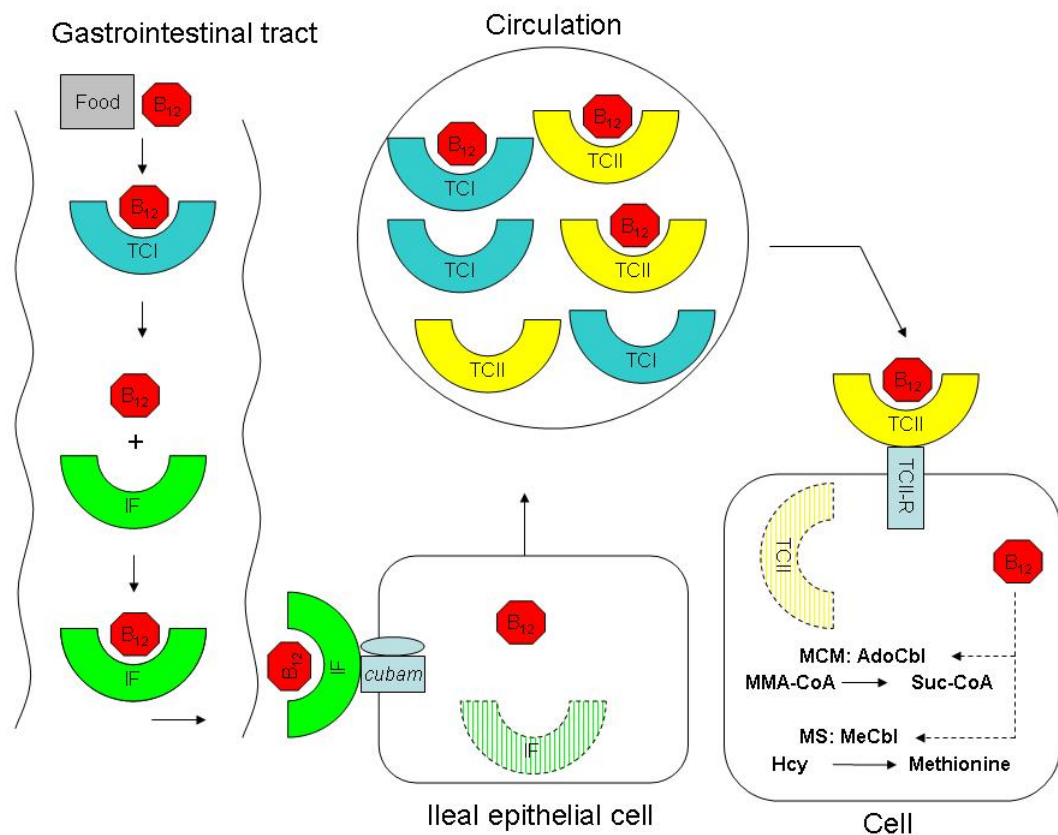


Figure 1.2 Vitamin B₁₂ Absorption (Adapted from (18))

In gastrointestinal tract, B₁₂ is released from food source, first bound to transcobalamin I (TCI) and then to intrinsic factor (IF). Intrinsic factor bound B₁₂ is then transported into blood stream through ileal epithelial cells. This process is mediated by cubilin–amniotin complex (*cubam*). Plasma B₁₂ is bound to TCI or transcobalamin II (TCII).

The transcobalamin II bound B₁₂ is taken into target cells mediated by transcobalamin II receptor and serves as the precursor of cofactors for MS and MCM.

B₁₂ (Red) = cobalamin; TCI (Blue) = transcobalamin I (haptocorrin); IF (Green) = intrinsic factor; TCII (Yellow) = transcobalamin II; *cubam* (Light blue) = cubilin–amnionless complex; TCII-R = transcobalamin II receptor; MMA-CoA = methylmalonyl-CoA; Suc-CoA = succinyl-CoA; Hcy = homocysteine.

1.3.2 The cellular vitamin B₁₂ pathway and genetic disorders

The vitamin B₁₂ pathway in human generally refers to the process by which cells convert cobalamin to its functional cofactors and utilize them in MS and MCM. Genetic disorders represented by nine complementation groups (*cblA-H* and *mut*) have blocks in this process. The genes responsible for some of these complementation groups have been identified (*mut*, *cblA*, *cblB*, *cblC*, *cblE*, and *cblG*), and the genes responsible for complementation groups of *cblD*, *cblF* and *cblH* remain unknown so far (figure 1.3).

TCII-Cbl is thought to be taken into the lysosomes of the target cell. In the lysosome, Cbl is released by protease digestion of TCII and then transported, presumably by a transporter encoded by *cblF*, into the cytoplasm. Once in the cytoplasm, it undergoes an unknown process mediated by *cblC*, *cblD*, and/or *cblH* and becomes usable by MS or enters the mitochondrion where it becomes usable by MCM, the latter sequences involving cbl reduction and addition of the methyl group or adenosyl group, respectively.

metH, was cloned by screening a plasmid DNA library of *E. coli* by complementation of a strain deficient in both methionine synthases(*metE*⁻ and *metH*) by Banerjee *et al* in 1989(19). The B₁₂ binding domain was crystallized and its structure was elucidated by Drennan *et al* in 1994(20).

The human *MTR* gene was identified independently in 1996 and 1997 by three groups, including our lab(21-23), at the time when the genome sequences of *C. elegans* became available. All the three groups took advantage of sequence conservation of MCM between prokaryotic and eukaryotic species and PCR based approaches, except that our group obtained the full-length cDNA sequence using the PCR approach exclusively while the other two groups isolated cDNA clones to help obtain full-length sequences.

1.3.2.1.2 Methionine synthase reductase (MSR); complementation group, *cblE*; gene symbol, *MTRR*

As described previously, MS bound cobalamin cofactor acts as an intermediate methyl carrier between methyltetrahydrofolate and homocysteine. The reaction involves a two-step transmethylation. In the first step, the methyl group of methyltetrahydrofolate is transferred to the MS bound cob(I)alamin to generate the intermediate MS bound methylcob(III)alamin. In the next step, the methyl group is transferred to homocysteine to generate methionine and the enzyme returns to the MS bound cob(I)alamin state which is ready for the next two-step cycle. During the cycling, MS bound cob(I)alamin is highly reactive and may be oxidized to cob(II)alamin state in which the MS is inactive. To circumvent this continuous leakage, in both *E. coli* and humans, cob(II)alamin is reductively methylated to the methylcob(III)alamin state in which MS activity is restored.

MSR accounts for this reductive methylation in humans, while in *E. coli* a two-flavoprotein system, consisting of flavodoxin and a flavodoxin reductase, which had been well-characterized before the human system was identified serves a similar role(24,25). The cloning of the cDNA for human methionine synthase reductase was done by Leclerc *et al* from our lab in 1998 (22). With the prediction that the human counterpart of the *E. coli* reductive methylation system would be a single protein with the combination of flavodoxin and NADPH-dependent flavodoxin reductase, they postulated that the candidate protein would bind FMN, FAD, and NADPH, resembling cytochrome P-450 reductase and NO synthase. Taking advantage of then available human EST database and *C. elegans* genome sequences, they cloned the cDNA for *MTRR* gene. Mutations were found in *cblE* patients. The gene structure was subsequently determined (26). Interestingly, it predicted an alternative splicing event in which a translation product of 725 amino acids, extending beyond the initial 698-amino-acid protein, was also predicted. The additional 27 amino acids at the N-terminus were found to harbor a potential mitochondrial leader sequence when it was evaluated with a program called TRANSPEP. The mitochondrial leader sequence, if it truly exists, implies a mitochondrial role for MSR in addition to the cytosolic role of reactivating MS.

1.3.2.1.3 Methylmalonyl-CoA Mutase (MCM); complementation group, *mut*; gene symbol *MUT*

Methylmalonyl-CoA mutase is a mitochondrial enzyme that catalyzes the isomerization of L-methylmalonyl-CoA to succinyl-CoA. It is the last enzyme in the propionate pathway by which a number of compounds including odd chain fatty acids,

branched-chain amino acids, and cholesterol are converted to succinyl-CoA and therefore enter the Krebs cycle. To be functional, MCM needs AdoCbl as its cofactor. Dietary deficiency of vitamin B₁₂ or blocks in intestinal or cellular uptake or cellular processing of cobalamin (i.e. *cblF*, *cblC*, *cblD*, *cblB*, and *cblA*) lead to secondary-defects in MCM activity. The human MCM protein was purified from placenta and liver, respectively, by the Allen and Rosenberg groups in early 1980s (27,28). The cDNA for the gene (*MUT*) responsible for *mut* complementation group was cloned in 1988 by immunological screening against human placenta and liver cDNA expression libraries by the latter group (29). In their study, they took advantage of an antibody against human MCM and cDNA libraries constructed in a λ gt11 vector. Full-length *mut* cDNA was isolated thereafter by the collaboration of the two groups using polymerase chain reaction (PCR) from the human liver cDNA library by priming with sequences from the 5' end of the longest cDNA described above and sequences in the phage vector(30). In a previous study done by Fenton *et al* in 1984, MCM had been shown to be taken up by mitochondria with proteolytic processing to form mature MCM(28). In the study described here, the N-terminal end of the mature MCM from human liver was also sequenced and compared to the full-length nucleotide sequence. The amino terminal sequence began from the 33rd amino acid of the predicted full-length peptide. These data showed that human MCM has a mitochondrial leader sequence of the first 32 amino acids. Many mutations of MCM in patient cell lines have been identified since then and significantly contributed to the molecular genetic understanding of the disorder(31,32). The X-ray crystal structure of the MCM from *Propionibacterium shermanii* was published (33,34).

1.3.2.1.4 ATP: cobalamin adenosyltransferase (MMAB protein); complementation group *cblB*; gene symbol *MMAB*

This complementation group corresponds to a mitochondrial defect in the synthesis of AdoCbl. It had been shown years ago that the defect was in the adenosyltransferase required to add adenosyl to reduced cobalamin (35,36). The gene was cloned by two groups. Leal *et al* in 2003 cloned the gene by screening a bovine cDNA expression library for complementation of a *Salmonella enterica* strain deficient in adenosyltransferase (37). They initially screened about 250,000 clones from a human liver cDNA expression library but did not find any positive clones. A bovine cDNA library was then used in screening and 4 identical positive clones were found. The clone represented partial amino acid sequence of the putative bovine adenosyltransferase and was 29% identical to the N-terminal domain of PduO (the adenosyltransferase of *S. enterica*). Using BLASTP searches, they found an I.M.A.G.E. clone representing the full-length of the human ortholog, which encodes a 250-amino-acid protein with 25% identity to PduO. Subsequently, the authenticity of the clones was confirmed when the recombinant proteins showed adenosyltransferase activity *in vitro*. In our laboratory, Dobson *et al* in 2002 cloned this gene and the gene responsible for *cblA* (refer to the next section) by searching bacterial and archaeal genomes for gene clusters (operons) containing the gene for methylmalonyl-CoA mutase (38,39). The rationale was that in many prokaryotic genomes, the genes for functionally related proteins are located in close proximity to each other in operons. Dobson *et al* examined the MCM operon in the COG database at NCBI and found two genes adjacent to MCM in *Archaeoglobous*

fulgidus. One of them eventually led to the identification of the gene responsible for *cblA* in a similar approach described here below. The other represented an unknown gene in *Archaeoglobous fulgidus*, which encoded a protein with sequence similarity to PduO. The sequence of this protein was used in BLASTP searches of the nonredundant database at NCBI identifying an unknown human protein with sequence similarity to PduO and the protein in *A. fulgidus*. Importantly, mutations were found for this gene in *cblB* patients, which confirmed it was the human cobalamin adenosyltransferase gene. More mutations were identified thereafter in *cblB* patients (40-42) and studies on the mechanism of cobalamin adenosylation and the impact of mutations on ATP and cobalamin binding were reported by Leal *et al*, Saridakis *et al*, and Zhang *et al* in our laboratory(37,43,44). The bacterial ortholog from *Thermoplasma acidophilum* as well as the human protein have been crystallized (43,45).

1.3.2.1.5 Putative methylmalonyl-CoA mutase chaperone (MMAA protein); complementation group *cblA*; gene symbol *MMAA*

Patients from this group have a block in AdoCbl synthesis and initial expectations were that the defect would be in a cobalamin reductase. The gene was cloned in the same study by Dobson *et al* as described above in the gene cluster with the genes for mutase (MCM) and adenosyltransferase (39). Examination of the predicted functional domains of the protein based on the deduced amino acid sequence suggested it belongs to the G3E family of P-loop GTPases. Over 23 mutations have been identified in the *cblA* patients since then (41,42,46). The work done by Korotkova *et al* and by Padovani *et al* suggests that it binds and hydrolyzes GTP and interacts with the mutase, suggesting it might be a

mutase chaperone (47-49). The work done by Dobson and Froese in our lab showed the *E. coli* orthologue of MMAA (*YgfD*) is also GTP-dependent and interacts with the orthologue of mutase (*Sbm*) (unpublished).

1.3.2.1.6 Putative TonB function; complementation group *cbIC*; gene symbol *MMACHC*

The *cbIC* group is the largest of the complementation groups with over 200 patient cell lines assigned to it (larger than the sum of all other groups). These patients have homocysteinemia (uria) and methylmalonic acidemia (uria) and therefore have a defect in the common pathway leading to the synthesis of both cofactors. The locus of *MMACHC* was first mapped to chromosome 1p34-1p32 region using linkage studies involving some ten families with affected children (50). As reported in 2006, Lerner-Ellis *et al* (51) refined this interval by haplotype analysis and homozygosity mapping. Our laboratory contributed to the analysis by screening for candidate genes in this region that showed a potential sequence relationship to bacterial genes involved in vitamin B₁₂ metabolism. The gene was ultimately found by brute force sequencing by Lerner-Ellis *et al*. Mutations were found in one gene in this area which was then confirmed as the gene for *cbIC* (51). Mutations were identified in genomic DNAs from 190 of 204 *cbIC* patients. The cellular defect in MS and MCM of *cbIC* was complemented by expressing wild type *MMACHC* in immortalized *cbIC* fibroblasts. A survey for motifs that resemble motifs in bacterial genes with cobalamin-related functions revealed that *MMACHC* had a putative cobalamin-binding domain and the C-terminal amino acid sequence of *MMACHC* was found to have a 51% similarity to part of the *E. coli* tonB protein which is involved

in transport of iron and cobalamin across the periplasmic membrane of the bacterium. In human cells, such a protein might be involved in bringing B₁₂ into the cytosol across the lysosomal membrane. It is a possible accessory protein to transport across the lysosomal membrane or it might act as a cytosolic B₁₂ chaperone.

1.3.2.2 Complementation groups without a gene identification

The genes for *clbF*, *cblD* and *cblH* are the missing pieces of the vitamin B₁₂ pathway. The gene for *cblF* may represent a lysosomal protein that is responsible for the transport of cobalamin out of lysosome into cytosol (see below). The genes for *cblD* and *cblH* may represent the transport or/and modification of cobalamin after it is released from lysosome.

1.3.2.2.1 Lysosomal efflux protein; complementation group *cblF*

Cultured fibroblasts from a patient with methylmalonyl aciduria were found not to be able to release unmetabolized, nonprotein-bound vitamin B₁₂ into the cytosol and therefore accumulated it in the lysosome. (52). Complementation studies were done thereafter and a new complementation (*cblF*) distinct from the previous complementation groups that had been identified (*cblA*, *cblB*, *cblC* & *cblD*), was established (53). The intracellular Cbl distribution was visualized in fibroblasts from *cblF* patients and from normal humans by electron microscope autoradiography in a study done by Vassiliadis *et al* 1991 and over 60% of radioisotope labelled cobalamin was found in the lysosome of patients' fibroblast while only 4.7% was found in that of normals. These results support

the hypothesis that the defect in *cbfF* is due to the failure of transport of Cbl across the lysosomal membrane into cytosol (54).

The *cbfF* patients have combined symptoms of homocystinemia (uria), and methylmalonic acidemia (uria) since both of pathways are blocked. The clinical symptoms at diagnosis, the age of presentation and complication varies among patients (55).

1.3.2.2.2 Complementation group *cbfD*

The clinical case for the *cbfD* complementation group was first reported as two affected brothers from a consanguineous family by Goodman *et al* in 1970(56). Later, in 1978, by complementation studies of cultured fibroblasts, Willard *et al* showed that they belonged to a complementation group distinct from *mut*, *cbfA*, *cbfB*, and *cbfC* patients, although they shared the same biochemical features with *cbfC*, i.e. deficient activity of both MCM and MS as well as decreased level of their respective cofactors AdoCbl and MeCbl(57). Later complementation studies showed *cbfD* was distinct from *cbfF*(53). The two cases for this sibship remained the only identified *cbfD* cases until 2004 when Suormala *et al* reported 3 new patients representing two isolated variants with deficient activity of either MCM or MS, but not both(58). These findings put *cbfD* in the unique position that it causes either isolated or combined deficiency of AdoCbl and MeCbl synthesis. As discussed by Suormala *et al*, two likely models were proposed to explain the exceptional findings. One explanation is that the *cbfD* protein contains at least three different domains, one for B₁₂ binding, one for presentation of reduced B₁₂ to MS, one for

presentation to mitochondria. Another explanation is that two isoforms of *cblD* proteins exist by alternative splicing, one for MS activity and the other for mitochondrial function.

1.3.2.2.3 Complementation group *cblH*

The first and only case for the *cblH* complementation group was first reported in 1990 by Cooper *et al* in a complementation study among fibroblasts from a patient with the same clinical and biochemical phenotype as the *cblA* complementation group(59). Originally, it had been designated as a *cblA* variant. Then, in 2000, Watkins et al (60)reported that this cell line complemented each of 28 *cblA* cell lines and therefore ruled out the likelihood of interallelic complementation within the *cblA* group. It has been designated as a distinct complementation group (*cblH*) since then (60). However, as discussed by Suomala *et al* in 2004, it is possible that this *cblH* line actually belongs to the *cblD* complementation group because no complementation tests have ever been done between the two groups (58).

Chapter Two: Complementation-based cloning to identify vitamin B₁₂ genes

2.1 Introduction

Vitamin B₁₂ (cobalamin, Cbl) is the precursor of cofactors for two enzymes in humans. The mitochondrial enzyme methylmalonyl-CoA mutase (MCM) requires 5'-deoxyadenosylcobalamin (AdoCbl) as cofactor while the cytoplasmic enzyme methionine synthase (MS) requires methylcobalamin (MeCbl). Humans do not make cobalamin but acquire it from food sources. Dietary Cbl is absorbed in the intestine, transported in the blood stream, and taken up by the cells (refer to chapter one for detailed review). The intracellular processing of Cbl is not fully understood. Cbl is first released from the lysosome into the cytosol in the form of cob(III)alamin (Co³⁺) oxidation state. Cob(III)alamin is reduced to cob(I)alamin and methylated in the cytosol or adenosylated in the mitochondrion. Any defects that block the steps from precursor to cofactors will lead to dysfunction of the cofactor-dependent enzymes, which will result in methylmalonic acidemia if production of AdoCbl is impaired and homocystinemia if production of MeCbl is impaired. In humans, genetic defects of intracellular utilization of Cbl are represented by 9 complementation groups (*cblA-H* & *mut*). *mut* and *cblG* are the complementation groups that represent MCM and MS defects, respectively. *cblA*, *cblB* and *cblH* are complementation groups in which only AdoCbl synthesis is impaired. *cblE* represents methionine synthase reductase (MSR), which is required to keep MS active by reductive methylation of cob(II)alamin on MS, which becomes oxidized to the cob(II) state over time. *cblF* and *cblC* represent earlier steps in the pathway and cause combined defects of MCM and MS. The *cblD* defect causes either isolated or combined deficiency

of MeCbl and AdoCbl synthesis. The genes for *mut*, *cblA*, *cblB*, *cblC*, *cblE*, *cblG* have been previously identified (refer to Chapter one for detailed review). *cblB* (MMAB) encodes a mitochondrial ATP:cobalamin adenosyltransferase. *cblA* (MMAA) also encodes a mitochondrial protein. Its function has not been determined but bacterial orthologs have GTPase activity and interact with MCM. The gene for *cblC* (MMACHC) has been cloned. It acts at an early step but its function remains unclear. The genes for *cblF*, *cblD*, and *cblH* have not been identified yet. For all three complementation groups, AdoCbl synthesis is defective. *cblF* is thought to encode a lysosomal Cbl transporter to release Cbl from the lysosome into the cytosol. Only a few cases have been reported for *cblD* and one case of *cblH* has been reported. As nobody did complementation test between *cblD* and *cblH*, the possibility that they represent one complementation group should not be ruled out (see chapter 3 for recent update on *cblD* and *cblH*).

The intracellular AdoCbl synthesis can be evaluated by monitoring MCM activity by [¹⁴C]-propionate incorporation, an autoradiographic procedure (61) which detects [¹⁴C]-propionate incorporation into trichloroacetic-acid-precipitable material by fibroblast monolayer *in situ*. The [¹⁴C]-propionate is taken up, converted to its CoA derivative and metabolized in two steps by propionyl-CoA carboxylase (PCC) and methylmalonyl-CoA epimerase (MCEE) to L-methylmalonyl-CoA. Mutants with defective AdoCbl synthesis or a primary defect of the MCM cannot metabolize [¹⁴C]-propionate further, while normal cells incorporate the isotope into labeled macromolecules via succinyl-CoA and the Krebs cycle. As AdoCbl synthesis is defective in *cblF*, *cblH* and *cblD*, mutant human fibroblasts have much lower incorporation than normal human fibroblasts. The incorporation of the radioisotope is monitored *in situ* by autoradiography, where the

intensity of silver grains over cells is the indicator of incorporation levels (figure 2.1 and figure 2.2). This makes a complementation-based cDNA library screen for genes responsible for these complementation groups possible. In this study, two cDNA libraries are used to screen the clones that can “complement”/correct the incorporation defects of *cblH* and *cblF* human fibroblasts.

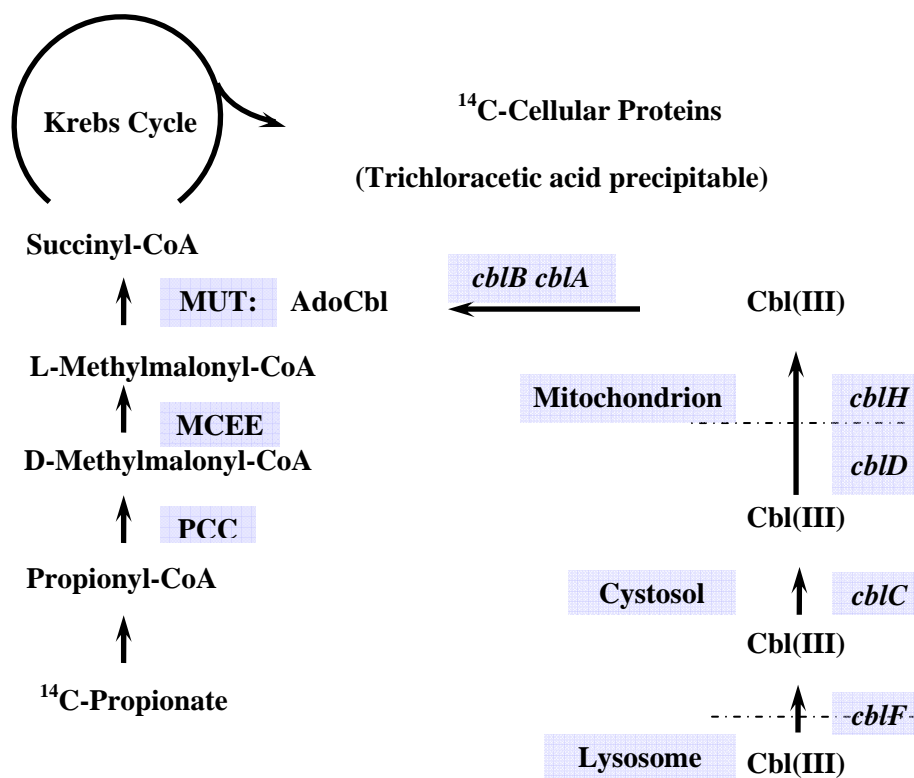


Figure 2.1 Principle of [^{14}C]-propionate incorporation to monitor B_{12} metabolism.

2.2 Materials and methods:

2.2.1 Cell lines

Fibroblast lines from control human and patients used in this study were provided by David Rosenblatt at the Canadian Repository for Mutant Human Cell Strains (McGill University). The *cblF* lines were WG1350 and WG2432. The *cblH* line was WG1437. The *cblB* line was WG2325. The *cblC* was WG2473. The normal human fibroblast lines were MCH39 and MCH65.

2.2.2 Tissue culture

Normal and patient fibroblast cells were cultured in Dulbecco's Minimal Essential Medium (DMEM Invitrogen) supplemented with 5% fetal bovine serum, 5% calf bovine serum, 1× Antibiotic/Antimycotic. Cells were incubated at 37 °C in 5% CO₂, water saturated atmosphere, washed with Dulbecco's phosphate buffered saline with calcium or magnesium, and harvested with 0.25% trypsin (Invitrogen).

2.2.3 Preparation of polyA+ RNA

Total cellular RNA was prepared from HepG2 cells using Qiagen RNeasy Kits according to manufacturer-suggested procedure. PolyA+ RNA was then isolated from the total cellular RNA using a Stratagene "Absolutely mRNA™ Purification Kit" according to manufacturer-suggested procedure. The final concentration of polyA+ RNA was determined by RiboGreen RNA assay kit (Molecular Probes) and was 0.01 ~ 0.02 µg/µl.

2.2.4 Gene delivery of human fibroblasts

2.2.4.1 Microinjection

Microinjection is a method to effectively deliver nucleic acids into human skin fibroblasts. A microinjection and micromanipulation system (InjectMan NI2 and

FemtoJet from Eppendorf) was used for adherent cell microinjection. Needles for microinjection were pulled by a micropipette puller (Model P-87, Sutter Instrument). The glass capillaries used to make needle were purchased from World Precision Instrument (TW120F-3). The parameters used to make needles were as follows: Heat=595, Pull=165, Velocity= 100, Time= 135. PolyA+ RNA was injected into cytoplasm while plasmids containing expressible cDNA were injected into the nuclei of cells for proper expression.

2.2.4.2 Nucleofection

Primary human skin fibroblasts are difficult to transfect with exogenous DNA for all kinds of non-viral transfection systems. Recently, AMAXA developed an electroporation based system called “nucleofection” and claimed high delivery efficiency with previously difficult primary cells. The equipment (Nucleofector I) and reagents (basic fibroblast kit) were bought and relatively high delivery efficiency (~25%) on our human skin fibroblasts was obtained. In a typical experiment in this study, 1 ug of plasmid DNA and a half million fibroblasts were used in each nucleofection. The preset program “U-20” designed for high transfection rate but with low survival rate was used.

2.2.5 [^{14}C]-*Propionate incorporation in human cells*

Cells were plated onto Lab-Tek or Lab-Tek II chamber slides (Nunc) and incubated overnight under the condition described in section 2.2.2. Cells on the slides were then rinsed with phosphate-buffered saline once and incubated with labeling mix (Puck's saline (table 1) containing 15% dialyzed fetal bovine serum (Invitrogen 26400-044), 500 μM L-alanine, 10 mM glucose, 100 μM , 1 $\mu\text{Ci/ml}$, 10 Ci/mole [$1\text{-}^{14}\text{C}$]-sodium propionate (Moravek Biochemicals, MC-201)) at 37 °C in a 5% CO_2 , water saturated

atmosphere for 10 hours to overnight. At the end of the incubation, the labelling mix was discarded and the chamber frames removed. The slides with a monolayer of cells were washed 3 times in 0.9% sodium chloride solution for 5 minutes, fixed in 3.7% formaldehyde / 0.9% sodium chloride solution for 10 minutes, washed in 5% ice-cold trichloroacetic acid for 5 minutes and washed in water for 10 minutes. The slides were then air dried or ethanol dried.

Components	g/l
NaCl	7.4
KCl	0.285
Na ₂ HPO ₄ ·7H ₂ O	0.29
MgSO ₄ ·7H ₂ O	0.154
CaCl ₂ ·2H ₂ O	0.016
KH ₂ PO ₄	0.083
Phenol Red	0.005
NaHCO ₃	1.2

Table 2.1 Puck's saline

2.2.6 Autoradiography

In the dark room, dried slides from section 2.2.5 were vertically dipped into pre-warmed emulsion (42 °C), and placed horizontally with the side with cells facing up until the emulsion solidified. Slides coated with emulsion were put into a light-proof box and stored in a cold room (~ 4°C) to expose 2 to 4 weeks. With all the solutions at 14 °C, the slides were developed in D-19 developer (Kodak) for 3-4 minutes, washed in 2% acetic acid for 30 seconds, fixed in fixer (Kodak) for 3-4 minutes and then washed twice in

water. The slides were then air dried and examined for silver grains over fixed cells using a light microscope.

Note: Emulsion NTB-2 (Kodak) was initially used until its production was discontinued and replaced by NTB (Kodak) in this study. Emulsion was used directly or 1:1 diluted in double distilled water.

2.2.7 cDNA library screening

2.2.7.1 cDNA library from Clontech

A mammalian expressible human liver cDNA library was purchased from Clontech to use as a source of clones for complementation-based cloning in mutant fibroblast cell lines. This library was constructed in a pEXp1 vector and mammalian transcription is driven by a CMV promoter upstream of inserted cDNA. The library was amplified once and contains 1.0×10^7 independent clones with cDNA size ranging from 0.5- 3.5kb and average 1.6 kb. The library was re-amplified once after it was received by plating on LB agar plates overnight. For each plate, the *E. coli* clones (about 8000 ~ 10,000 per plate) were collected into 2ml tube with 20% glycerol/LB and stored in -80 °C freezer. Plasmids were extracted (Qiagen Miniprep Kit) from *E. coli* from 30 tubes and used for nucleofection of mutant fibroblasts and these cells were tested for restoration of [¹⁴C]-propionate incorporation. The strategy to screen this library is to identify a positive pool in the initial screen, followed by through several rounds of sib selection to identify a single clone. Briefly, once a positive pool has been assigned, this pool will be further divided into smaller pools and the same procedure will be applied to these pools to assign positive outcomes. After several rounds of sib selection, a pure clone will be identified.

2.2.7.2 cDNA library from OriGene

A second mammalian expressible human liver cDNA library (“rapid-screen arrayed cDNA library”) was purchased from OriGene to use as a source of clones for complementation-based cloning in mutant fibroblast cell lines. The library was designed to contain large inserts and the mammalian expression of these inserts was driven by CMV promoter (pCMV6-XL4 vector). Unlike conventional cDNA library, in which all clones are mixed in one tube, to make this library, “sub-plates” with 96 wells each were first made. Each subplate well was prepared as a collection of 50 randomly picked clones. A replica sub-plate made up of 96 wells of 50 clones each was pooled and plated into a single well of the “Master plate”. Therefore the 96 wells of the “Master plate” contain 5000 clones each (x100 wells = 500,000 clones). In that way, a positive well on the master plate makes it possible to screen the corresponding subplate with the expectation that one or more wells of the subplate will also give a positive outcome. To screen this library, plasmid DNA from each well of the master plate (approximately 500,000 clones in total) was microinjected or nucleofected into mutant fibroblasts and these cells were tested for restoration of [^{14}C]-propionate incorporation. If a positive well is identified, the corresponding sub-plate will be ordered from OriGene. Another round of functional testing will be done for each well of the “sub-plate” to identify positive wells that can restore the [^{14}C]-propionate incorporation in mutant cells. Positive wells would now represent a sampling of 50 clones. To analyze these, the contents of positive wells of the “sub-plate” will be plated for colonies. Colonies will be picked and propagated for plasmid DNA and tested for restoration of the [^{14}C]-propionate incorporation in mutant cells.

2.3 Results

2.3.1 Selection of culture conditions and tester mutant cell lines

Depending on how severe the defect is, mutant cells will show different residual ability to incorporate [^{14}C]-propionate, which will result in “background” after autoradiography. The background [^{14}C]-propionate incorporation of mutant cell lines (2 for *cbiD*, 2 for *cbiF*, and 1 for *cbiH*) was very high when cultured in the MEM- α medium used by the Canadian Repository for Mutant Human Cell Strains (data not shown). Several cell culture conditions were tested to bring down the background, including increasing culture temperature for temperature sensitive mutants and using low B₁₂ medium. There was no effect on ^{14}C -propionate incorporation level of mutant cell lines cultured in higher temperature. But the background was successfully reduced to a very low level for the *cbiH* cell line (figure 2.3 c.) by growing the cells in low B₁₂ medium as described in section 2.2.2, a condition that did not affect normal cells. Two additional *cbiF* cell lines were also identified with low background in low B₁₂ medium by testing several cell lines (figure 2.3 e.). In addition, a *cbiB* cell line was selected with low background that could be used as positive control for complementation since the gene defect is already known (figure 2.3 a.). Two *cbiD* cell lines were available for the lab. However, both of them were so “leaky” that they were hard to distinguish from normal fibroblasts when tested for [^{14}C]-propionate incorporation level. No additional *cbiD* cell lines were requested because of the likelihood that *cbiH* and *cbiD* represented the same gene ((58); see in Chapter 3 for recent update).

2.3.2 Complementation by polyA+ RNA from HepG2 cells

To assess whether rare transcripts corresponding to B₁₂ genes could be detected in complementation experiments, polyA+ RNA isolated from HepG2 cells was microinjected into mutant cells of *cblB*, *cblF*, and *cblH*. Injected cells were subjected to [¹⁴C]-propionate incorporation assay. The incorporation was restored for most of injected cells in all three cases (figure 2.3). This demonstrates that the injected RNA contains the specific transcript required to restore function to the genetic defect. Previous work done by this lab confirmed that the correction of [¹⁴C]-propionate incorporation of mutant fibroblasts is not due to physical damage of mutant cells caused by microinjection but is rather due to transgene molecules injected(62).

The correction by polyA+ RNA showed that the target polyA+ RNA species were at high enough frequency to functionally complement mutant defectst. This implies that cDNA made from similar tissue sources would have a high enough titer of the relevant cDNA to complement the mutant defect as well. This experiment represents proof-of-principle that a well made cDNA library could be used to identify the affected gene.

2.3.3 Screening the Clontech library

2.3.3.1 Test of the system

To generate a positive control for the complementation experiment and test the Clontech cDNA library, a *cblB* clone (*MMAB* gene) was isolated from the library by former PostDoc Xue Yang and named as pEXp1-MMAB. The insert of this clone was sequenced and confirmed as full length. *cblB* mutant fibroblasts were microinjected and nucleofected with the plasmid to see if this plasmid would correct ¹⁴C-propionate incorporation. It turned out the clone successfully corrected *cblB* mutant cell line either

nucleofected (figure 2.4 a.) or injected (figure 2.4 c.). This proved that the expression vector system works.

2.3.3.2 Library screening

30 pools of Clontech library (about 10,000 clones per pool) were screened for complementation of the *cblF* or *cblH* defects. Pools of plasmid DNA were delivered into mutant cells by nucleofection. No positive pools were identified for *cblF*. An ambiguous positive pool (#12) was found for *cblH* (figure 2.5). This pool was divided into smaller pools (1000 clones per pool, 30 pools to fully cover the original pool) and each pool was screened. However, no positive pools were identified.

2.3.4 Screening the Origene Master Plate

2.3.4.1 Test of the system

To test the OriGene library, individual expressible cDNA clones were purchased from OriGene that were isolated from their rapid-screen arrayed cDNA library. These included cDNA clones corresponding to human acid beta-galactosidase, *cblA*, *cblB*, and *mut*. These clones contained full-length coding regions and were built in the same (pCMV6-XL4) or similar vectors (pCMV6-XL6) as the one used in the OriGene cDNA library purchased (pCMV6-XL4). The human acid beta-galactosidase clone was obtained because the beta-galactosidase activity assay, which was an X-Gal-based procedure as used for bacterial LacZ, would take overnight while it would take a week or more to test the cDNA clones for *cblA*, *cblB*, and *mut* by restoration of [¹⁴C]-propionate incorporation. A G_{M1} gangliosidosis cell line (beta-galactosidase deficient) was tested at acid pH for X-gal staining after nucleofection and the clone was shown to restore the blue straining resulting from the X-gal substrate (date not shown). For the B₁₂ defects (*cblA*,

cblB and *mut*), the pure clones were nucleofected into corresponding mutant cells and assayed for [^{14}C]-propionate incorporation. All of them were functional and corrected the defect using this assay (figure 2.6 experiments were done by Marko Vujanovic, another graduate student).

2.3.4.2 Library screening

Each of 48 wells of the OrigGene “Master Plate” (96 wells in total) has been screened by microinjection of aliquots of plasmid (0.1 $\mu\text{g}/\mu\text{l}$) into the *cblH* or *cblF* mutant cells followed by [^{14}C]-propionate incorporation assay. No positive well was identified. In addition, the cDNAs from each of the remaining wells of the master plate (48 wells) have been nucleofected into the *cblF* mutant cells. These slides are currently being exposed to emulsion. The result will be examined later and will not be described in this thesis.

2.4 Discussion

As summarized in the first chapter, different clues have led to the identification of several genes involved in vitamin B₁₂ processing. The scarcity of clinical cases and the lack of information on counterparts in other organisms for the genes of the remaining complementation groups make it hard to reproduce the previous successes (i.e. immunological screening, positional cloning, cloning approaches based on conservation of sequences across species, etc.).

However, the ability to monitor ^{14}C -propionate incorporation in skin fibroblasts derived from patients by autoradiography makes a functional/complementation screening method possible. Classically, cDNA library screening based on complementation has been widely used in yeast and bacterial systems. With the emergence of mammalian

expressible cDNA libraries and effective gene transfer methods, a few successful cases have been reported in which genes were isolated in mammalian systems based on correcting mutant phenotype by complementation, e.g. Torben et al. cloned the CDG-IIc, a GDP-fucose transporter, using a human liver cDNA library (Clontech) to complement patient fibroblasts (63).

For complementation-based cloning to work for the vitamin B₁₂ disorders, it is necessary that the mutant cell lines to be used as testers have very low background in the ¹⁴C-propionate incorporation experiment. In this study, the *cblH* cell line was shown to be very deficient in [¹⁴C]-propionate incorporation, whereas the two *cblF* cell lines were not as “tight” as *cblH* but were low enough in isotope incorporation for a complementation-based experiment. The major aim of this thesis was to establish the conditions to make possible the isolation of complementing cDNAs.

A key result was to show that total HepG2 polyA⁺ RNA could correct *cblB*, *cblH* and *cblF* when microinjected into mutant cells (figure 2.3 b. d. f.) This meant the required mRNA species at the frequency present in unfractionated polyA⁺ RNA was sufficient to permit complementation even at the low volumes used in microinjection. Attempts were made to fractionate polyA⁺ RNA. However, none of the attempts yielded a practical method for the ultimate identification because of the fragile nature of RNA. However, as HepG2 is a liver carcinoma cell line, the correction by its polyA⁺ RNA indicated that the liver should be a very good source for B₁₂ gene transcripts. Indeed, when known B₁₂ genes were checked for their expression profile by analysis of EST counts available from NCBI, liver was one of the three best sources for B₁₂ gene transcripts (ranging from 18 (for *MMAA*) to 192 (for *MUT*) copies per million; two other

good sources were brain and kidney). In a published study, the average injection volume for HeLa cells determined by injecting ^{125}I -labeled serum albumin solution was about 1×10^{-13} L (64). Human fibroblasts are much larger than HeLa cells in volume and likely can receive at least ten-fold larger injection volumes, i.e. at least 1×10^{-11} L. Thus, as many as 100,000 transcripts would be microinjected into each cell based on this estimation. Given the estimated transcript frequency, this number is consistent with the result that most of the injected cells were positive for silver grains. This result implies that only a few copies of gene transcripts are sufficient to rescue the defect. The RNA experiment provided the basis for the cDNA library to work, where only a few copies of target cDNA clones would be required to be delivered into each cell.

All the full-length clones tested were functional (pEXp1-MMAB for the Clontech library and *cblA*, *cblB*, *mut*, beta-galactosidase for the OriGene library). This result not only proved that the expression system based on these vectors would work as long as the insert was full-length, but also established the experimental system for the cDNA library screen (i.e. successful gene delivery, complementation, and $[^{14}\text{C}]$ -propionate incorporation).

Both microinjection and nucleofection were used in this study. Each method has benefits and drawbacks. Microinjection would deliver several thousand (or more) copies of polyA⁺ RNA or plasmids into each cell. Relatively few cells were injected (50 to 100) and were assessed as injected fields on each slide. This gave a sensitivity of 50-100 times the injection count. However, this was still a very labour intensive method, especially given that 2 or 3 sets of slides for each pool and cell line were used in the screen. It was initially thought that microinjection would provide better results simply because more

cDNAs were delivered into individual cells compared to nucleofection. However, in a series of experiments in which pure clones (*cblA*, *cblB*, and *mut*) were delivered either by microinjection or nucleofection for restoration of [^{14}C]-propionate incorporation of corresponding mutant cells (experiments done by another student Marko Vujanovic in our lab), nucleofection provided comparable results with those by microinjection. On the slides done by nucleofection, virtually all the cells were positive with different levels of silver grains covered on top of them (data not shown). As nucleofection was less labour intensive, it was therefore used in the *cblF* screen for the remaining 48 wells of the OriGene “Master Plate”. The only drawback is that nucleofection uses many more cells and DNA. Because vigorously growing, early passage cells need to be used for these experiments, it has been challenging to get enough cells for primary fibroblasts used in this study. Alternatively, immortalized cell lines could be made to surmount this obstacle.

In the screen of the Clontech library, an ambiguous positive pool (#12) was found for *cblH* (figure 2.5 d.). Two sets of slides (approximately 20,000 cells) were examined for positive cells. The cluster of cells shown by figure 2.5 d. was the only positive cluster of cells found. The cluster of positive cells could be explained by crossfeeding from a single positive cell, which is often observed by this procedure. The positive pool was subdivided into 30 smaller pools (1000 clones each) which were screened. More positive cells were expected to be seen as the target clones should be 10 times enriched in at least one of the sub-pools. However, no positive cells were seen on any of the slides examined. It was very likely that the positive cells seen initially were false positives. In retrospect, however, it might be necessary to repeat the original pool (#12) for the reoccurrence of positive cells.

The number of independent clones to be screened to saturate a library can be calculated using the equation: $N = \ln(1-95\%)/\ln(1-f)$, where N stands for number of independent clones to be screened and f stands for the frequency of the target clones in the library (65). For example, by applying this equation, if the frequency of the target clone is 1 in 10^7 , the number of clones to be screened would be 3×10^7 ; if the frequency of the target clone is 1 in 1000, the number of independent clones to be screened is about 3000. A frequency for B_{12} gene transcripts should range from 1 in 5,000 to 1 in 50,000 if a liver source was used. Therefore, about 15,000 to 150,000 clone needs to be screened for complete coverage of this library. However, the calculation was based on the assumption that every transcript was successfully converted into full-length cDNA clones, and therefore the quality of the cDNA libraries (i.e. the frequency of full-length clones) became the “loose-end” in this study. Two cDNA libraries have been screened in this project. About 300,000 and 250,000 clones were screened, respectively, from the Clontech and OriGene libraries. These numbers were well over the number of clones to be screened for completeness if full-length clones dominated the libraries. It is hard to evaluate the quality of the library although individual library manufacturers always claim a high percentage of full-length clones.

An attempt to isolate all the known B_{12} genes and holocarboxylase synthase (HCS) from the Clontech library by a hybridization-based method was done by the former PostDoc Xue Yang in our lab. Only the cDNA clone for *cblB* was isolated (the pEXp1-MMAB in section 2.3.3.1). Though uneven amplification of the library clones was suspected in her experiment, failure to isolate a majority of the attempted clones hinted that the quality of the cDNA library was not great. The OriGene liver cDNA

library was chosen for a second try because it was thought to be of higher quality. The library was constructed for larger inserts and clones were individually picked to avoid loss of clones due to uneven amplification. And because of this, fewer clones needed to be screened than from conventional libraries to achieve equal representation. Further, clones were impossible to lose in the “sub-plate” once a positive well was identified, because each well of the master plate was put together directly by combining the 96 wells of each subplate so that any clone in the Master plate had to have come from at least one well of the subplate. Despite all these advantages, it is still hard to evaluate the percentage of full-length cDNA clones in this library unless a number of clones are selected to examine “full-lengthness”. OriGene obtained full-length cDNA clones that cover 98% of the human genome (<http://www.origene.com/cdna/>), most of which were isolated from Rapid-screen cDNA panels by PCR-based screen. However, these clones were probably screened across as many as 13 Rapid-screen cDNA panels. It would be interesting to know how many of known B₁₂ genes were isolated from the liver Rapid-screen panel used in this study. Unfortunately, this information isn’t released by OriGene.

2.5 Future directions

The gene for *cbiD/cbiH* in the course of this study was identified by another research group (personal communication). They confirmed that the same gene accounted for both *cbiD* and *cbiH* complementation groups by identifying mutations accounting for both groups. *cbiF* is yet to be identified. For the cDNA library-based screen, the rest of 48 wells need to be examined to compensate for the potentially low frequency of full-length clones. If a pure clone is identified that corrects mutant cells, the corresponding

gene should be sequenced in multiple mutant cell lines for mutations. This will provide the ultimate authenticity of the gene responsible for *cblF*.

Meanwhile the approaches to clone *cblF* should not be limited to screening cDNA libraries. Though polyA⁺ RNA has been hard to work with, polyA⁺ RNA from HepG2 cells nevertheless provided us a starting point since complementation was achieved. It was reported polyA⁺ RNA could be size fractionated long time ago (66). If it can be repeated by us, we would at least know the size of the transcript for *cblF* and the enriched polyA⁺ RNA could be a good source to make a cDNA library to screen.

A comparative genomics approach is also described in the next chapter to identify B₁₂ genes.

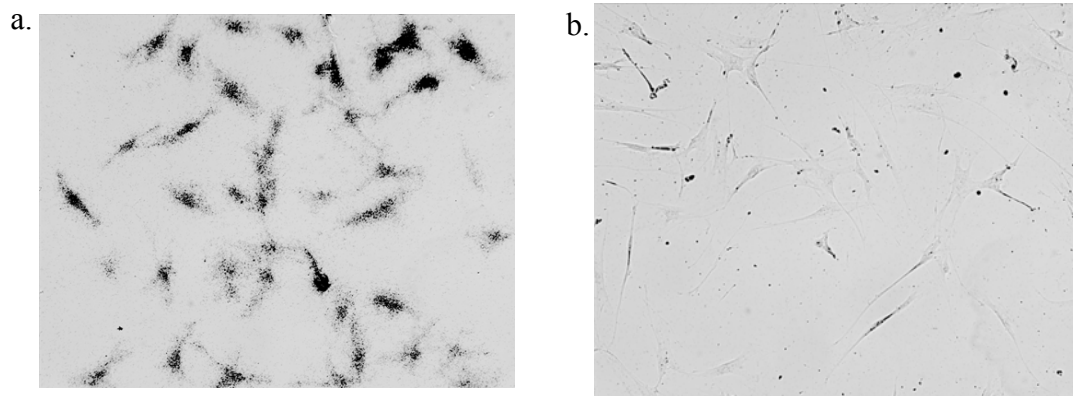


Figure 2.2 Autoradiography of TCA-precipitable protein after [^{14}C]-propionate incorporation

a. MCH39 (normal fibroblasts) **b.** WG2473 (*cbIc* mutant fibroblasts); **a.** shows that normal fibroblasts have high level of [^{14}C]-propionate incorporation and the intensity of silver grains over cells is very high. While in **b.**, mutant fibroblasts defective in *cbIc* gene are impaired in incorporating [^{14}C]-propionate and almost no silver grains are seen over the cells in this cell line shown here.

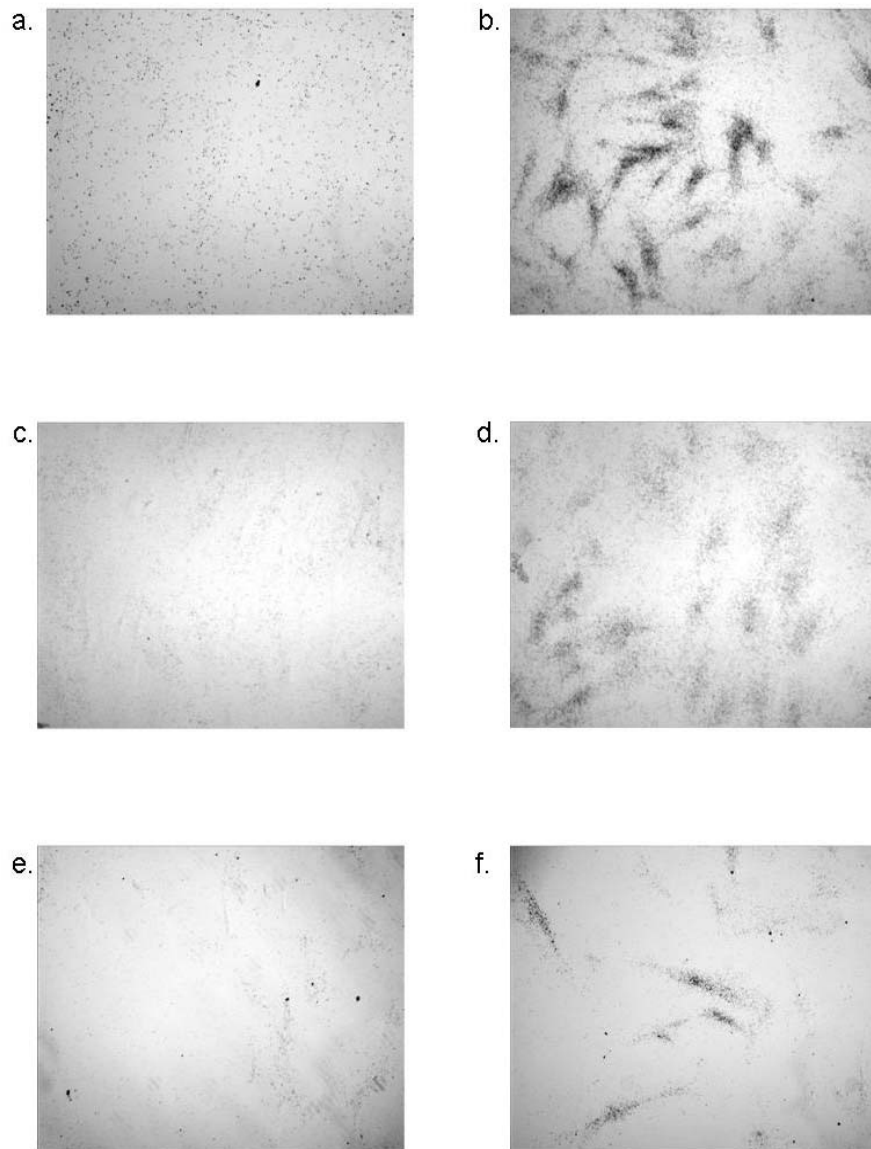


Figure 2.3 Microinjection of HepG2 mRNA

a. c. e. show uninjected control cells for *cblB*, *cblF*, and *cblH*, respectively; **b. d. f.** show injected cells for *cblB*, *cblF*, and *cblH*, respectively; Compared to the uninjected controls, intense silver grains can be observed over many cells in the injected areas for all the three cases, which indicates restoration of [^{14}C]-propionate incorporation in these cells by injected HepG2 polyA⁺ RNA.

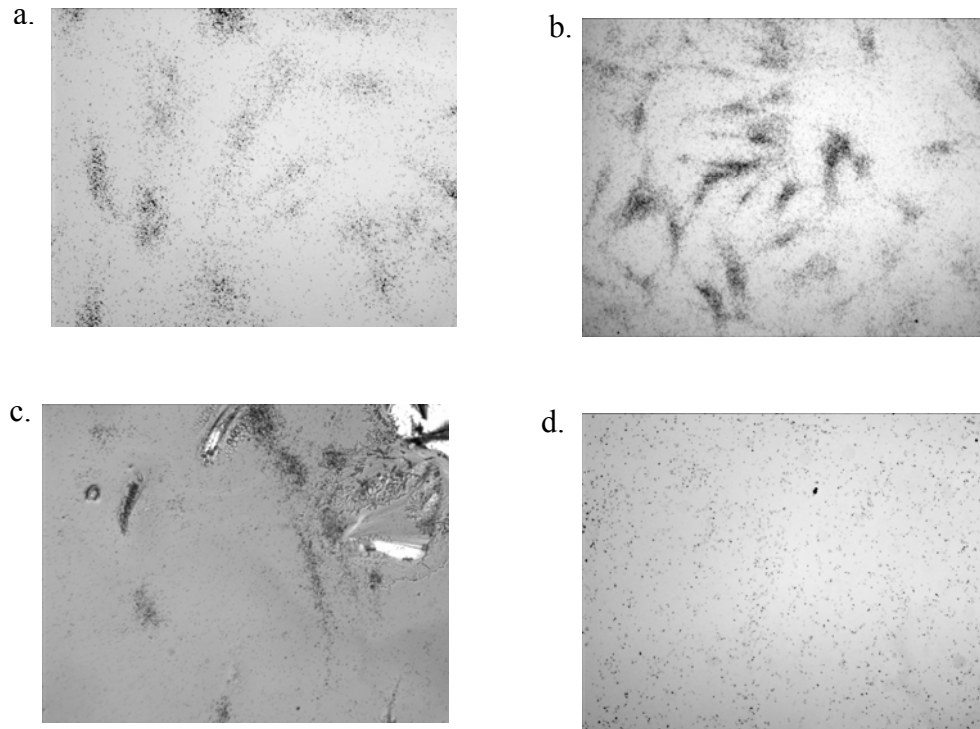


Figure 2.4 A *cblB* clone isolated from the library: pEXp1_MMAB

a. *cblB* cells nucleofected with pEXp1_MMAB **b.** *cblB* cells microinjected by polyA(+) HepG2 RNA (Positive control for injection and [14 C]-propionate incorporation) **c.** *cblB* cells microinjected with pEXp1_MMAB **d.** *cblB* cells (Negative control)
This figure shows the restoration of [14 C]-propionate incorporation of *cblB* mutant cells by nucleofecting or injecting the pEXp1-MMAB clone.

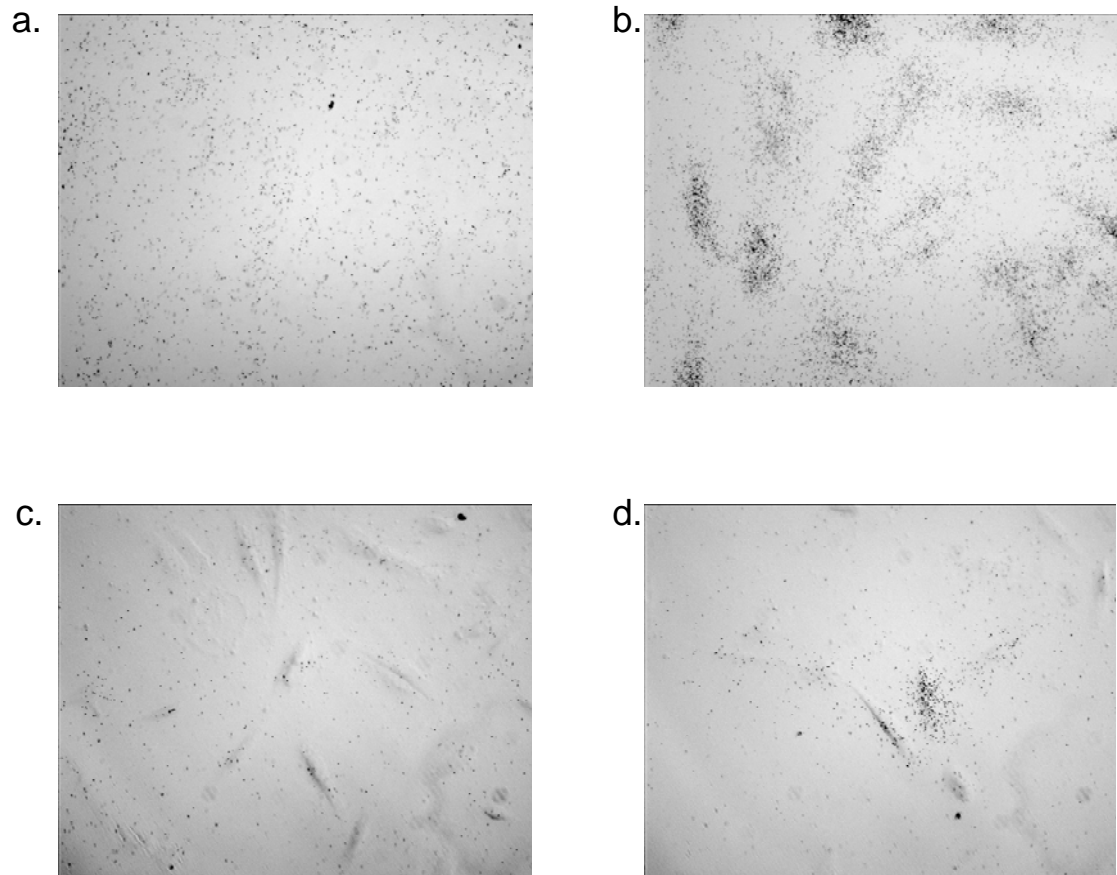


Figure 2.5 Screening the Clontech library

a. *cblB* control cells (negative control for $[^{14}\text{C}]$ -propionate incorporation) **b.** *cblB* cells nucleofected by pEXp1-MMAB (positive control for $[^{14}\text{C}]$ -propionate incorporation) **c.** *cblH* control cells nucleofected by a GFP expression vector (Positive control for nucleofection and negative control for $[^{14}\text{C}]$ -propionate incorporation). **d.** *cblH* cells nucleofected by sub-pool #12; **d.** shows the only positive cells observed on the slides examined. Note the cluster of 4 cells with higher intensity of silver grains over them compared to the control (**c.**).

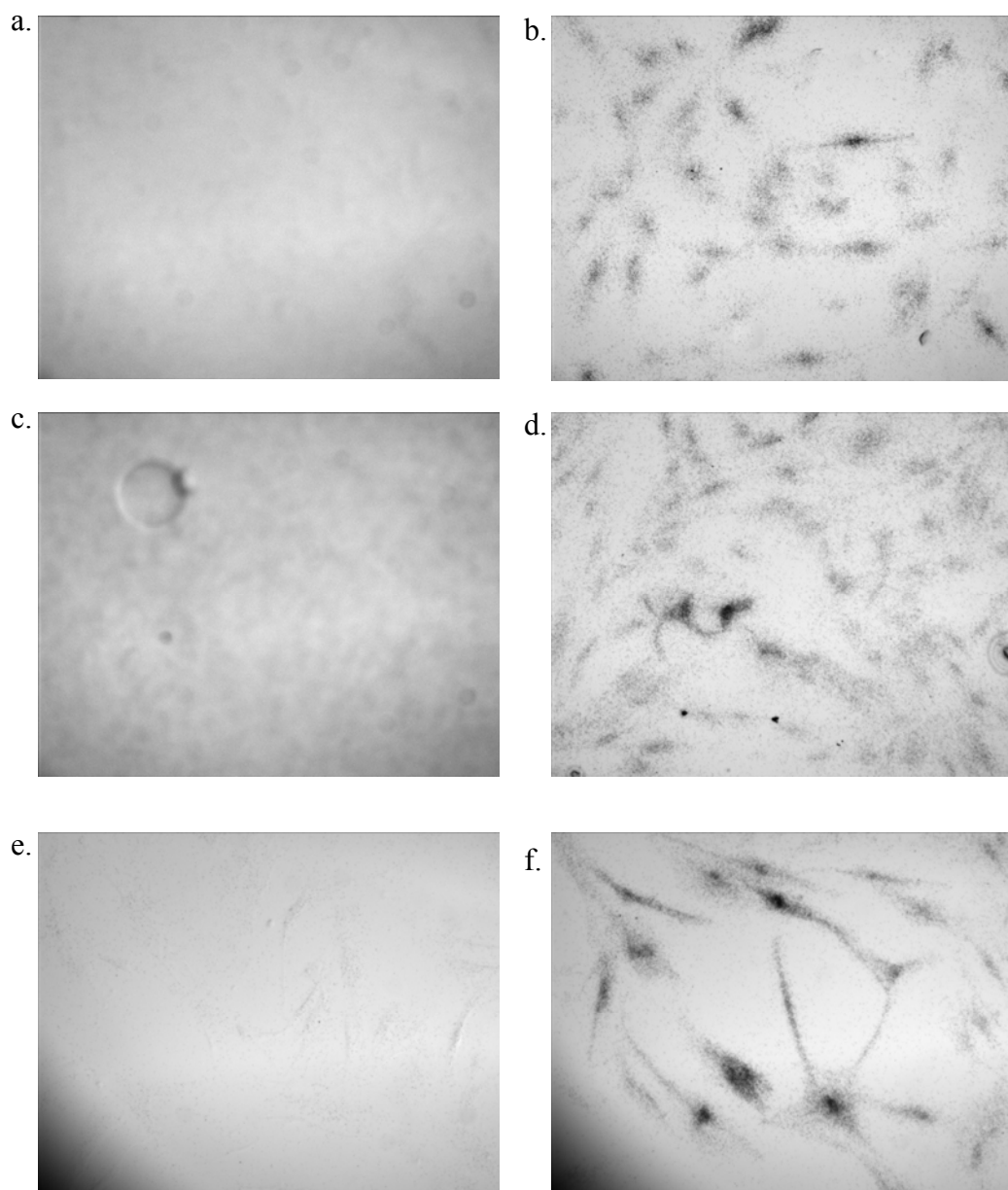


Figure 2.6 Test of the cDNA clones from OriGene

a. *cbIB* control cells **b.** *cbIB* nucleofected with the *cbIB* cDNA clone from OriGene

c. *mut* control cells **d.** *mut* cells nucleofected with the *mut* cDNA clone from OriGene

e. *cbIA* control cells **f.** *cbIA* cells nucleofected with the *cbIA* cDNA clone from OriGene

This figure shows the restoration of [¹⁴C]-propionate incorporation in these cells by complementing them with corresponding expressible cDNA clones purchased from OriGene.

Chapter Three: Identification of gene candidates by a comparative genomics approach

3.1 Introduction

Vitamin B₁₂ (cobalamin, Cbl) utilization in humans is required for two enzymes. One is the methylmalonyl-CoA mutase (MCM, EC 5.4.99.2), a mitochondrial enzyme that catalyzes the isomerization of methylmalonyl-CoA to succinyl-CoA. This enzyme requires a vitamin B₁₂ derivative, adenosylcobalamin (AdoCbl) for activity. The other is the 5'-methyltetrahydrofolate-homocysteine S-methyltransferase (Methionine synthase, MS, EC 2.1.1.13), a cytoplasmic enzyme that catalyzes the remethylation of homocysteine to form methionine. This enzyme requires another cobalamin derivative, methylcobalamin to function. The intracellular transport and modification of Cbl to form AdoCbl and MeCbl is not completely understood. However, genes involved in B₁₂ utilization are evident by complementation groups (*cblA-H* and *mut*). Genes for *cblA*, *cblB*, *cblC*, *cblE*, *cblG*, and *mut* have been identified previously (refer to chapter one for detailed review of these genes).

In this chapter, I will describe a bioinformatics approach to generate a list of candidate genes for *cblD*, *cblH*, and *cblF* (refer to chapter one for detailed review of these complementation groups), and discuss ways to narrow down the list of genes and methods to test individual candidates. Historically, the genes for *cblA* and *cblB* were identified by Melissa Dobson in our lab through a bioinformatics approach ((38,39); See chapter one for detailed review). She made use of the NCBI COG (cluster of orthologous groups) database to examine the genes in close proximity with *MUT* in prokaryotic genomes. Two genes were found in the same operon with *MUT* in *Archaeoglobus*

fulgidus and their human homologs turned out to be the genes for *cblA* and *cblB*. This method could not be extended to search eukaryotic genomes because their genes are not usually organized in the form of operons, where genes in the same pathway are close to each other. However, a genome subtraction approach can be used across eukaryotic genomes to generate a B₁₂ gene list.

The approach is based on the hypothesis that the unidentified B₁₂ processing genes (i.e. *cblD*, *cblH*, and *cblF*) are present in the genomes of all B₁₂ users but not in the non-B₁₂ users. Select groups of eukaryotes (B₁₂ users) require vitamin B₁₂ for the activity of their B₁₂ dependent enzymes (MCM and MS) while others (Non-B₁₂ users) do not have vitamin B₁₂ dependent enzymes and thus do not use vitamin B₁₂ (13,67,67). The known genes (i.e. *MMAA*, *MMAB*, *MMACHC*, *MTR*, *MTRR* and *MUT* for *cblA*, *cblB*, *cblC*, *cblE*, *cblG*, and *mut*, respectively) for the intracellular utilization of vitamin B₁₂ are conserved among B₁₂ users such as human, mouse, and worms (68,69). However, these genes are not found in non- B₁₂ users such as fruit fly and yeasts. Thus, by comparing the genomes of B₁₂ users and non-B₁₂ users, a list of genes that are shared between B₁₂ users but which are not in the genomes of non-B₁₂ users can be generated and this list should include all the B₁₂ genes. With help of other information, the list of genes can be further narrowed down and good candidates can be tested.

The conversion from methylmalonyl-CoA to succinyl-CoA by MCM in B₁₂ users is the last step in the metabolism of propionate prior to entry into the Krebs cycle. Propionate is the catabolism product of odd chain fatty acids, branched chain amino acids, methionine, threonine, and side chain of cholesterol. In non-B₁₂ users such as insects, it is thought to be converted into acetate via 3-hydroxypropionate. In B₁₂ users,

propionate is converted to propionyl-CoA which is carboxylated to form D-methylmalonyl-CoA by the propionyl-CoA carboxylase (PCC), isomerized to L-methylmalonyl-CoA by D-methylmalonyl-CoA epimerase (MCE) and rearranged to form succinyl-CoA by MCM (refer to figure 2.1 in Chapter Two). It seems the AdoCbl dependent MCM is an integral part of this pathway and therefore the genes encoding PCC and MCE (i.e. *PCCA*, *PCCB* coding for the α and β subunits of PCC and *MCEE* coding for MCE) are classified as B₁₂ genes in this study.

Eukaryotic B₁₂ users and non-users are determined experimentally by examining vitamin B₁₂ levels, propionate metabolism into succinate or B₁₂-dependent growth (13,67). With the complete genome information for many eukaryotic species, users and non-users can be identified using a bioinformatics approach by examining the existence of genes for B₁₂ utilization (i.e. *MMAA*, *MMAB*, *MMACHC*, *MTR*, *MTRR*, *MUT*, *PCCA*, *PCCB*, and *MCEE*). Therefore, to identify previously unknown B₁₂ users and non-users for this study, B₁₂ genes were examined for existence in complete eukaryotic genomes available through NCBI in this study. Organisms classified as B₁₂ users and non-B₁₂ users could then be used in the genome subtraction procedure.

The genome comparison among B₁₂ users and non-B₁₂ users is based on homolog identification among species. In this study, the question of whether a gene is shared by B₁₂ users but non-B₁₂ users can be rephrased as whether an ortholog of this gene is identified in B₁₂ users but not in non-B₁₂ users. Two independent homolog (orthologs and paralogs) databases are used in this study, HomoloGene and InParanoid. HomoloGene (70) (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=homologene>) is a system for automatic detection of homologs among annotated genes of 18 completely sequenced

eukaryotic genomes. Homologs of each gene among species are clustered in a group (A HomoloGene group). InParanoid eukaryotic ortholog database (71,72) (<http://InParanoid.sbc.su.se/>) is a collection of pairwise comparisons between 26 completely sequenced eukaryotic genomes. Homologs between two species are clustered in a group (InParanoid cluster).

3.2 Methods

3.2.1 BLASTP

The sequence alignment tool at the NCBI BLAST web site was used to survey B₁₂ genes across eukaryotic genomes (<http://www.ncbi.nlm.nih.gov/BLAST/>). The protein sequence corresponding to each of these genes for human is used as a query in the BLASTP interface (73) (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=euk) to do sequence similarity searches in the complete protein database(s) of an individual species (e.g. *Danio rerio* (zebrafish)) or a group of species (e.g. Fungi). Default values are used for optional parameters. A significant hit (usually with an expect value far less than 0.001 and a bit score more than 100) indicates existence of a homologous gene in the genome of the tested species.

3.2.2 Comparison of selected eukaryotic genomes

3.2.2.1 Generate candidate gene list using HomoloGene

For the HomoloGene comparison, human (*Homo sapiens*) and a nematode (*C. elegans*) are chosen as the B₁₂ users while a protozoan (*P.falciparum*), two plants (*A.thaliana* and *O.sativa*), two insects (*A.gambiae* and *D.melanogaster*), a group of fungi (*M.grisea*, *S.pombe*, *S.cerevisiae*, *E.gossypii* *N.crassa*, *K. Lactis*) are chosen as the non-

B₁₂ users. A HomoloGene group list is generated at the web interface using a Boolean query in the “organism” column that combines (using “AND” or “NOT”) the gene sets of each species coded by NCBI taxonomy ID (Box 3.1). The returned result is a list of HomoloGene groups. Each group represents a gene and the list of groups therefore represents the candidate gene list *per se*.

```
txid9606 [Organism] AND txid6239 [Organism] NOT (txid148305
[Organism] OR txid33169 [Organism] OR txid3702 [Organism] OR
txid4530 [Organism] OR txid4896 [Organism] OR txid4932 [Organism]
OR txid5141 [Organism] OR txid7165 [Organism] OR txid7227
[Organism] OR txid28985 [Organism])
```

Green=B₁₂ users; Red= non-B₁₂ users

Box 3.1 The Boolean query used in HomoloGene WWW site

3.2.2.2 Generate candidate gene list using InParanoid

The procedure for comparison based on the InParanoid database is done differently because of the pairwise format of the data (figure 3.1). In other words, the comparison offered by InParanoid database is between two species for each analysis. Human (*Homo sapiens*), a nematode (*C. elegans*), and *Dictyostelium discoideum* are chosen as B₁₂ users while *D.melanogaster*, *S.cerevisiae*, *S.pombe*, *O.sativa* and *A.thaliana* are chosen as non-B₁₂ users. To build the B₁₂ candidate gene list, 7 sets of InParanoid clusters are downloaded (figure 3.1 shows two example clusters of the *Homo sapiens* and *C. elegans* set). These sets are *Homo sapiens* and *C. elegans* (4588 clusters),

Homo sapiens and *Dictyostelium discoideum* (3111 clusters), *Homo sapiens* and *D. melanogaster* (5534 clusters), *Homo sapiens* and *S. cerevisiae* (2137 clusters), *Homo sapiens* and *S. pombe* (2400 clusters), *Homo sapiens* and *Arabidopsis thaliana* (3231 clusters) and *Homo sapiens* and *Oryza sativa* (2774 clusters). For each set, data are imported into a Microsoft Excel spread sheet where data are processed in such way that only the lines that contain Ensembl proteins IDs for human entries are retained. Each set now represents a list of human proteins that have homologs in *C. elegans*, *D. discoideum*, *D. melanogaster*, *S. pombe*, *S. cerevisiae*, *A. thaliana*, and *O. sativa*, respectively.

Proteins IDs that are in lists for both species of B₁₂ users (i.e. *C. elegans* and *D. discoideum*) but not in lists of non-B₁₂ users (i.e. *D. melanogaster*, *S. pombe*, *S. cerevisiae*, *A. thaliana*, and *O. sativa*) are extracted to form a new list. This list contains human proteins IDs that have homologs in *C. elegans* and *D. discoideum* but not in *D. melanogaster*, *S. pombe*, *S. cerevisiae*, *A. thaliana*, and *O. sativa*. A corresponding Ensembl gene ID list is generated via Ensembl BioMart (Dec 2005 Ensembl release 36, NCBI *Homo sapiens* build 35, <http://dec2005.archive.ensembl.org/Multi/martview>).

Using the retrieved gene list, an up-to-date gene list with annotation is made via current version of Ensembl BioMart (June 2007, Ensembl release 45, NCBI *Homo sapiens* build 36, (<http://www.ensembl.org/biomart/martview>)).

Inparalog and Orthologs clusters for *Homo sapiens* (Ensembl) and *Caenorhabditis elegans* (Wormbase) (4588 clusters in total)

Cluster #1				
Protein ID	Species	Score	Bootstrap	Name
ENSP00000351750	H.sapiens	1	100%	Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain 1) (DHC1). [Source:Uniprot/SWISSPROT;Acc:Q14204]
WBGene00000962	C.elegans	1	100%	

Cluster #2				
Protein ID	Species	Score	Bootstrap	Name
ENSP00000304350	H.sapiens	1	100%	Pre-mRNA processing splicing factor 8 (Splicing factor Pp8) (PRP8 homolog) (220 kDa U5 snRNP-specific protein) (p220). [Source:Uniprot/SWISSPROT;Acc:Q6P2Q9]
WBGene00004187	C.elegans	1	100%	

Cluster #3				
Protein ID	Species	Score	Bootstrap	Name
ENSP00000353174	H.sapiens	1	100%	Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RyR2) (RYR-2) (Cardiac muscle ryanodine receptor-calcium release channel) (hRyR-2). [Source:Uniprot/SWISSPROT;Acc:Q92736]
ENSP00000354852	H.sapiens	0.5		Ryanodine receptor 3 (Brain-type ryanodine receptor) (RyR3) (RYR-3) (Brain ryanodine receptor-calcium release channel). [Source:Uniprot/SWISSPROT;Acc:Q15413]
ENSP00000352608	H.sapiens	0.456		Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RyR1) (RYR-1) (Skeletal muscle calcium release channel). [Source:Uniprot/SWISSPROT;Acc:P21817]
WBGene00006801	C.elegans	1	100%	

Figure 3.1 Example clusters of the data set for *Homo sapiens* (Ensembl) and *Caenorhabditis elegans* (Wormbase)

Three clusters are shown here. 4588 clusters in total are in this set. Each cluster (e.g. Cluster #3) consists of a seed-paralog pair (the scores which reflect the distance between inparalogs, usually equal or very close to 1 for them) and their inparalogs (The scores usually less than 1 for them). Bootstrap values reflect the confidence that the original seed-ortholog pair is true ortholog and usually equal or are close to 100%. When the data are processed in excel, only seed-paralog pairs with a score no less than 99% are retained. For these pairs, human protein IDs are extracted in one column (i.e. “ENSP00000351750”, ENSP00000304350” and ENSP00000354852 etc.). The genome comparison is based on these human protein IDs.

3.3 Results and discussion

3.3.1 Survey of vitamin B₁₂ utilization in eukaryotic genomes

B₁₂ genes were surveyed across 34 complete eukaryotic genomes available from NCBI (as summarized in table 3.1). The results agreed with the commonly accepted theory (67) that insects, fungi, and plants do not use B₁₂ and therefore have no B₁₂ genes. Unreported previously, B₁₂ genes were found in all three species of protozoa surveyed (*Leishmania braziliensis*, *Leishmania infantum* JPCM5, and *Dictyostelium discoideum* AX4). B₁₂ genes were also found in the only echinoderm surveyed (*Strongylocentrotus purpuratus* (sea urchin)). However, the list of species was far from complete and more eukaryotic B₁₂ users could be found if genome data for more species are available. For example, it has been reported that some algae use B₁₂ (74,75) and have the B₁₂ dependent methionine synthase in their genomes (76). Recently, an early divergent fungus (*Blastocladiella emersonii*) was found to harbour B₁₂ genes such as *MUT*, *MMAB*, *PCCA*, *PCCB*, *MCEE* in the uncompleted genome (77). Though not surveyed in this study, the protozoa *Euglena gracilis* had been identified as B₁₂ users and served as an organism for B₁₂ bioassay several decades ago (78-81). Vitamin B₁₂ synthesis (10) and utilization based on literature and findings in this study are summarized in figure 3.2.

It is noticeable that multiple significant hits were found in both B₁₂ users and non-B₁₂ users for *PCCA*, *PCCB*, and *MTRR* (table 3.1). This resulted from the presence of corresponding paralogs across all surveyed genomes. *PCCA/B* encodes propionyl-CoA carboxylase which has other paralogous carboxylases such as acetyl-CoA carboxylase, pyruvate carboxylase and methylcrotonoyl-CoA carboxylase while *MTRR* encodes methionine synthase reductase, of which paralogs such as nitric oxide synthase 1,

NADPH-cytochrome-P450 oxidoreductase and sulfite reductase exist across the genomes surveyed.

The genes for transcobalamin I, II (*TCI* and *TCII*) and intrinsic factor (IF) were also surveyed (table 3.1, Last column). The genes for transcobalamin I, II, and intrinsic factor have originated by gene duplication (82) and they are important B₁₂ carrier proteins for its absorption and transport in humans. The genes for them were only found conserved in vertebrates in this study, though it was hard to conclude all of the three genes were present in some species because it was hard to distinguish them using BLASTP or by annotation. However, there was definitely no indication of their existence in other B₁₂ users (i.e. sea urchin, worms, and the protozoa) and non-B₁₂ users. This suggests that the other B₁₂ users must have adopted different means to acquire their vitamin B₁₂ from their environment.

Finally, the genes encoding cubilin and amnionless (*CUBN* and *AMN*), the intrinsic factor-cobalamin receptor complex were surveyed (Data not shown in table 3.1). For cubilin, a protein with over 3500 amino acids, full-length hits were found in all multicellular organisms, including all vertebrates, echinoderm (*Strongylocentrotus purpuratus* (sea urchin), insects (fruit fly and honey bee), and the nematode (*C. elegans*). No significant hits were found in fungi, plants, and protozoa. In addition, the N terminus of cubilin was found to be similar to the N termini of Notch and Jagged proteins, while the C terminus was found to be similar to some Cub-domain containing proteins and bone-morphogenic proteins. For amnionless, a smaller protein with about 450 amino acids, significant hits were found in vertebrates and sea urchin but not in other eukaryotes (“weak” hits in fruit fly, i.e. hits with low expect values but bit scores less than 100).

Compared to the cobalamin carrier gene (i.e. *IF*), the broader distribution of the receptor genes suggests that this receptor complex has functions other than transporting B₁₂.

Taxa		Genomes	Genes									
			<i>MTR</i>	<i>MTRR^a</i>	<i>MUT</i>	<i>MMAA</i>	<i>MMAB</i>	<i>MMACHC</i>	<i>PCCA^c</i>	<i>PCCB^c</i>	<i>MCEE</i>	<i>TCI, II, or IF</i>
Vertebrates (11)	Mammals(9)	Bos taurus (cow) Build 3.1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Canis familiaris (dog)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Homo sapiens (human) Build 36	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Macaca mulatta (rhesus macaque)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Monodelphis domestica (gray short-tailed opossum)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Mus musculus (mouse) Build 37	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Ornithorhynchus anatinus (platypus)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Pan troglodytes (chimpanzee)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		Rattus norvegicus (rat)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Vertebrate s(2)	Danio rerio (zebrafish)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

		Gallus gallus (chicken)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Invertebrates (4)	Insects(2)	Drosophila melanogaster	No	Yes	No	No	No	No	Yes	Yes	No	No
		Apis mellifera (honey bee) Amel_4.0	No	Yes	No	No	No	No	Yes	Yes	No	No
	Nematode(1)	Caenorhabditis elegans	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
	Echinoderm(1)	Strongylocentrotus purpuratus (sea urchin) Build 2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Fungi (15)	Ascomycota(12)	Aspergillus fumigatus Af293	No	Yes	No	No	No	No	No	Yes	Yes	No	No
		Aspergillus terreus NIH2624	No	No	No	No	No	No	No	Yes	Yes	No	No
		Magnaporthe grisea 70-15	No	No	No	No	No	No	No	Yes	Yes	No	No
		Candida albicans SC5314	No	No	No	No	No	No	No	Yes	No	No	No
		Candida glabrata CBS 138	No	Yes	No	No	No	No	No	No	No	No	No

		Debaryomyces hansenii CBS767	No	Yes	No	No	No	No	No	Yes	No	No	No
		Eremothecium gossypii	No	No	No	No	No	No	No	Yes	No	No	No
		Kluyveromyces lactis NRRL Y-1140	No	Yes	No	No	No	No	No	Yes	No	No	No
		Pichia stipitis CBS 6054	No	Yes	No	No	No	No	No	Yes	No	No	No
		Saccharomyces cerevisiae	No	Yes	No	No	No	No	No	Yes	No	No	No

		Yarrowia lipolytica CLIB122	No	Yes	No	No	No	No	No	Yes	Yes	No	No
		Schizosaccharomyces pombe 972h-	No	Yes	No	No	No	No	No	Yes	No	No	No
	Basidiomycota (2)	Cryptococcus neoformans var. neoformans B-3501A	No	Yes	No	No	No	No	No	Yes	Yes	No	No
		Cryptococcus neoformans var. neoformans JEC21	No	Yes	No	No	No	No	No	Yes	Yes	No	No
	Microsporidia (1)	Encephalitozoon cuniculi GB-M1	No	No	No	No	No	No	No	No	No	No	No

<i>Plants (2)</i>	Monocotyledon (1)	Arabidopsis thaliana (thale cress)	No	Yes	No	No	No	No	No	No	No	No	No
		Oryza sativa (rice)	No	Yes	No	No	No	No	No	No	No	No	No
<i>Protozoa (3)</i>	Kinetoplastida (2)	Leishmania braziliensis	Yes	Yes	Yes	Yes	Yesb	No	Yes	Yes	Yes	Yes	No
		Leishmania infantum JPCM5	Yes	Yes	Yes	Yes	Yesb	No	Yes	Yes	Yes	Yes	No
	Others (1)	Dictyostelium discoideum AX4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

Table 3.1 Survey of B₁₂ genes, *TC1*, *TC2* and *IF* across eukaryotic genomes

Protein queries used in search are MTR: NP_000245, MTRR: NP_002445, MUT: NP_000246, MMACHC: NP_056321, MMAA: NP_758454 MMAB: NP_443077, PCC α : NP_000273, PCC β : NP_000523, MCEE: NP_115990, TC2 (Transcobalamin): NP_000346, TC1 (Haptocorrin): NP_001053, IF (Intrinsic factor): NP_005133. “Yes” means a significant hit is found for a given gene in the corresponding genome; “No” means no significant hit is found for a given gene in the corresponding genome. Note: a. “Yes” doesn’t necessarily mean MTRR in that genome because it has paralogs such as nitric oxide synthase 1, NADPH-cytochrome-P450 oxidoreductase and Sulfite reductase across the genomes surveyed; b. these hits have reasonable low expect values with bit scores lower than 100; c. “Yes” doesn’t necessarily mean PCC α and PCC β in that genome because they have other carboxylases such as Acetyl-CoA carboxylase, pyruvate carboxylase and methylcrotonoyl-Coenzyme A carboxylase 1 that is hard to be distinguished from each other using BLASTP.

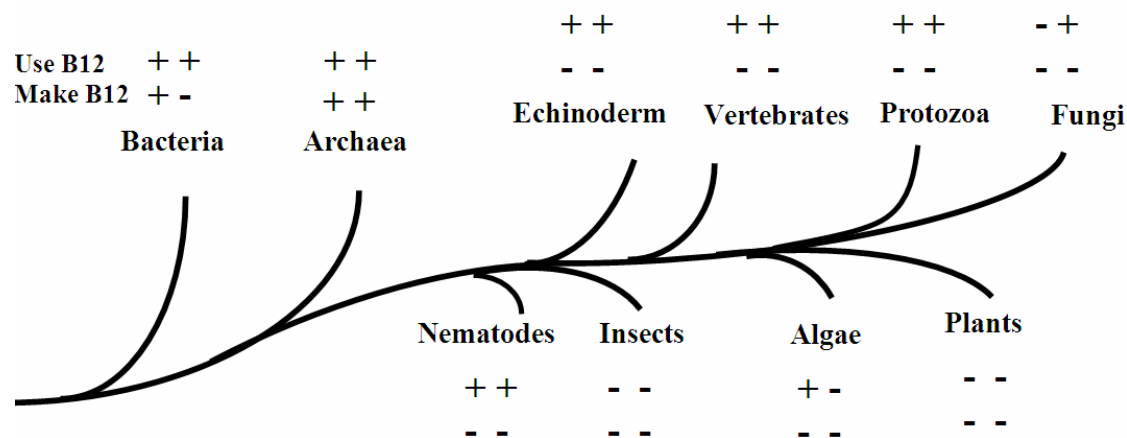


Figure 3.2 Synthesis and utilization of Vitamin B₁₂

The B₁₂ synthesis and utilization across living organisms are summarized based on previous literature (10,67,83) and the findings in this study. Prokaryotes are both users and makers while eukaryotes lost the ability to synthesis B₁₂. Echinoderm, vertebrates,

nematodes, and protozoa surveyed are using B₁₂. Insects, fungi, and plants are not using B₁₂, except that an early divergent aquatic fungus was reported to harbor B₁₂ gene in its genome. Some algae use B₁₂ and some do not.

3.3.2 Comparison of selected eukaryotic genomes

In the HomoloGene based genome comparison (summarized in figure 3.3), Human (19,493 HomoloGene groups) and *C. elegans* (*C. elegans* 4,973 HomoloGene groups) were chosen as B₁₂ users. Though available, other vertebrate genomes were not chosen for comparison because they are too close to humans, therefore not only would using these genomes offer little subtraction but also could undermine the quality of results if they were not annotated as well as the human genome. In contrast, all the non-B₁₂ users (25,412 HomoloGene groups in total) available were used in the comparison to maximize the subtraction. There were 6,687 HomoloGene groups shared by Human and the non-B₁₂ users, 3,953 HomoloGene groups shared by Human and *C. elegans* and 4,135 HomoloGene groups shared by the non-B₁₂ users and *C. elegans*.

The candidate gene list included 584 HomoloGene groups that were shared by Human and *C. elegans* but not with non-B₁₂ users. It represented less than 3% of the human genome. HomoloGene groups representing *PCCB* (#447), *MCEE* (#13078), *MUT* (#20097), *MMAA* (#14586), *MMAB* (#12680), *MMACHC* (#12082), and *MTRR* (#11419) were found in this list (The numbers in the parentheses correspond to HomoloGene IDs). Missing HomoloGene groups that had been predicted in the list (i.e. those of *PCCA* and *MTR*) were examined. *PCCA* (#236) was not in the list because its group contained several fungi members. This was not a surprise as discussed in the previous section that *PCCA* could be confused with other carboxylases when compared by BLASTP. *MTR* was subtracted from the list due to the existence of a gene member from *A.gambiae* in that group. The record for this gene was discontinued in the more recent build of database and therefore it could be just a wrongly annotated gene. On the contrary, *MTRR* and

PCCB, which also have paralogs with high degree of sequence similarities, ended up in the list. This implies that though based on BLASTP search, the building procedure used by HomoloGene provides a certain level of sophistication to distinguish paralogs. Overall, the existence of most of (if not all) B₁₂ genes in this list suggests that it is very likely that the genes for *cblF*, *cblH*, and *cblD* are also in this list. The subtraction is so effective that this list represents only about 3% (584/19493) of the human genome.

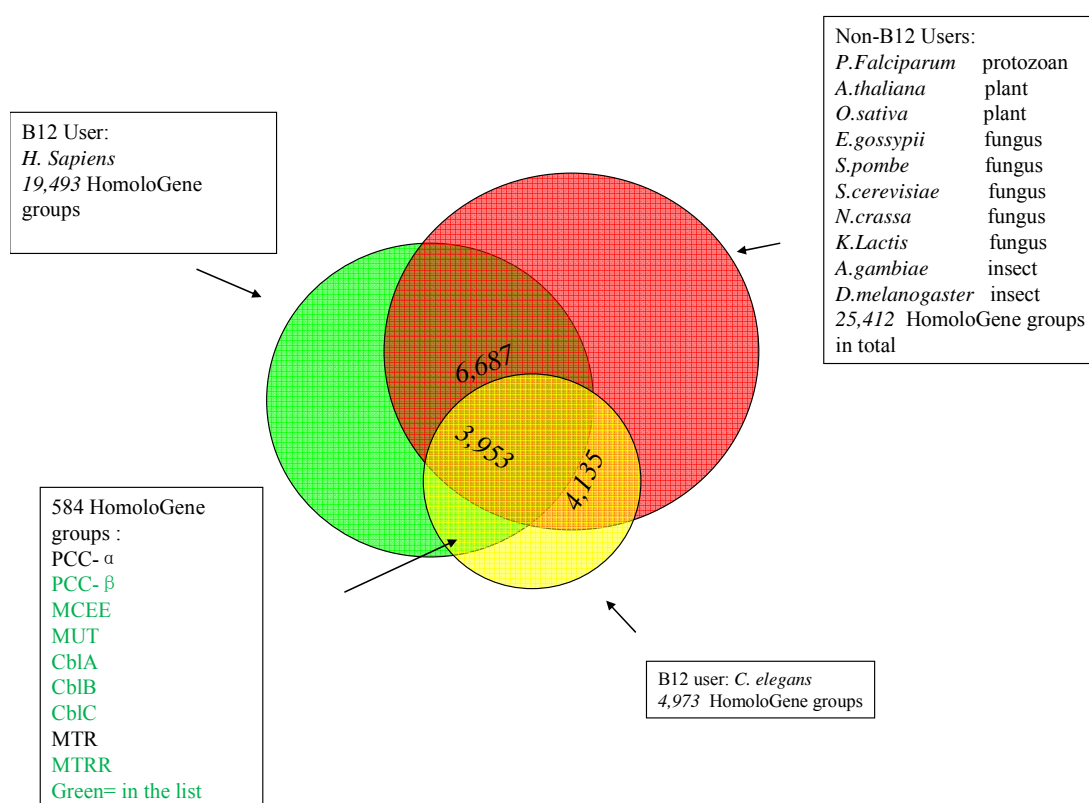


Figure 3.3 Comparison based on HomoloGene.

Numbers within the Venn diagram intersects are the number of HomoloGene groups in common in the intersect spaces. 6,687 HomoloGene groups shared by Human and the non-B12 users, 3,953 HomoloGene groups shared by Human and *C. elegans* and 4,135

HomoloGene groups shared by the non-B12 users and *C. elegans*. The candidate gene list contains 584 Homologene groups with 7 of 9 B12 genes included.

The comparison based on InParanoid was aimed at achieving further subtraction by adding another newly found B₁₂ user, *Dictyostelium discoideum*, to the analysis (Refer to section 3.2.2.2 for detailed procedure). Indeed, a shorter list with 259 genes was generated, which contains all the B₁₂ genes except for MMACHC, MTR, and MTRR (figure 3.4, table A.2). In the perfect world, the HomoloGene list should include the InParanoid list. In the 259 genes for which a corresponding HomoloGene group could be identified, there were 50 genes in common in the two lists (table 3.2). They included five B₁₂ genes (*MMAA MMAB MUT PCCB* and *MCEE*). Considering the different data sources (NCBI versus Ensembl), different collections of genomes, and different building procedures used to generate these two lists, the result was satisfactory. It would be hard to tell which list was better. It was decided to consider the candidates from both lists to avoid the possibility of missing out on candidate target genes, though the second list was originally intended to help shorten the first list.

Note: For complete gene list generated using HomoloGene, please go to the WWW site (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=homologene>) and search using the Boolean query in Box 3.1. Alternatively, a corresponding human gene list with annotations is shown as table A.1 in the appendix. For a complete gene list generated using InParanoid, please refer to table A.2. in the appendix. Based on description, about 50% of entries are known genes; about 30% are genes with predicted families or domains; only about 20% are completely unknown genes. Based on annotation, about

30% of corresponding proteins for these genes have at least one transmembrane domain and about 20% of corresponding proteins for these genes have at least one signal domain.

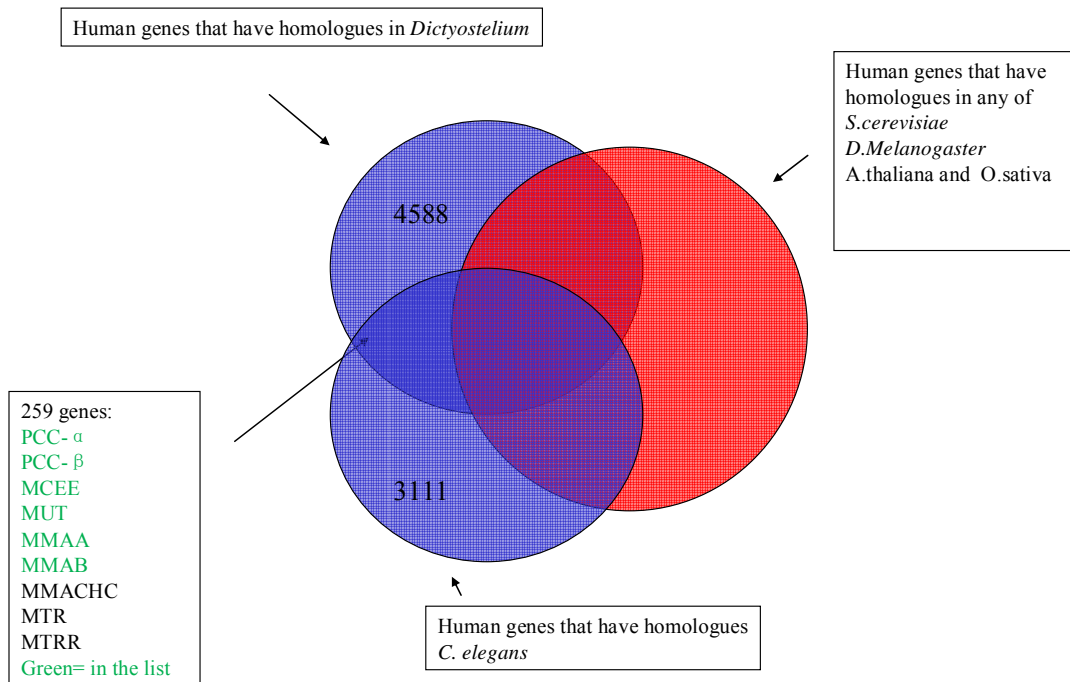


Figure 3.4 Comparison based on InParanoid

Numbers within the Venn diagram intersects are the number of human genes that have homologs in *Dictyostelium* and *C.elegans*, respectively. The area overlapped with the two blue sets but not with the red set represents the list of 259 B₁₂ candidate genes.

3.4 Future directions

During the course of this study, the gene for *cblD/cblH* was identified by another research group. It was shown that a single gene corresponds to both groups (unpublished data). It is encouraging to notice that *cblD/cblH* gene is in both lists and is even in the list of 50 (table 3.2). This further proves the efficacy of this approach and makes it encouraging that the gene for *cblF* may also be present.

The gene for *cblF* is likely to encode a lysosomal membrane B₁₂ transporter. In order to narrow the list of candidates, first, the annotations of the genes in the derived should be examined and genes with known functions would be ruled out. The remaining genes should be classified into categories with different priorities. Genes with complete unknown functions or with characteristics such as membrane spanning domains, lysosomal targeting, B₁₂ binding motif etc. should be put into high priority. Lysosomal databases (84) can be used to help in this process.

The approach to test candidate genes will include obtaining the Origene cDNA for transfection into *cblF* cells and use of an RNAi procedure to block propionate incorporation in *C. elegans*. For each individual candidate gene, loss-of-function test can be used by monitoring [¹⁴C]-propionate incorporation levels of *C. elegans* when its expression is silenced by RNAi ((68), and unpublished data). In addition, a large number of genes have been knocked out in *C. elegans* are available through the Gene Knockout Consortium (<http://celeganskoconsortium.omrf.org/>). Thus, knockout strains, if available, can be evaluated. Alternatively, using human cells, complementation tests can be done in *cblF* mutant cells if expressible full-length clones are available for purchase from

OriGene. Sequencing the gene in *cbIF* mutant cells for mutations can also be done and will provide the ultimate confirmation of the gene identity.

The finding that *Dictyostelium* is an apparent B₁₂ user might lead us to establish an excellent model organism to study vitamin B₁₂ metabolism in eukaryotes. The beauty of this creature as a model organism is that it has a fully sequenced haploid genome and most of the molecular genetic techniques typically associated with *S. cerevisiae* are also available (i.e. Targeted gene disruption using homologous recombination and random insertional mutagenesis). *C. elegans* has been established as a model organism by our lab and others. To take advantage of the uniqueness of *Dictyostelium* and establish it as a new model organism would extend our capacity to study B₁₂ metabolism.

The loss-of-function screen using model organisms will not only help us to identify the B₁₂ genes represented by human genetic diseases but also be equally powerful to identify B₁₂ genes that are not represented by human disease. Meanwhile, the candidate gene lists generated in this study theoretically include all B₁₂ genes regardless of whether they are manifested in human disease or not. A screen of the candidate genes using *C. elegans* or *Dictyostelium* would be very promising to find more B₁₂ genes.

In this study, only 34 completed eukaryotic genomes available through NCBI were surveyed for B₁₂ genes. There are currently 65 completed genomes for eukaryotes (<http://genomesonline.org/gold.cgi?want=Published+Complete+Genomes&pubsort=Domain>). These include completed genomes for a few protists and non-insects arthropods, which were not covered in this study. For completeness, they should be surveyed for B₁₂ genes as well. In addition, a paper published recently reported that the adenosylcobalamin dependent class II ribonucleotide reductase was present in the protozoan *Euglena gracilis*

as well as in *Dictyostelium Discoideum* (81). This is an exceptional finding because the enzyme is usually present only in prokaryotes. It should be included as a B₁₂ gene in future surveys.

Finally, the comparative genome approach is shown to be very effective in generating B₁₂ gene candidates. Whether or not this approach can be applied to other pathways is unclear at this stage. Vitamin B₁₂ is an ancient molecule and this is the reason it can be found so widely distributed in prokaryotes and eukaryotes. Nevertheless, it is worth keeping this concept in mind in the future gene searches.

Table 3.2 The 50 genes shared by the two lists

B₁₂ genes were labelled with green blocks. Gene for *cbID* was labelled with a yellow block. Chr = Chromosome; TM= Transmembrane Domain; SD= Singal Domain

Ensembl Gene ID	Description	Chr	TD	SD
ENSG00000008256	Cytohesin-3 (PH, SEC7 and coiled-coil domain-containing protein 3) (ARF nucleotide-binding site opener 3) (Protein ARNO3) (General receptor of phosphoinositides 1) (Grp1). [Source:Uniprot/SWISSPROT;Acc:O43739]	7		
ENSG00000008015	Presenilin-1 (EC 3.4.23.-) (PS-1) (Protein S182) [Contains: Presenilin-1 NTF subunit; Presenilin-1 CTF subunit; Presenilin-1 CTF12 (PS1-CTF12)]. [Source:Uniprot/SWISSPROT;Acc:P49768]	14	1	
ENSG000000084110	Histidine ammonia-lyase (EC 4.3.1.3) (Histidase). [Source:Uniprot/SWISSPROT;Acc:P42357]	12		
ENSG00000008836	Sodium bicarbonate transporter-like protein 11 (Sodium borate cotransporter 1) (NaBC1) (Bicarbonate transporter-related protein 1) (Solute carrier family 4 member 11). [Source:Uniprot/SWISSPROT;Acc:Q8NBS3]	20	1	
ENSG00000100034	Ca(2+)/calmodulin-dependent protein kinase phosphatase (EC 3.1.3.16) (CaM-kinase phosphatase) (CaMKPase) (Partner of PIX 2) (hFEM-2) (Protein phosphatase 1F). [Source:Uniprot/SWISSPROT;Acc:P49593]	22		

ENSG00000101160	Cathepsin Z precursor (EC 3.4.22.-) (Cathepsin X) (Cathepsin P). [Source:Uniprot/SWISSPROT;Acc:Q9UBR2]	20	1	1
ENSG00000101843	26S proteasome non-ATPase regulatory subunit 10 (26S proteasome regulatory subunit p28) (Gankyrin). [Source:Uniprot/SWISSPROT;Acc:O75832]	X		
ENSG00000103642	Serine beta-lactamase-like protein LACTB, mitochondrial precursor (EC 3.4.-.-). [Source:Uniprot/SWISSPROT;Acc:P83111]	15		
ENSG00000104969	Small glutamine-rich tetratricopeptide repeat-containing protein A (Vpu-binding protein) (UBP). [Source:Uniprot/SWISSPROT;Acc:O43765]	19		
ENSG00000105135	ilvB (bacterial acetolactate synthase)-like isoform 1 [Source:RefSeq_peptide;Acc:NP_006835]	19	1	1
ENSG00000107537	Phytanoyl-CoA dioxygenase, peroxisomal precursor (EC 1.14.11.18) (Phytanoyl-CoA alpha-hydroxylase) (PhyH) (Phytanic acid oxidase). [Source:Uniprot/SWISSPROT;Acc:O14832]	10		
ENSG00000108239	TBC1 domain family member 12. [Source:Uniprot/SWISSPROT;Acc:O60347]	10		
ENSG00000108848	Cisplatin resistance-associated overexpressed protein (cAMP regulatory element-associated protein 1) (CRE-associated protein 1) (CREAP-1) (Luc7A) (Okadaic acid-inducible phosphoprotein OA48-18). [Source:Uniprot/SWISSPROT;Acc:O95232]	17		
ENSG00000113552	Glucosamine-6-phosphate isomerase (EC 3.5.99.6) (Glucosamine-6-phosphate deaminase) (GNPDA) (GlcN6P deaminase) (Oscillin). [Source:Uniprot/SWISSPROT;Acc:P46926]	5		

ENSG00000114054	Propionyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit beta) (Propanoyl-CoA:carbon dioxide ligase subunit beta). [Source:Uniprot/SWISSPROT;Acc:P05166]	3		1
ENSG00000115307	Ancient ubiquitous protein 1 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9Y679]	2		
ENSG00000116791	Quinone oxidoreductase (EC 1.6.5.5) (NADPH:quinone reductase) (Zeta- crystallin). [Source:Uniprot/SWISSPROT;Acc:Q08257]	1		
ENSG00000118514	aldehyde dehydrogenase 8A1 isoform 2 [Source:RefSeq_peptide;Acc:NP_739577]	6		
ENSG00000124370	Methylmalonyl-CoA epimerase, mitochondrial precursor (EC 5.1.99.1) (DL-methylmalonyl-CoA racemase). [Source:Uniprot/SWISSPROT;Acc:Q96PE7]	2		
ENSG00000132773	Target of EGR1 protein 1. [Source:Uniprot/SWISSPROT;Acc:Q96GM8]	1		
ENSG00000133243	BTB/POZ domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q9BX70]	19		
ENSG00000134242	Tyrosine-protein phosphatase non-receptor type 22 (EC 3.1.3.48) (Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP) (Lymphoid phosphatase) (LyP). [Source:Uniprot/SWISSPROT;Acc:Q9Y2R2]	1		
ENSG00000134996	Osteoclast-stimulating factor 1. [Source:Uniprot/SWISSPROT;Acc:Q92882]	9		
ENSG00000138760	Lysosome membrane protein 2 (Lysosome membrane protein II) (LIMP II) (Scavenger receptor class B member 2) (85 kDa lysosomal membrane sialoglycoprotein) (LGP85) (CD36 antigen-like 2). [Source:Uniprot/SWISSPROT;Acc:Q14108]	4	1	1
ENSG00000139344	Probable imidazolonepropionase (EC 3.5.2.7) (Amidohydrolase domain- containing protein 1). [Source:Uniprot/SWISSPROT;Acc:Q96NU7]	12		

ENSG00000139428	Cob(I)yrinic acid a,c-diamide adenosyltransferase, mitochondrial precursor (EC 2.5.1.17) (Cob(I)alamin adenosyltransferase) (Methylmalonic aciduria type B protein). [Source:Uniprot/SWISSPROT;Acc:Q96EY8]	12		
ENSG00000142188	Transmembrane protein 50B (HCV p7-transregulated protein 3). [Source:Uniprot/SWISSPROT;Acc:P56557]	21	1	
ENSG00000144959	Arylacetamide deacetylase-like 1 (EC 3.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q6PIU2]	3	1	
ENSG00000146085	Methylmalonyl-CoA mutase, mitochondrial precursor (EC 5.4.99.2) (MCM) (Methylmalonyl-CoA isomerase). [Source:Uniprot/SWISSPROT;Acc:P22033]	6		
ENSG00000147010	SH3 domain-containing kinase-binding protein 1 (Cbl-interacting protein of 85 kDa) (Human Src-family kinase-binding protein 1) (HSB-1) (CD2-binding protein 3) (CD2BP3). [Source:Uniprot/SWISSPROT;Acc:Q96B97]	X		
ENSG00000147251	dedicator of cytokinesis 11 [Source:RefSeq_peptide;Acc:NP_653259]	X		
ENSG00000151498	Acyl-CoA dehydrogenase family member 8, mitochondrial precursor (EC 1.3.99.-) (ACAD-8) (Isobutyryl-CoA dehydrogenase) (Activator- recruited cofactor 42 kDa component) (ARC42). [Source:Uniprot/SWISSPROT;Acc:Q9UKU7]	11		
ENSG00000151611	Methylmalonic aciduria type A protein, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q8IVH4]	4		

ENSG00000154845	Serine/threonine-protein phosphatase 4 regulatory subunit 1. [Source:Uniprot/SWISSPROT;Acc:Q8TF05]	18		
ENSG00000157014	TatD DNase domain-containing deoxyribonuclease 2 (EC 3.1.21.-). [Source:Uniprot/SWISSPROT;Acc:Q93075]	3		
ENSG00000159199	ATP synthase lipid-binding protein, mitochondrial precursor (EC 3.6.3.14) (ATP synthase proteolipid P1) (ATPase protein 9) (ATPase subunit C). [Source:Uniprot/SWISSPROT;Acc:P05496]	17	1	
ENSG00000159650	Probable urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate hydrolase). [Source:Uniprot/SWISSPROT;Acc:Q96N76]	3		
ENSG00000159733	Zinc finger FYVE domain-containing protein 28. [Source:Uniprot/SWISSPROT;Acc:Q9HCC9]	4		
ENSG00000162441	Protein LZIC (Leucine zipper and ICAT homologous domain-containing protein) (Leucine zipper and CTNNBIP1 domain-containing protein). [Source:Uniprot/SWISSPROT;Acc:Q8WZA0]	1		
ENSG00000162591	Multiple epidermal growth factor-like domains 6 precursor (EGF-like domain-containing protein 3) (Multiple EGF-like domain protein 3). [Source:Uniprot/SWISSPROT;Acc:O75095]	1	1	1
ENSG00000162819	Uncharacterized protein C1orf58. [Source:Uniprot/SPTREMBL;Acc:Q5VW32]	1		
ENSG00000164124	transmembrane protein 144 (TMEM144), mRNA [Source:RefSeq_dna;Acc:NM_018342]	4	1	
ENSG00000164535	Sn1-specific diacylglycerol lipase beta (EC 3.1.1.-) (DGL-beta) (KCCR13L). [Source:Uniprot/SWISSPROT;Acc:Q8NCG7]	7	1	1

ENSG00000165097	Flavin-containing amine oxidase domain-containing protein 1 (EC 1.-.-.-). [Source:Uniprot/SWISSPROT;Acc:Q8NB78]	6		
ENSG00000168288	Uncharacterized protein C2orf25, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q9H3L0]	2		
ENSG00000175711	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1 [Source:RefSeq_peptide;Acc:NP_001009905]	17		
ENSG00000187546	CDNA FLJ16237 fis, clone HCASM2002754 (Hypothetical protein FLJ16237). [Source:Uniprot/SPTREMBL;Acc:Q6ZNB7]	7	1	
ENSG00000188167	similar to F40B5.2b (FLJ45032), mRNA [Source:RefSeq_dna;Acc:NM_001039770]	3	1	
ENSG00000197408	Cytochrome P450 2B6 (EC 1.14.14.1) (CYP2B6) (P450 IIB1). [Source:Uniprot/SWISSPROT;Acc:P20813]	19		1
ENSG00000197915	Repetin. [Source:Uniprot/SWISSPROT;Acc:Q6XPR3]	1		

3.5 Appendix

3.5.1 Table for gene list generated using HomoloGene

Table A.1 Gene list generated using HomoloGene

This list of genes was generated from the list of HomoloGene groups. B₁₂ genes were labelled with green blocks. Gene for cblD was labelled with a yellow block. Chr = Chromosome; TM= Transmembrane Domain; SD= Singal Domain

EntrezGene ID	Ensembl Gene ID	Description	Chr	TD	SD
10000	ENSG00000117020	RAC-gamma serine/threonine-protein kinase (EC 2.7.11.1) (RAC-PK-gamma) (Protein kinase Akt-3) (Protein kinase B, gamma) (PKB gamma) (STK-2). [Source:Uniprot/SWISSPROT;Acc:Q9Y243]	1		
10007	ENSG00000113552	Glucosamine-6-phosphate isomerase (EC 3.5.99.6) (Glucosamine-6-phosphate deaminase) (GNPDA) (GlcN6P deaminase) (Oscillin). [Source:Uniprot/SWISSPROT;Acc:P46926]	5		
10036	ENSG00000167670	Chromatin assembly factor 1 subunit A (CAF-1 subunit A) (Chromatin assembly factor I p150 subunit) (CAF-I 150 kDa subunit) (CAF-Ip150). [Source:Uniprot/SWISSPROT;Acc:Q13111]	19		

10079	ENSG00000054793	Probable phospholipid-transporting ATPase IIA (EC 3.6.3.1) (ATPase class II type 9A) (ATPase IIA). [Source:Uniprot/SWISSPROT;Acc:O75110]	20	1	
10100	ENSG00000134198	Tetraspanin-2 (Tspan-2). [Source:Uniprot/SWISSPROT;Acc:O60636]	1	1	1
10188	ENSG00000061938	Activated CDC42 kinase 1 (EC 2.7.10.2) (ACK-1) (Tyrosine kinase non-receptor protein 2). [Source:Uniprot/SWISSPROT;Acc:Q07912]	3		
10217	ENSG00000144677	CTD small phosphatase-like protein (CTDSP-like) (Small C-terminal domain phosphatase 3) (Small CTD phosphatase 3) (SCP3) (Nuclear LIM interactor-interacting factor 1) (NLI-interacting factor 1) (NIF-like protein) (RBSP3) (YA22 protein) (HYA22). [Source:Uniprot/SWISSPROT;Acc:O15194]	3		
10220	ENSG00000135414	Growth/differentiation factor 11 precursor (GDF-11) (Bone morphogenetic protein 11). [Source:Uniprot/SWISSPROT;Acc:O95390]	12		1
10342	ENSG00000114354	Protein TFG (TRK-fused gene protein). [Source:Uniprot/SWISSPROT;Acc:Q92734]	3		
10455	ENSG00000198721	Peroxisomal 3,2-trans-enoyl-CoA isomerase (EC 5.3.3.8) (Dodecenoyl-CoA isomerase) (Delta(3),delta(2)-enoyl-CoA isomerase) (D3,D2-enoyl-CoA isomerase) (DBI-related protein 1) (DRS-1) (Hepatocellular carcinoma-associated antigen 88) (Renal carcinoma antigen 88). [Source:Uniprot/SWISSPROT;Acc:O75521]	6		

10559	ENSG00000164414	Uncharacterized protein C6orf165. [Source:Uniprot/SWISSPROT;Acc:Q8IYR0]	6	1	1
10568	ENSG00000157765	Sodium-dependent phosphate transport protein 2B (Sodium/phosphate cotransporter 2B) (Na(+)/Pi cotransporter 2B) (Sodium-phosphate transport protein 2B) (Na(+)-dependent phosphate cotransporter 2B) (NaPi-2b) (Solute carrier family 34 member 2) (NaPi3b). [Source:Uniprot/SWISSPROT;Acc:O95436]	4	1	
10690	ENSG00000172461	Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 9) (FucT-IX). [Source:Uniprot/SWISSPROT;Acc:Q9Y231]	6	1	1
10695	ENSG00000137161	trinucleotide repeat containing 5 [Source:RefSeq_peptide;Acc:NP_006577]	6		1
10735	ENSG00000101972	Cohesin subunit SA-2 (Stromal antigen 2) (SCC3 homolog 2). [Source:Uniprot/SWISSPROT;Acc:Q8N3U4]	X		
10765	ENSG00000117139	Jumonji, AT rich interactive domain 1B (RBP2-like) [Source:RefSeq_peptide;Acc:NP_006609]	1		
10771	ENSG00000015171	Zinc finger MYND domain-containing protein 11 (Adenovirus 5 E1A-binding protein) (BS69 protein). [Source:Uniprot/SWISSPROT;Acc:Q15326]	10		

10846	ENSG00000112541	cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A (EC 3.1.4.17). [Source:Uniprot/SWISSPROT;Acc:Q9Y233]	6		
10861	ENSG00000145217	Sulfate anion transporter 1 (SAT-1) (Solute carrier family 26 member 1). [Source:Uniprot/SWISSPROT;Acc:Q9H2B4]	4	1	
10892	ENSG00000172175	Mucosa-associated lymphoid tissue lymphoma translocation protein 1 (EC 3.4.22.-) (MALT lymphoma-associated translocation) (Paracaspase). [Source:Uniprot/SWISSPROT;Acc:Q9UDY8]	18		
10994	ENSG00000105135	ilvB (bacterial acetolactate synthase)-like isoform 1 [Source:RefSeq_peptide;Acc:NP_006835]	19	1	1
11040	ENSG00000102096	Serine/threonine-protein kinase Pim-2 (EC 2.7.11.1) (Pim-2h). [Source:Uniprot/SWISSPROT;Acc:Q9P1W9]	X		
1105	ENSG00000153922	Chromodomain-helicase-DNA-binding protein 1 (EC 3.6.1.-) (ATP-dependent helicase CHD1) (CHD-1). [Source:Uniprot/SWISSPROT;Acc:O14646]	5		
11152	ENSG00000196998	WD repeat domain phosphoinositide-interacting protein 4 (WIPI-4) (WD repeat protein 45). [Source:Uniprot/SWISSPROT;Acc:Q9Y484]	X		
11160	ENSG00000147475	SPFH domain-containing protein 2 precursor. [Source:Uniprot/SWISSPROT;Acc:O94905]	8		1

11165	ENSG00000112664	Diphosphoinositol polyphosphate phosphohydrolase 1 (EC 3.6.1.52) (DIPP-1) (Diadenosine 5',5'''-P1,P6-hexaphosphate hydrolase 1) (EC 3.6.1.-) (Nucleoside diphosphate-linked moiety X motif 3) (Nudix motif 3). [Source:Uniprot/SWISSPROT;Acc:O95989]	6		
11186	ENSG00000068028	Ras association domain-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9NS23]	3		
1124	ENSG00000106069	Beta-chimaerin (Beta-chimerin) (Rho GTPase-activating protein 3). [Source:Uniprot/SWISSPROT;Acc:P52757]	7		
11254	ENSG00000087916	Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+) (Amino acid transporter ATB0+) (Solute carrier family 6 member 14). [Source:Uniprot/SWISSPROT;Acc:Q9UN76]	X	1	
11328	ENSG00000122642	FK506-binding protein 9 precursor (EC 5.2.1.8) (Peptidyl-prolyl cis- trans isomerase) (PPIase) (Rotamase). [Source:Uniprot/SWISSPROT;Acc:O95302]	7		1
114034	ENSG00000132773	Target of EGR1 protein 1. [Source:Uniprot/SWISSPROT;Acc:Q96GM8]	1		
114294	ENSG00000103642	Serine beta-lactamase-like protein LACTB, mitochondrial precursor (EC 3.4.-.-). [Source:Uniprot/SWISSPROT;Acc:P83111]	15		
114987	ENSG00000148225	WD repeat protein 31. [Source:Uniprot/SWISSPROT;Acc:Q8NA23]	9		

115201	ENSG00000101844	Cysteine protease ATG4A (EC 3.4.22.-) (Autophagy-related protein 4 homolog A) (hAPG4A) (Autophagin-2) (Autophagy-related cysteine endopeptidase 2) (AUT-like 2 cysteine endopeptidase). [Source:Uniprot/SWISSPROT;Acc:Q8WYN0]	X		
115817	ENSG00000157379	Dehydrogenase/reductase SDR family member 1 (EC 1.1.-.-). [Source:Uniprot/SWISSPROT;Acc:Q96LJ7]	14		
1164	ENSG00000123975	Cyclin-dependent kinases regulatory subunit 2 (CKS-2). [Source:Uniprot/SWISSPROT;Acc:P33552]	9		
116988	ENSG00000133612	Centaurin-gamma 3 (ARF-GAP with GTP-binding protein-like, ankyrin repeat and pleckstrin homology domains 3) (AGAP-3) (MR1-interacting protein) (MRIP-1) (CRAM-associated GTPase) (CRAG). [Source:Uniprot/SWISSPROT;Acc:Q96P47]	7		
1183	ENSG00000073464	Chloride channel protein 4 (ClC-4). [Source:Uniprot/SWISSPROT;Acc:P51793]	X	1	
119391	ENSG00000065621	Glutathione transferase omega-2 (EC 2.5.1.18) (GSTO-2). [Source:Uniprot/SWISSPROT;Acc:Q9H4Y5]	10		
1207	ENSG00000074201	Methylosome subunit pICln (Chloride conductance regulatory protein ICln) (I(Cln)) (Chloride channel, nucleotide sensitive 1A) (Chloride ion current	11		

		inducer protein) (CICI) (Reticulocyte pICln). [Source:Uniprot/SWISSPROT;Acc:P54105]			
121214	ENSG00000170426	orphan short-chain dehydrogenase / reductase [Source:RefSeq_peptide;Acc:NP_683695]	12		
121278	ENSG00000139287	Tryptophan 5-hydroxylase 2 (EC 1.14.16.4) (Tryptophan 5-monooxygenase 2) (Neuronal tryptophan hydroxylase). [Source:Uniprot/SWISSPROT;Acc:Q8IWU9]	12		
122664	ENSG00000179636	Protein p25-beta. [Source:Uniprot/SWISSPROT;Acc:P59282]	14		
122970	ENSG00000177465	Acyl-coenzyme A thioesterase 4 (EC 3.1.2.2) (Acyl-CoA thioesterase 4) (Peroxisomal acyl coenzyme A thioester hydrolase Ib) (Peroxisomal long-chain acyl-coA thioesterase Ib) (PTE-Ib) (PTE-2b). [Source:Uniprot/SWISSPROT;Acc:Q8N9L9]	14		
124923	ENSG00000167524	Uncharacterized serine/threonine-protein kinase SgK494 (EC 2.7.11.1) (Sugen kinase 494). [Source:Uniprot/SWISSPROT;Acc:Q96LW2]	17		
1259	ENSG00000198515	cGMP-gated cation channel alpha 1 (CNG channel alpha 1) (CNG-1) (CNG1) (Cyclic nucleotide-gated channel alpha 1) (Cyclic nucleotide-gated channel, photoreceptor) (Cyclic nucleotide-gated cation channel 1) (Rod photoreceptor cGMP-gated channel subunit alph	4	1	

		[Source:Uniprot/SWISSPROT;Acc:P29973]			
126119	ENSG00000161677	Josephin-2 (Josephin domain-containing protein 2). [Source:Uniprot/SWISSPROT;Acc:Q8TAC2]	19		
126767	ENSG00000188984	Arylacetamide deacetylase-like 3 (EC 3.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q5VUY0]	1		
1272	ENSG0000018236	Contactin-1 precursor (Neural cell surface protein F3) (Glycoprotein gp135). [Source:Uniprot/SWISSPROT;Acc:Q12860]	12		1
127700	ENSG00000116885	Protein OSCP1 (Organic solute transport protein 1) (hOSCP1) (Oxidored-nitro domain-containing protein 1). [Source:Uniprot/SWISSPROT;Acc:Q8WVF1]	1		
1285	ENSG00000169031	Collagen alpha-3(IV) chain precursor (Goodpasture antigen) [Contains: Tumstatin]. [Source:Uniprot/SWISSPROT;Acc:Q01955]	2		
131601	ENSG00000163870	Integral membrane protein GPR175. [Source:Uniprot/SWISSPROT;Acc:Q86W33]	3	1	
131669	ENSG00000159650	Probable urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate hydrolase).	3		

		[Source:Uniprot/SWISSPROT;Acc:Q96N76]			
132321	ENSG00000151470	Uncharacterized protein C4orf33. [Source:Uniprot/SPTREMBL;Acc:Q8N1A6]	4		
133121	ENSG00000164303	Ectonucleotide pyrophosphatase/phosphodiesterase family member 6 precursor (EC 3.1.-.-) (E-NPP6) (NPP-6) [Contains: Ectonucleotide pyrophosphatase/phosphodiesterase family member 6 soluble form]. [Source:Uniprot/SWISSPROT;Acc:Q6UWR7]	4		1
133923	ENSG00000164185	Zinc finger protein 474. [Source:Uniprot/SWISSPROT;Acc:Q6S9Z5]	5		
1363	ENSG00000109472	Carboxypeptidase E precursor (EC 3.4.17.10) (CPE) (Carboxypeptidase H) (CPH) (Enkephalin convertase) (Prohormone-processing carboxypeptidase). [Source:Uniprot/SWISSPROT;Acc:P16870]	4		1
137735	ENSG00000174429	Actin-binding Rho-activating protein (Striated muscle activator of Rho-dependent signaling) (STARS). [Source:Uniprot/SWISSPROT;Acc:Q8N0Z2]	8		
137868	ENSG00000185053	Zeta-sarcoglycan (Zeta-SG) (ZSG1). [Source:Uniprot/SWISSPROT;Acc:Q96LD1]	8	1	
139065	ENSG00000179542	SLIT and NTRK-like protein 4 precursor. [Source:Uniprot/SWISSPROT;Acc:Q8IW52]	X	1	1

1396	ENSG00000182351	Cysteine-rich protein 1 (Cysteine-rich intestinal protein) (CRIP) (Cysteine-rich heart protein) (hCRHP). [Source:Uniprot/SWISSPROT;Acc:P50238]	14		
139818	ENSG00000147251	dedicator of cytokinesis 11 [Source:RefSeq_peptide;Acc:NP_653259]	X		
140	ENSG00000121933	Adenosine A3 receptor. [Source:Uniprot/SWISSPROT;Acc:P33765]	1	1	
140609	ENSG00000151414	Serine/threonine-protein kinase Nek7 (EC 2.7.11.1) (NimA-related protein kinase 7). [Source:Uniprot/SWISSPROT;Acc:Q8TDX7]	1		
1429	ENSG00000116791	Quinone oxidoreductase (EC 1.6.5.5) (NADPH:quinone reductase) (Zeta-crystallin). [Source:Uniprot/SWISSPROT;Acc:Q08257]	1		
144165	ENSG00000139174	Prickle-like protein 1 (REST/NRSF-interacting LIM domain protein 1). [Source:Uniprot/SWISSPROT;Acc:Q96MT3]	12		
144193	ENSG00000139344	Probable imidazolonepropionase (EC 3.5.2.7) (Amidohydrolase domain-containing protein 1). [Source:Uniprot/SWISSPROT;Acc:Q96NU7]	12		
146712	ENSG00000175711	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1 [Source:RefSeq_peptide;Acc:NP_001009905]	17		
148362	ENSG00000162819	Uncharacterized protein C1orf58. [Source:Uniprot/SPTREMBL;Acc:Q5VW32]	1		
149469	ENSG00000198198	Uncharacterized protein KIAA0467. [Source:Uniprot/SWISSPROT;Acc:Q5T011]	1		

150159	ENSG00000164037	Na ⁺ /H ⁺ exchanger like domain containing [Source:RefSeq_peptide;Acc:NP_631912]	4	1	
151393	ENSG00000115841	family with sequence similarity 82, member A (FAM82A), mRNA [Source:RefSeq_dna;Acc:NM_144713]	2		
152110	ENSG00000163491	Serine/threonine-protein kinase Nek10 (EC 2.7.11.1) (NimA-related protein kinase 10). [Source:Uniprot/SWISSPROT;Acc:Q6ZWH5]	3		
1522	ENSG00000101160	Cathepsin Z precursor (EC 3.4.22.-) (Cathepsin X) (Cathepsin P). [Source:Uniprot/SWISSPROT;Acc:Q9UBR2]	20	1	1
152503	ENSG00000109686	SH3 domain protein D19 [Source:RefSeq_peptide;Acc:NP_001009555]	4		
153129	ENSG00000177058	CDNA FLJ90709 fis, clone PLACE1007881 (Hypothetical protein FLJ90709). [Source:Uniprot/SPTREMBL;Acc:Q8NBW4]	5	1	
1538	ENSG00000183035	Cylicin-1 (Cylicin I) (Multiple-band polypeptide I). [Source:Uniprot/SWISSPROT;Acc:P35663]	X		
153830	ENSG00000145860	ring finger protein 145 (RNF145), mRNA [Source:RefSeq_dna;Acc:NM_144726]	5	1	
153918	ENSG00000118491	UPF0418 protein C6orf94. [Source:Uniprot/SWISSPROT;Acc:Q5TFG8]	6		
1555	ENSG00000197408	Cytochrome P450 2B6 (EC 1.14.14.1) (CYP11B6) (P450 I1B1). [Source:Uniprot/SWISSPROT;Acc:P20813]	19		1

1562	ENSG00000108242	Cytochrome P450 2C18 (EC 1.14.14.1) (CYPIIC18) (P450-6B/29C). [Source:Uniprot/SWISSPROT;Acc:P33260]	10		1
158046	ENSG00000130045	Thioredoxin domain-containing protein C9orf121. [Source:Uniprot/SWISSPROT;Acc:Q5VZ03]	9		
158158	ENSG00000165105	RAS and EF hand domain containing [Source:RefSeq_peptide;Acc:NP_689786]	9		
1605	ENSG00000173402	Dystroglycan precursor (Dystrophin-associated glycoprotein 1) [Contains: Alpha-dystroglycan (Alpha-DG); Beta-dystroglycan (Beta- DG)]. [Source:Uniprot/SWISSPROT;Acc:Q14118]	3	1	1
1612	ENSG00000196730	Death-associated protein kinase 1 (EC 2.7.11.1) (DAP kinase 1). [Source:Uniprot/SWISSPROT;Acc:P53355]	9		
163404	ENSG00000117598	phosphatidic acid phosphatase type 2d isoform 2 [Source:RefSeq_peptide;Acc:NP_001010861]	1	1	1
163486	ENSG00000162701	DENN/MADD domain containing 1B [Source:RefSeq_peptide;Acc:NP_659414]	1	1	
164091	ENSG00000182749	Membrane progesterin receptor alpha (mPR alpha) (Progesterin and adipoQ receptor family member VII). [Source:Uniprot/SWISSPROT;Acc:Q86WK9]	1	1	
164395	ENSG00000131044	tubulin tyrosine ligase-like family, member 9	20		

		[Source:RefSeq_peptide;Acc:NP_001008409]			
1666	ENSG00000104325	2,4-dienoyl-CoA reductase, mitochondrial precursor (EC 1.3.1.34) (2,4-dienoyl-CoA reductase [NADPH]) (4-enoyl-CoA reductase [NADPH]). [Source:Uniprot/SWISSPROT;Acc:Q16698]	8		
166614	ENSG00000170390	Serine/threonine-protein kinase DCAMKL2 (EC 2.7.11.1) (Doublecortin- like and CAM kinase-like 2). [Source:Uniprot/SWISSPROT;Acc:Q8N568]	4		
166785	ENSG00000151611	Methylmalonic aciduria type A protein, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q8IVH4]	1		
167127	ENSG00000168671	UDP glycosyltransferase 3 family, polypeptide A2 [Source:RefSeq_peptide;Acc:NP_777574]	5	1	1
169966	ENSG00000174016	Protein FAM46D. [Source:Uniprot/SWISSPROT;Acc:Q8NEK8]	X		
171392	ENSG00000197372	zinc finger protein 675 [Source:RefSeq_peptide;Acc:NP_612203]	19		
1749	ENSG00000105880	Homeobox protein DLX-5. [Source:Uniprot/SWISSPROT;Acc:P56178]	7		
1777	ENSG00000105612	Deoxyribonuclease-2-alpha precursor (EC 3.1.22.1) (Deoxyribonuclease II alpha) (DNase II alpha) (Acid DNase) (Lysosomal DNase II) (R31240_2). [Source:Uniprot/SWISSPROT;Acc:O00115]	19		1
1811	ENSG00000091138	Chloride anion exchanger (Protein DRA) (Down-regulated in adenoma) (Solute carrier family 26 member 3).	7	1	

		[Source:Uniprot/SWISSPROT;Acc:P40879]			
1813	ENSG00000149295	D(2) dopamine receptor (Dopamine D2 receptor). [Source:Uniprot/SWISSPROT;Acc:P14416]	11	1	
1820	ENSG00000116017	AT-rich interactive domain-containing protein 3A (ARID domain- containing protein 3A) (B-cell regulator of IgH transcription) (Bright) (E2F-binding protein 1). [Source:Uniprot/SWISSPROT;Acc:Q99856]	19		
1827	ENSG00000159200	Calcipressin-1 (Down syndrome critical region protein 1) (Myocyte- enriched calcineurin-interacting protein 1) (MCIP1) (Adapt78). [Source:Uniprot/SWISSPROT;Acc:P53805]	21		
1837	ENSG00000134769	Dystrobrevin alpha (Dystrobrevin-alpha) (Dystrophin-related protein 3). [Source:Uniprot/SWISSPROT;Acc:Q9Y4J8]	18		
1850	ENSG00000184545	Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16) (Dual specificity protein phosphatase hVH-5). [Source:Uniprot/SWISSPROT;Acc:Q13202]	11		
1948	ENSG00000125266	Ephrin-B2 precursor (EPH-related receptor tyrosine kinase ligand 5) (LERK-5) (HTK ligand) (HTK-L). [Source:Uniprot/SWISSPROT;Acc:P52799]	13	1	1
1953	ENSG00000162591	Multiple epidermal growth factor-like domains 6 precursor (EGF-like	1	1	

		domain-containing protein 3) (Multiple EGF-like domain protein 3). [Source:Uniprot/SWISSPROT;Acc:O75095]			
1960	ENSG00000179388	Early growth response protein 3 (EGR-3) (Zinc finger protein pilot). [Source:Uniprot/SWISSPROT;Acc:Q06889]	8		
196463	ENSG00000151176		12	1	1
197407	ENSG00000180035	Zinc finger protein 553. [Source:Uniprot/SWISSPROT;Acc:Q96MX3]	16		
1995	ENSG00000196361	ELAV-like protein 3 (Hu-antigen C) (HuC) (Paraneoplastic cerebellar degeneration-associated antigen) (Paraneoplastic limbic encephalitis antigen 21). [Source:Uniprot/SWISSPROT;Acc:Q14576]	19		
2002	ENSG00000126767	ETS domain-containing protein Elk-1. [Source:Uniprot/SWISSPROT;Acc:P19419]	X		
200894	ENSG00000169379	ADP-ribosylation factor-like protein 13B (ADP-ribosylation factor-like protein 2-like 1) (ARL2-like protein 1). [Source:Uniprot/SWISSPROT;Acc:Q3SXY8]	3		
2045	ENSG00000135333	Ephrin type-A receptor 7 precursor (EC 2.7.10.1) (Tyrosine-protein kinase receptor EHK-3) (EPH homology kinase 3) (Receptor protein- tyrosine kinase HEK11). [Source:Uniprot/SWISSPROT;Acc:Q15375]	6	1	1

207	ENSG00000142208	RAC-alpha serine/threonine-protein kinase (EC 2.7.11.1) (RAC-PK-alpha) (Protein kinase B) (PKB) (C-AKT). [Source:Uniprot/SWISSPROT;Acc:P31749]	14		
2139	ENSG00000064655	Eyes absent homolog 2 (EC 3.1.3.48). [Source:Uniprot/SWISSPROT;Acc:O00167]	20		
2167	ENSG00000170323	Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding protein) (ALBP) (A-FABP). [Source:Uniprot/SWISSPROT;Acc:P15090]	8		
2192	ENSG00000077942	Fibulin-1 precursor. [Source:Uniprot/SWISSPROT;Acc:P23142]	22		1
2200	ENSG00000166147	Fibrillin-1 precursor. [Source:Uniprot/SWISSPROT;Acc:P35555]	15		1
221656	ENSG00000165097	Flavin-containing amine oxidase domain-containing protein 1 (EC 1.-.-.-). [Source:Uniprot/SWISSPROT;Acc:Q8NB78]	6		
221955	ENSG00000164535	Sn1-specific diacylglycerol lipase beta (EC 3.1.1.-) (DGL-beta) (KCCR13L). [Source:Uniprot/SWISSPROT;Acc:Q8NCG7]	7	1	1
222223	ENSG00000164659	KIAA1324-like (KIAA1324L), mRNA [Source:RefSeq_dna;Acc:NM_152748]	7	1	
2242	ENSG00000182511	Proto-oncogene tyrosine-protein kinase Fes/Fps (EC 2.7.10.2) (C-Fes). [Source:Uniprot/SWISSPROT;Acc:P07332]	15		
22808	ENSG00000158186	Ras-related protein M-Ras precursor (Ras-related protein R-Ras3).	3		

		[Source:Uniprot/SWISSPROT;Acc:O14807]			
22904	ENSG00000064932	KIAA0963 [Source:RefSeq_peptide;Acc:NP_055778]	19		
22990	ENSG00000100731	Pecanex-like protein 1 (Pecanex homolog). [Source:Uniprot/SWISSPROT;Acc:Q96RV3]	14	1	
23017	ENSG00000135472	Fas apoptotic inhibitory molecule 2 (Lifeguard protein) (Transmembrane BAX inhibitor motif-containing protein 2). [Source:Uniprot/SWISSPROT;Acc:Q9BWQ8]	12	1	
23025	ENSG00000130477	Unc-13 homolog A (Munc13-1). [Source:Uniprot/SWISSPROT;Acc:Q9UPW8]	19		
23057	ENSG00000157064	Nicotinamide mononucleotide adenylyltransferase 2 (EC 2.7.7.1) (NMN adenylyltransferase 2). [Source:Uniprot/SWISSPROT;Acc:Q9BZQ4]	1		
23062	ENSG00000103365	ADP-ribosylation factor-binding protein GGA2 (Golgi-localized, gamma ear- containing, ARF-binding protein 2) (Gamma-adaptin-related protein 2) (VHS domain and ear domain of gamma-adaptin) (Vear). [Source:Uniprot/SWISSPROT;Acc:Q9UJY4]	16		
23063	ENSG00000062650	Wings apart-like protein homolog (Friend of EBNA2 protein). [Source:Uniprot/SWISSPROT;Acc:Q7Z5K2]	10		

23072	ENSG00000002746	E3 ubiquitin-protein ligase HECW1 (EC 6.3.2.-) (HECT, C2 and WW domain-containing protein 1) (NEDD4-like ubiquitin-protein ligase 1) (hNEDL1). [Source:Uniprot/SWISSPROT;Acc:Q76N89]	7		
23102	ENSG00000167202	TBC1 domain family, member 2B [Source:RefSeq_peptide;Acc:NP_055894]	15		
23111	ENSG00000133104	Spartin (Trans-activated by hepatitis C virus core protein 1). [Source:Uniprot/SWISSPROT;Acc:Q8N0X7]	13		
23116	ENSG00000198718	Uncharacterized protein KIAA0423. [Source:Uniprot/SWISSPROT;Acc:Q9Y4F4]	14		1
23125	ENSG00000108509	Calmodulin-binding transcription activator 2. [Source:Uniprot/SWISSPROT;Acc:O94983]	17		
23164	ENSG00000133030	Myosin phosphatase Rho-interacting protein (Rho-interacting protein 3) (M-RIP) (RIP3) (p116Rip). [Source:Uniprot/SWISSPROT;Acc:Q6WCQ1]	17		
23167	ENSG00000132294	Protein EFR3-like. [Source:Uniprot/SWISSPROT;Acc:Q14156]	8		
2318	ENSG00000128591	Filamin-C (Gamma-filamin) (Filamin-2) (Protein FLNc) (Actin-binding- like protein) (ABP-L) (ABP-280-like protein). [Source:Uniprot/SWISSPROT;Acc:Q14315]	7		

2321	ENSG00000102755	Vascular endothelial growth factor receptor 1 precursor (EC 2.7.10.1) (VEGFR-1) (Vascular permeability factor receptor) (Tyrosine-protein kinase receptor FLT) (Flt-1) (Tyrosine-protein kinase FRT) (Fms-like tyrosine kinase 1). [Source:Uniprot/SWISSPROT;Acc:P17948]	13		1
23211	ENSG00000130749	Zinc finger CCCH domain-containing protein C19orf7. [Source:Uniprot/SWISSPROT;Acc:Q9UPT8]	19		
23228	ENSG00000154822	phospholipase C-like 2 [Source:RefSeq_peptide;Acc:NP_055999]	3		
23232	ENSG00000108239	TBC1 domain family member 12. [Source:Uniprot/SWISSPROT;Acc:O60347]	10		
23247	ENSG00000047578	KIAA0556 (KIAA0556), mRNA [Source:RefSeq_dna;Acc:NM_015202]	16		
23266	ENSG00000117114	Latrophilin-2 precursor (Calcium-independent alpha-latrotoxin receptor 2) (Latrophilin homolog 1) (Lectomedin-1). [Source:Uniprot/SWISSPROT;Acc:O95490]	1	1	1
23275	ENSG00000186866	GDP-fucose protein O-fucosyltransferase 2 precursor (EC 2.4.1.221) (Peptide-O-fucosyltransferase 2) (O-FucT-2). [Source:Uniprot/SWISSPROT;Acc:Q9Y2G5]	21		1
23322	ENSG00000103494		16		
23333	ENSG00000173852	DPY19L1 protein. [Source:Uniprot/SPTREMBL;Acc:Q4G151]	7	1	

23443	ENSG00000117620	UDP-N-acetylglucosamine transporter (Golgi UDP-GlcNAc transporter) (Solute carrier family 35 member A3). [Source:Uniprot/SWISSPROT;Acc:Q9Y2D2]	1	1	1
23457	ENSG00000150967	ATP-binding cassette sub-family B member 9 precursor (ATP-binding cassette transporter 9) (ABC transporter 9 protein) (TAP-like protein) (TAPL) (hABCB9). [Source:Uniprot/SWISSPROT;Acc:Q9NP78]	12	1	1
23484	ENSG00000104660	Leptin receptor overlapping transcript-like 1. [Source:Uniprot/SWISSPROT;Acc:O95214]	8	1	
23516	ENSG00000104635	solute carrier family 39 (zinc transporter), member 14 [Source:RefSeq_peptide;Acc:NP_056174]	8	1	1
23617	ENSG00000182490	Testis-specific serine/threonine-protein kinase 2 (EC 2.7.11.1) (TSSK- 2) (Testis-specific kinase 2) (TSK-2) (Serine/threonine-protein kinase 22B) (DiGeorge syndrome protein G). [Source:Uniprot/SWISSPROT;Acc:Q96PF2]	22		
23729	ENSG00000197417	Carbohydrate kinase-like protein (EC 2.7.1.-). [Source:Uniprot/SWISSPROT;Acc:Q9UJH6]	17		
24	ENSG00000198691	Retinal-specific ATP-binding cassette transporter (ATP-binding cassette sub- family A member 4) (RIM ABC transporter) (RIM protein) (RmP) (Stargardt	1	1	1

		disease protein). [Source:Uniprot/SWISSPROT;Acc:P78363]			
25	ENSG00000097007	Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.10.2) (p150) (c- ABL) (Abelson murine leukemia viral oncogene homolog 1). [Source:Uniprot/SWISSPROT;Acc:P00519]	9		
2525	ENSG00000171124	Galactoside 3(4)-L-fucosyltransferase (EC 2.4.1.65) (Blood group Lewis alpha-4-fucosyltransferase) (Lewis FT) (Fucosyltransferase 3) (FucT- III). [Source:Uniprot/SWISSPROT;Acc:P21217]	19		1
2526	ENSG00000196371	Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 4) (FucT-IV) (Fuc-TIV) (ELAM-1 ligand fucosyltransferase). [Source:Uniprot/SWISSPROT;Acc:P22083]	11	1	
2529	ENSG00000180549	Alpha-(1,3)-fucosyltransferase (EC 2.4.1.-) (Galactoside 3-L-fucosyltransferase) (Fucosyltransferase 7) (FucT-VII) (Selectin-ligand synthase). [Source:Uniprot/SWISSPROT;Acc:Q11130]	9	1	1
2534	ENSG00000010810	Proto-oncogene tyrosine-protein kinase Fyn (EC 2.7.10.2) (p59-Fyn) (Protooncogene Syn) (SLK). [Source:Uniprot/SWISSPROT;Acc:P06241]	6		
254102	ENSG00000173442	Signal-induced proliferation-associated protein 1 (Sipa-1) (GTPase-activating protein Spa-1) (p130 SPA-1).	11		

		[Source:Uniprot/SWISSPROT;Acc:Q96FS4]			
2542	ENSG00000137700	Glucose-6-phosphate translocase (Glucose-5-phosphate transporter) (Solute carrier family 37 member 4). [Source:Uniprot/SWISSPROT;Acc:O43826]	11	1	
254428	ENSG00000133065	solute carrier family 41 member 1 [Source:RefSeq_peptide;Acc:NP_776253]	1	1	
255426	ENSG00000146090	RasGEF domain family, member 1C [Source:RefSeq_peptide;Acc:NP_778232]	5		
25780	ENSG00000152689	RAS guanyl-releasing protein 3 (Calcium and DAG-regulated guanine nucleotide exchange factor III) (Guanine nucleotide exchange factor for Rap1). [Source:Uniprot/SWISSPROT;Acc:Q8IV61]	2		
25844	ENSG00000137207	Protein YIPF3 (YIP1 family member 3) (Killer lineage protein 1) (Natural killer cell-specific antigen KLIP1). [Source:Uniprot/SWISSPROT;Acc:Q9GZM5]	6	1	
25864	ENSG00000042022	Abhydrolase domain-containing protein 14A (EC 3.-.-.-). [Source:Uniprot/SWISSPROT;Acc:Q9BUJ0]	3		1
25917	ENSG00000134077	THUMP domain-containing protein 3. [Source:Uniprot/SWISSPROT;Acc:Q9BV44]	3		
2595	ENSG00000092529	Neutral alpha-glucosidase C (EC 3.2.1.20).	15		

		[Source:Uniprot/SWISSPROT;Acc:Q8TET4]			
25961	ENSG00000166321	Nucleoside diphosphate-linked moiety X motif 13 (EC 3.-.-) (Nudix motif 13) (Protein KiSS-16). [Source:Uniprot/SWISSPROT;Acc:Q86X67]	10		
25974	ENSG00000132763	Methylmalonic aciduria and homocystinuria type C protein. [Source:Uniprot/SWISSPROT;Acc:Q9Y4U1]	1		
26033	ENSG00000107518	attractin-like 1 [Source:RefSeq_peptide;Acc:NP_997186]	10	1	
26038	ENSG00000116254	Chromodomain helicase-DNA-binding protein 5 (EC 3.6.1.-) (ATP-dependent helicase CHD5) (CHD-5). [Source:Uniprot/SWISSPROT;Acc:Q8TDI0]	1		
26059	ENSG00000187672	ERC protein 2. [Source:Uniprot/SWISSPROT;Acc:O15083]	3		
261734	ENSG00000131697	Nephrocystin-4 (Nephroretinin). [Source:Uniprot/SWISSPROT;Acc:O75161]	1		
26191	ENSG00000134242	Tyrosine-protein phosphatase non-receptor type 22 (EC 3.1.3.48) (Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP) (Lymphoid phosphatase) (LyP). [Source:Uniprot/SWISSPROT;Acc:Q9Y2R2]	1		
2623	ENSG00000102145	Erythroid transcription factor (GATA-binding factor 1) (GATA-1) (Eryf1) (GF-1) (NF-E1 DNA-binding protein). [Source:Uniprot/SWISSPROT;Acc:P15976]	X		
26276	ENSG00000184056	Vacuolar protein sorting-associated protein 33B (hVPS33B).	15		

		[Source:Uniprot/SWISSPROT;Acc:Q9H267]			
26281	ENSG00000078579	Fibroblast growth factor 20 (FGF-20). [Source:Uniprot/SWISSPROT;Acc:Q9NP95]	8		
26289	ENSG00000154027	Adenylate kinase isoenzyme 5 (EC 2.7.4.3) (ATP-AMP transphosphorylase). [Source:Uniprot/SWISSPROT;Acc:Q9Y6K8]	1		
2629	ENSG00000177628	Glucosylceramidase precursor (EC 3.2.1.45) (Beta-glucocerebrosidase) (Acid beta-glucosidase) (D-glucosyl-N-acylsphingosine glucohydrolase) (Alglucerase) (Imiglucerase). [Source:Uniprot/SWISSPROT;Acc:P04062]	1		1
2650	ENSG00000187210	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase (EC 2.4.1.102) (Core 2 branching enzyme) (Core2-GlcNAc-transferase) (C2GNT) (Core 2 GNT). [Source:Uniprot/SWISSPROT;Acc:Q02742]	9	1	1
26578	ENSG00000134996	Osteoclast-stimulating factor 1. [Source:Uniprot/SWISSPROT;Acc:Q92882]	9		
266675	ENSG00000142959	Bestrophin-4 (Vitelliform macular dystrophy 2-like protein 2). [Source:Uniprot/SWISSPROT;Acc:Q8NFU0]	1	1	
27012	ENSG00000164794	potassium channel, subfamily V, member 1 [Source:RefSeq_peptide;Acc:NP_055194]	8	1	

27034	ENSG00000151498	Acyl-CoA dehydrogenase family member 8, mitochondrial precursor (EC 1.3.99.-) (ACAD-8) (Isobutyryl-CoA dehydrogenase) (Activator- recruited cofactor 42 kDa component) (ARC42). [Source:Uniprot/SWISSPROT;Acc:Q9UKU7]	11		
27106	ENSG00000105643	Arrestin domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q8TBH0]	19		
27130	ENSG00000119509	Inversin (Inversion of embryo turning homolog) (Nephrocystin-2). [Source:Uniprot/SWISSPROT;Acc:Q9Y283]	9		
27131	ENSG00000089006	Sorting nexin-5. [Source:Uniprot/SWISSPROT;Acc:Q9Y5X3]	20		
27163	ENSG00000138744	N-acylethanolamine-hydrolyzing acid amidase precursor (EC 3.5.1.-) (N-acylsphingosine amidohydrolase-like) (ASAH-like protein) (Acid ceramidase-like protein). [Source:Uniprot/SWISSPROT;Acc:Q02083]	4	1	1
27165	ENSG00000135423	Glutaminase liver isoform, mitochondrial precursor (EC 3.5.1.2) (GLS) (L-glutamine amidohydrolase) (L-glutaminase). [Source:Uniprot/SWISSPROT;Acc:Q9UI32]	12		
27249	ENSG00000168288	Uncharacterized protein C2orf25, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q9H3L0]	2		
27319	ENSG00000180828	Class B basic helix-loop-helix protein 5 (bHLHB5).	8		

		[Source:Uniprot/SWISSPROT;Acc:Q8NFJ8]			
27348	ENSG00000136816	Torsin B precursor (Torsin family 1 member B). [Source:Uniprot/SWISSPROT;Acc:O14657]	9	1	1
2742	ENSG00000101958	Glycine receptor subunit alpha-2 precursor. [Source:Uniprot/SWISSPROT;Acc:P23416]	X	1	1
27436	ENSG00000143924	Echinoderm microtubule-associated protein-like 4 (EMAP-4) (Restrictedly overexpressed proliferation-associated protein) (Ropp 120). [Source:Uniprot/SWISSPROT;Acc:Q9HC35]	2		
2766	ENSG00000137198	GMP reductase 1 (EC 1.7.1.7) (Guanosine 5'-monophosphate oxidoreductase 1) (Guanosine monophosphate reductase 1). [Source:Uniprot/SWISSPROT;Acc:P36959]	6		
2781	ENSG00000128266	Guanine nucleotide-binding protein G(z) subunit alpha (G(x) alpha chain) (Gz-alpha). [Source:Uniprot/SWISSPROT;Acc:P19086]	22		
28232	ENSG00000176463	Solute carrier organic anion transporter family member 3A1 (Solute carrier family 21 member 11) (Sodium-independent organic anion transporter D) (Organic anion-transporting polypeptide D) (OATP-D) (Organic anion transporter polypeptide-related protein 3) [Source:Uniprot/SWISSPROT;Acc:Q9UIG8]	15	1	

283629	ENSG00000139908	Testis-specific serine/threonine-protein kinase 4 (EC 2.7.11.1) (TSSK- 4) (Testis-specific kinase 4) (TSK-4) (Serine/threonine-protein kinase 22E). [Source:Uniprot/SWISSPROT;Acc:Q6SA08]	14		
283767	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
284013	ENSG00000182853	Vitelline membrane outer layer protein 1 homolog precursor. [Source:Uniprot/SWISSPROT;Acc:Q7Z5L0]	17		1
285704	ENSG00000174136	RGM domain family member B precursor. [Source:Uniprot/SWISSPROT;Acc:Q6NW40]	5		
286380	ENSG00000187559	forkhead box D4 like 3 [Source:RefSeq_peptide;Acc:NP_955390]	9		
2870	ENSG00000198055	G protein-coupled receptor kinase 6 (EC 2.7.11.16) (G protein-coupled receptor kinase GRK6). [Source:Uniprot/SWISSPROT;Acc:P43250]	5		
2882	ENSG00000116157	Glutathione peroxidase 7 precursor (EC 1.11.1.9) (CL683). [Source:Uniprot/SWISSPROT;Acc:Q96SL4]	1		1
2890	ENSG00000155511	Glutamate receptor 1 precursor (GluR-1) (GluR-A) (GluR-K1) (Glutamate receptor ionotropic, AMPA 1) (AMPA-selective glutamate receptor 1). [Source:Uniprot/SWISSPROT;Acc:P42261]	5	1	1
2911	ENSG00000152822	Metabotropic glutamate receptor 1 precursor (mGluR1).	6	1	1

		[Source:Uniprot/SWISSPROT;Acc:Q13255]			
2914	ENSG00000124493	Metabotropic glutamate receptor 4 precursor (mGluR4). [Source:Uniprot/SWISSPROT;Acc:Q14833]	6	1	1
292	ENSG00000005022	ADP/ATP translocase 2 (Adenine nucleotide translocator 2) (ANT 2) (ADP,ATP carrier protein 2) (Solute carrier family 25 member 5) (ADP,ATP carrier protein, fibroblast isoform). [Source:Uniprot/SWISSPROT;Acc:P05141]	X	1	
2950	ENSG00000084207	Glutathione S-transferase P (EC 2.5.1.18) (GST class-pi) (GSTP1-1). [Source:Uniprot/SWISSPROT;Acc:P09211]	11		
29895	ENSG00000180209	Myosin regulatory light chain 2, skeletal muscle isoform (Fast skeletal myosin light chain 2) (MLC2B). [Source:Uniprot/SWISSPROT;Acc:Q96A32]	16		
29904	ENSG00000103319	Elongation factor 2 kinase (EC 2.7.11.20) (eEF-2 kinase) (eEF-2K) (Calcium/calmodulin-dependent eukaryotic elongation factor 2 kinase). [Source:Uniprot/SWISSPROT;Acc:O00418]	16		
29958	ENSG00000132837	Dimethylglycine dehydrogenase, mitochondrial precursor (EC 1.5.99.2) (ME2GLYDH). [Source:Uniprot/SWISSPROT;Acc:Q9UI17]	5		
29966	ENSG00000196792	Striatin-3 (Cell-cycle autoantigen SG2NA) (S/G2 antigen).	14		

		[Source:Uniprot/SWISSPROT;Acc:Q13033]			
29985	ENSG00000141873	solute carrier family 39 (zinc transporter), member 3 isoform a [Source:RefSeq_peptide;Acc:NP_653165]	19	1	1
29994	ENSG00000123636	Bromodomain adjacent to zinc finger domain protein 2B (hWALp4). [Source:Uniprot/SWISSPROT;Acc:Q9UIF8]	2		
30011	ENSG00000147010	SH3 domain-containing kinase-binding protein 1 (Cbl-interacting protein of 85 kDa) (Human Src-family kinase-binding protein 1) (HSB-1) (CD2-binding protein 3) (CD2BP3). [Source:Uniprot/SWISSPROT;Acc:Q96B97]	X		
3006	ENSG00000187837	Histone H1.2 (Histone H1d). [Source:Uniprot/SWISSPROT;Acc:P16403]	6		
3007	ENSG00000124575	Histone H1.3 (Histone H1c). [Source:Uniprot/SWISSPROT;Acc:P16402]	6		
3008	ENSG00000168298	Histone H1.4 (Histone H1b). [Source:Uniprot/SWISSPROT;Acc:P10412]	6		
3017	ENSG00000158373	Histone H2B type 1-D (H2B.b) (H2B/b) (H2B.1 B) (HIRA-interacting protein 2). [Source:Uniprot/SWISSPROT;Acc:P58876]	6		
3034	ENSG00000084110	Histidine ammonia-lyase (EC 4.3.1.3) (Histidase). [Source:Uniprot/SWISSPROT;Acc:P42357]	12		
3062	ENSG00000137252	Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2). [Source:Uniprot/SWISSPROT;Acc:O43614]	6	1	

3174	ENSG00000164749	Hepatocyte nuclear factor 4-gamma (HNF-4-gamma). [Source:Uniprot/SWISSPROT;Acc:Q14541]	8		
3177	ENSG00000174669	Equilibrative nucleoside transporter 2 (Equilibrative nitrobenzylmercaptapurine riboside-insensitive nucleoside transporter) (Equilibrative NBMPR-insensitive nucleoside transporter) (Nucleoside transporter, ei-type) (Solute carrier family 29 member 2) (36 [Source:Uniprot/SWISSPROT;Acc:Q14542]	11	1	1
3187	ENSG00000169045	Heterogeneous nuclear ribonucleoprotein H (hnRNP H). [Source:Uniprot/SWISSPROT;Acc:P31943]	5		
3190	ENSG00000165119	Heterogeneous nuclear ribonucleoprotein K (hnRNP K) (Transformation up- regulated nuclear protein) (TUNP). [Source:Uniprot/SWISSPROT;Acc:P61978]	9		
3198	ENSG00000105991	Homeobox protein Hox-A1 (Hox-1F). [Source:Uniprot/SWISSPROT;Acc:P49639]	7		
3217	ENSG00000120087	Homeobox protein Hox-B7 (Hox-2C) (HHO.C1). [Source:Uniprot/SWISSPROT;Acc:P09629]	17		
3251	ENSG00000165704	Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPRT) (HGPRTase). [Source:Uniprot/SWISSPROT;Acc:P00492]	X		

326625	ENSG00000119420	Cob(D)yrinic acid a,c-diamide adenosyltransferase, mitochondrial precursor (EC 2.5.1.17) (Cob(D)alamin adenosyltransferase) (Methylmalonic aciduria type B protein). [Source:Uniprot/SWISSPROT;Acc:Q96EY8]	12		
3274	ENSG00000113749	Histamine H2 receptor (H2R) (Gastric receptor I). [Source:Uniprot/SWISSPROT;Acc:P25021]	5	1	1
337876	ENSG00000198108	Chondroitin sulfate synthase 3 (EC 2.4.1.175) (Glucuronosyl-N-acetylgalactosaminyl-proteoglycan 4-beta-N- acetylgalactosaminyltransferase II) (Chondroitin synthase 2) (N- acetylgalactosaminyl-proteoglycan 3-beta-glucuronosyltransferase II) (EC 2.4.1.226) [Source:Uniprot/SWISSPROT;Acc:Q70JA7]	5	1	1
338567	ENSG00000186795	potassium channel, subfamily K, member 18 [Source:RefSeq_peptide;Acc:NP_862823]	10	1	1
338917	ENSG00000119614	Homeobox protein CHX10 (Ceh-10 homeodomain-containing homolog). [Source:Uniprot/SWISSPROT;Acc:P58304]	14		
339123	ENSG00000161999	Similar to C. Elegans protein F17C8.5. [Source:Uniprot/SPTREMBL;Acc:Q96S16]	16		
340156	ENSG00000145949	Uncharacterized serine/threonine-protein kinase SgK085 (EC 2.7.11.1) (Sugen kinase 85). [Source:Uniprot/SWISSPROT;Acc:Q86YV6]	6		

342125	ENSG00000188869	Transmembrane channel-like protein 3. [Source:Uniprot/SWISSPROT;Acc:Q7Z5M5]	15	1	
349334	ENSG00000204828	Forkhead box protein D4-like 4 (Forkhead box protein D4B) (Myeloid factor-gamma). [Source:Uniprot/SWISSPROT;Acc:Q8WXT5]	9		
353189	ENSG00000173930	solute carrier organic anion transporter family, member 4C1 [Source:RefSeq_peptide;Acc:NP_851322]	5	1	
3619	ENSG00000149503	Inner centromere protein. [Source:Uniprot/SWISSPROT;Acc:Q9NQS7]	11		
3713	ENSG00000163207	Involucrin. [Source:Uniprot/SWISSPROT;Acc:P07476]	1		
373	ENSG00000113595	GTP-binding protein ARD-1 (ADP-ribosylation factor domain protein 1) (Tripartite motif-containing protein 23) (RING finger protein 46). [Source:Uniprot/SWISSPROT;Acc:P36406]	5		
373156	ENSG00000197448	Glutathione S-transferase kappa 1 (EC 2.5.1.18) (GST 13-13) (Glutathione S-transferase subunit 13) (GST class-kappa) (GSTK1-1) (hGSTK1). [Source:Uniprot/SWISSPROT;Acc:Q9Y2Q3]	7		
375323	ENSG00000156959	lipoma HMGIC fusion partner-like 4 [Source:RefSeq_peptide;Acc:NP_940962]	3	1	
3754	ENSG00000162975	Potassium voltage-gated channel subfamily F member 1 (Voltage-gated potassium channel subunit Kv5.1) (KH1).	2	1	

		[Source:Uniprot/SWISSPROT;Acc:Q9H3M0]			
375790	ENSG00000188157	Agrin precursor. [Source:Uniprot/SWISSPROT;Acc:O00468]	1		1
3775	ENSG00000135750	Potassium channel subfamily K member 1 (Inward rectifying potassium channel protein TWIK-1) (Potassium channel KCNO1). [Source:Uniprot/SWISSPROT;Acc:O00180]	1	1	1
377711	ENSG00000179832	CDNA FLJ35542 fis, clone SPLEN2002917. [Source:Uniprot/SPTREMBL;Acc:Q8NAD1]	8		
3790	ENSG00000170745	Potassium voltage-gated channel subfamily S member 3 (Voltage-gated potassium channel subunit Kv9.3) (Delayed-rectifier K(+) channel alpha subunit 3). [Source:Uniprot/SWISSPROT;Acc:Q9BQ31]	2	1	
3850	ENSG00000186442	Keratin, type II cytoskeletal 3 (Cytokeratin-3) (CK-3) (Keratin-3) (K3) (65 kDa cytokeratin). [Source:Uniprot/SWISSPROT;Acc:P12035]	12		
386675	ENSG00000205441	Keratin-associated protein 10-7 (Keratin-associated protein 10.7) (High sulfur keratin-associated protein 10.7) (Keratin-associated protein 18-7) (Keratin-associated protein 18.7). [Source:Uniprot/SWISSPROT;Acc:P60409]	21		
386677	ENSG00000205445	Keratin-associated protein 10-1 (Keratin-associated protein 10.1) (High sulfur keratin-associated protein 10.1) (Keratin-associated protein 18-1) (Keratin-	21		

		associated protein 18.1). [Source:Uniprot/SWISSPROT;Acc:P60331]			
387694	ENSG00000178217	SH2 domain containing 4B [Source:RefSeq_peptide;Acc:NP_997255]	10		
387893	ENSG00000183955	Histone-lysine N-methyltransferase, H4 lysine-20 specific (EC 2.1.1.43) (Histone H4-K20 methyltransferase) (H4-K20-HMTase) (SET domain- containing lysine methyltransferase 8) (SET domain-containing protein 8) (PR/SET domain-containing protein 07) (PR/SET07 [Source:Uniprot/SWISSPROT;Acc:Q9NQR1]	12		
388591	ENSG00000158286	ring finger protein 207 [Source:RefSeq_peptide;Acc:NP_997279]	1		
388697	ENSG00000197915	Repetin. [Source:Uniprot/SWISSPROT;Acc:Q6XPR3]	1		
388743	ENSG00000203697	Calpain 8 (NCL-2). [Source:Uniprot/SPTREMBL;Acc:A2CEN9]	1		
389058	ENSG00000204335	Transcription factor Sp5. [Source:Uniprot/SWISSPROT;Acc:Q6BEB4]	2		
389072	ENSG00000178385	CDNA FLJ41210 fis, clone BRALZ2014484. [Source:Uniprot/SPTREMBL;Acc:Q6ZWE6]	2		
390010	ENSG00000206148	HPX homeobox protein (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q9UD57]	4		
3908	ENSG00000196569	Laminin subunit alpha-2 precursor (Laminin M chain) (Merosin heavy chain). [Source:Uniprot/SWISSPROT;Acc:P24043]	6		1

392636	ENSG00000187546	CDNA FLJ16237 fis, clone HCASM2002754 (Hypothetical protein FLJ16237). [Source:Uniprot/SPTREMBL;Acc:Q6ZNB7]	7	1	
394	ENSG00000100852	Rho GTPase-activating protein 5 (p190-B). [Source:Uniprot/SWISSPROT;Acc:Q13017]	14		
3965	ENSG00000168961	Galectin-9 (HOM-HD-21) (Ecalectin). [Source:Uniprot/SWISSPROT;Acc:O00182]	17		
3992	ENSG00000149485	fatty acid desaturase 3 [Source:RefSeq_peptide;Acc:NP_068373]	11	1	
3995	ENSG00000149485	fatty acid desaturase 3 [Source:RefSeq_peptide;Acc:NP_068373]	11	1	
400224	ENSG00000175985		14		
4065	ENSG00000054219	Lymphocyte antigen 75 precursor (DEC-205) (gp200-MR6) (CD205 antigen). [Source:Uniprot/SWISSPROT;Acc:O60449]	2	1	1
4093	ENSG00000120693	Mothers against decapentaplegic homolog 9 (SMAD 9) (Mothers against DPP homolog 9) (Smad9) (Madh6). [Source:Uniprot/SWISSPROT;Acc:O15198]	13		
41	ENSG00000110881	amiloride-sensitive cation channel 2, neuronal isoform a [Source:RefSeq_peptide;Acc:NP_064423]	12		
4134	ENSG00000047849	Microtubule-associated protein 4 (MAP 4). [Source:Uniprot/SWISSPROT;Acc:P27816]	3		

4139	ENSG00000116141	Serine/threonine-protein kinase MARK1 (EC 2.7.11.1) (MAP/microtubule affinity-regulating kinase 1). [Source:Uniprot/SWISSPROT;Acc:Q9P0L2]	1		
414308	ENSG00000183748	mannose receptor, C type 1-like 1 [Source:RefSeq_peptide;Acc:NP_001009567]	10	1	1
416	ENSG00000062096	Arylsulfatase F precursor (EC 3.1.6.-) (ASF). [Source:Uniprot/SWISSPROT;Acc:P54793]	X	1	1
4209	ENSG00000116604	Myocyte-specific enhancer factor 2D. [Source:Uniprot/SWISSPROT;Acc:Q14814]	1		
4249	ENSG00000152127	Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase V (EC 2.4.1.155) (Mannoside acetylglucosaminyltransferase 5) (Alpha-mannoside beta-1,6-N-acetylglucosaminyltransferase) (N-acetylglucosaminyl-transferase V) (GNT-V) (GlcNAc-T V). [Source:Uniprot/SWISSPROT;Acc:Q09328]	2	1	1
427	ENSG00000104763	Acid ceramidase precursor (EC 3.5.1.23) (Acylsphingosine deacylase) (N-acylsphingosine amidohydrolase) (AC) (Putative 32 kDa heart protein) (PHP32) [Contains: Acid ceramidase subunit alpha; Acid ceramidase subunit beta]. [Source:Uniprot/SWISSPROT;Acc:Q13510]	8		1
4306	ENSG00000151623	Mineralocorticoid receptor (MR).	4		

		[Source:Uniprot/SWISSPROT;Acc:P08235]			
4360	ENSG00000120586	Macrophage mannose receptor 1 precursor (MMR) (CD206 antigen). [Source:Uniprot/SWISSPROT;Acc:P22897]	10	1	1
440	ENSG00000070669	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) (Glutamine-dependent asparagine synthetase) (Cell cycle control protein TS11). [Source:Uniprot/SWISSPROT;Acc:P08243]	7		
440107	ENSG00000187510	CDNA FLJ46688 fis, clone TRACH3011538. [Source:Uniprot/SPTREMBL;Acc:Q6ZR37]	12		
440233	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
440243	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
440320	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
440822	ENSG00000184571	Piwi-like protein 3. [Source:Uniprot/SWISSPROT;Acc:Q7Z3Z3]	22		
441168	ENSG00000188820	Protein FAM26F. [Source:Uniprot/SWISSPROT;Acc:Q5R3K3]	6	1	

4486	ENSG00000164078	Macrophage-stimulating protein receptor precursor (EC 2.7.10.1) (MSP receptor) (p185-Ron) (CD136 antigen) (CDw136) [Contains: Macrophage-stimulating protein receptor alpha chain; Macrophage-stimulating protein receptor beta chain]. [Source:Uniprot/SWISSPROT;Acc:Q04912]	3	1	1
4534	ENSG00000171100	Myotubularin (EC 3.1.3.48). [Source:Uniprot/SWISSPROT;Acc:Q13496]	X		
4552	ENSG00000124275	Methionine synthase reductase, mitochondrial precursor (EC 1.16.1.8) (MSR). [Source:Uniprot/SWISSPROT;Acc:Q9UBK8]	1	1	1
4594	ENSG00000146083	Methylmalonyl-CoA mutase, mitochondrial precursor (EC 5.4.99.2) (MCM) (Methylmalonyl-CoA isomerase). [Source:Uniprot/SWISSPROT;Acc:P22033]	1		
4625	ENSG00000092054	Myosin-7 (Myosin heavy chain 7) (Myosin heavy chain, cardiac muscle beta isoform) (MyHC-beta) (Myosin heavy chain slow isoform) (MyHC- slow). [Source:Uniprot/SWISSPROT;Acc:P12883]	14		
4632	ENSG00000168530	Myosin light chain 1, skeletal muscle isoform (MLC1F) (A1 catalytic) (Alkali myosin light chain 1). [Source:Uniprot/SWISSPROT;Acc:P05976]	2		
4648	ENSG00000169994	myosin VIIIB [Source:RefSeq_peptide;Acc:NP_001073996]	2		
4659	ENSG00000058272	Protein phosphatase 1 regulatory subunit 12A (Myosin phosphatase- targeting subunit 1) (Myosin phosphatase target subunit 1) (Protein phosphatase	12		

		myosin-binding subunit). [Source:Uniprot/SWISSPROT;Acc:O14974]			
4744	ENSG00000100285	Neurofilament heavy polypeptide (NF-H) (Neurofilament triplet H protein) (200 kDa neurofilament protein). [Source:Uniprot/SWISSPROT;Acc:P12036]	22		
4760	ENSG00000162992	Neurogenic differentiation factor 1 (NeuroD1) (NeuroD). [Source:Uniprot/SWISSPROT;Acc:Q13562]	2		
4850	ENSG00000080802	CCR4-NOT transcription complex subunit 4 (EC 6.3.2.-) (E3 ubiquitin protein ligase CNOT4) (CCR4-associated factor 4) (Potential transcriptional repressor NOT4Hp). [Source:Uniprot/SWISSPROT;Acc:O95628]	7		
4867	ENSG00000144061	Nephrocystin-1 (Juvenile nephronophthisis 1 protein). [Source:Uniprot/SWISSPROT;Acc:O15259]	2		
4889	ENSG00000164129	Neuropeptide Y receptor type 5 (NPY5-R) (NPY-Y5 receptor) (Y5 receptor) (NPYY5). [Source:Uniprot/SWISSPROT;Acc:Q15761]	4	1	
49855	ENSG00000140386	Zinc finger protein 291. [Source:Uniprot/SWISSPROT;Acc:Q9BY12]	15		

5002	ENSG00000110628	Organic cation transporter-like protein 2 (Imprinted multi-membrane spanning polyspecific transporter-related protein 1) (Efflux transporter-like protein) (Tumor-suppressing subchromosomal transferable fragment candidate gene 5 protein) (Tumor-suppressing [Source:Uniprot/SWISSPROT;Acc:Q96BI1]	11	1	
5015	ENSG00000165588	Homeobox protein OTX2 (Orthodenticle homolog 2). [Source:Uniprot/SWISSPROT;Acc:P32243]	14		
50488	ENSG00000141503	Misshapen-like kinase 1 (EC 2.7.11.1) (Mitogen-activated protein kinase kinase kinase 6) (MAPK/ERK kinase kinase kinase 6) (MEK kinase kinase 6) (MEKKK 6) (Misshapen/NIK-related kinase) (GCK family kinase MiNK). [Source:Uniprot/SWISSPROT;Acc:Q8N4C8]	17		
5051	ENSG00000158006	Platelet-activating factor acetylhydrolase 2, cytoplasmic (EC 3.1.1.47) (Serine-dependent phospholipase A2) (HSD-PLA2). [Source:Uniprot/SWISSPROT;Acc:Q99487]	1		
50636	ENSG00000146205	transmembrane protein 16G isoform NGEF long [Source:RefSeq_peptide;Acc:NP_001001891]	2	1	

5092	ENSG00000166228	Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96) (PHS) (4-alpha-hydroxy-tetrahydropterin dehydratase) (Phenylalanine hydroxylase-stimulating protein) (Pterin carbinolamine dehydratase) (PCD) (Dimerization cofactor of hepatocyte nuclear factor 1- α) [Source:Uniprot/SWISSPROT;Acc:P61457]	10		
5096	ENSG00000114054	Propionyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit beta) (Propanoyl-CoA:carbon dioxide ligase subunit beta). [Source:Uniprot/SWISSPROT;Acc:P05166]	1	1	1
51002	ENSG00000144034	TP53RK-binding protein (PRPK-binding protein). [Source:Uniprot/SWISSPROT;Acc:Q9Y3C4]	2		
51024	ENSG00000106404	Mitochondrial fission 1 protein (Fis1 homolog) (hFis1) (Tetratricopeptide repeat protein 11) (TPR repeat protein 11). [Source:Uniprot/SWISSPROT;Acc:Q9Y3D6]	7	1	
51060	ENSG00000117862	Thioredoxin domain-containing protein 12 precursor (EC 1.8.4.2) (Thioredoxin-like protein p19) (Endoplasmic reticulum protein ERp19) (ERp18) (hTLP19). [Source:Uniprot/SWISSPROT;Acc:O95881]	1	1	1
51086	ENSG00000116783	Serine/threonine-protein kinase TNNI3K (EC 2.7.11.1) (TNNI3- interacting kinase) (Cardiac ankyrin repeat kinase).	1		

		[Source:Uniprot/SWISSPROT;Acc:Q59H18]			
51092	ENSG00000149577	SID1 transmembrane family member 2 precursor. [Source:Uniprot/SWISSPROT;Acc:Q8NBJ9]	11	1	1
51099	ENSG00000011198	Abhydrolase domain-containing protein 5. [Source:Uniprot/SWISSPROT;Acc:Q8WTS1]	3		
51111	ENSG00000110066	Histone-lysine N-methyltransferase SUV420H1 (EC 2.1.1.43) (Suppressor of variegation 4-20 homolog 1) (Suv4-20h1) (Su(var)4-20 homolog 1). [Source:Uniprot/SWISSPROT;Acc:Q4FZB7]	11		
51232	ENSG00000150938	Cysteine-rich motor neuron 1 protein precursor (CRIM-1) (Cysteine-rich repeat-containing protein S52). [Source:Uniprot/SWISSPROT;Acc:Q9NZV1]	2	1	1
5128	ENSG00000059758	Serine/threonine-protein kinase PCTAIRE-2 (EC 2.7.11.22) (PCTAIRE-motif protein kinase 2). [Source:Uniprot/SWISSPROT;Acc:Q00537]	12		
51285	ENSG00000103710	RAS-like, family 12 protein [Source:RefSeq_peptide;Acc:NP_057647]	15		
51361	ENSG00000134709	Hook homolog 1 (h-hook1) (hHK1). [Source:Uniprot/SWISSPROT;Acc:Q9UJC3]	1		
51384	ENSG00000002745	Protein Wnt-16 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9UBV4]	7		
5139	ENSG00000172572	cGMP-inhibited 3',5'-cyclic phosphodiesterase A (EC 3.1.4.17) (Cyclic GMP-	12	1	

		inhibited phosphodiesterase A) (CGI-PDE A). [Source:Uniprot/SWISSPROT;Acc:Q14432]			
51458	ENSG00000140519	Ammonium transporter Rh type C (Rh family type C glycoprotein) (Rhesus blood group family type C glycoprotein) (Rh type C glycoprotein) (Tumor-related protein DRC2) (Rh glycoprotein kidney). [Source:Uniprot/SWISSPROT;Acc:Q9UBD6]	15	1	1
5151	ENSG00000073417	High-affinity cAMP-specific and IBMX-insensitive 3',5'-cyclic phosphodiesterase 8A (EC 3.1.4.17). [Source:Uniprot/SWISSPROT;Acc:O60658]	15		
51524	ENSG00000149483	transmembrane protein 138 (TMEM138), mRNA [Source:RefSeq_dna;Acc:NM_016464]	11	1	
51540	ENSG00000132330	selenocysteine lyase [Source:RefSeq_peptide;Acc:NP_057594]	2		
5156	ENSG00000134853	Alpha platelet-derived growth factor receptor precursor (EC 2.7.10.1) (PDGF-R-alpha) (CD140a antigen). [Source:Uniprot/SWISSPROT;Acc:P16234]	4	1	1
51567	ENSG00000111802	TRAF and TNF receptor-associated protein (ETS1-associated protein 2) (ETS1-associated protein II) (EAPII). [Source:Uniprot/SWISSPROT;Acc:O95551]	6		

51573	ENSG00000006007	Glycerophosphodiester phosphodiesterase 1 (EC 3.1.4.44) (Membrane interacting protein of RGS16) (RGS16-interacting membrane protein). [Source:Uniprot/SWISSPROT;Acc:Q9NZC3]	16	1	1
516	ENSG00000159199	ATP synthase lipid-binding protein, mitochondrial precursor (EC 3.6.3.14) (ATP synthase proteolipid P1) (ATPase protein 9) (ATPase subunit C). [Source:Uniprot/SWISSPROT;Acc:P05496]	17	1	
51663	ENSG00000056097	zinc finger RNA binding protein [Source:RefSeq_peptide;Acc:NP_057191]	5		
51747	ENSG00000108848	Cisplatin resistance-associated overexpressed protein (cAMP regulatory element-associated protein 1) (CRE-associated protein 1) (CREAP-1) (Luc7A) (Okadaic acid-inducible phosphoprotein OA48-18). [Source:Uniprot/SWISSPROT;Acc:O95232]	17		
51807	ENSG00000183785	Peroxisome assembly protein 26 (Peroxin-26). [Source:Uniprot/SWISSPROT;Acc:Q7Z412]	22		
5264	ENSG00000107537	Phytanoyl-CoA dioxygenase, peroxisomal precursor (EC 1.14.11.18) (Phytanoyl-CoA alpha-hydroxylase) (PhyH) (Phytanic acid oxidase). [Source:Uniprot/SWISSPROT;Acc:O14832]	10		
5290	ENSG00000121879	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (PtdIns-3- kinase	3		

		p110) (PI3K). [Source:Uniprot/SWISSPROT;Acc:P42336]			
5292	ENSG00000137193	Proto-oncogene serine/threonine-protein kinase Pim-1 (EC 2.7.11.1). [Source:Uniprot/SWISSPROT;Acc:P11309]	6		
53	ENSG00000134575	Lysosomal acid phosphatase precursor (EC 3.1.3.2) (LAP). [Source:Uniprot/SWISSPROT;Acc:P11117]	11	1	1
5319	ENSG00000170890	Phospholipase A2 precursor (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase) (Group IB phospholipase A2). [Source:Uniprot/SWISSPROT;Acc:P04054]	12		1
53344	ENSG00000204116	Cysteine-rich hydrophobic domain 1 protein (Brain X-linked protein). [Source:Uniprot/SWISSPROT;Acc:Q5VXU3]	X	1	
5375	ENSG00000147588	Myelin P2 protein. [Source:Uniprot/SWISSPROT;Acc:P02689]	8		
53834	ENSG00000127418	Fibroblast growth factor receptor-like 1 precursor (FGF receptor-like protein 1) (Fibroblast growth factor receptor 5) (FGFR-like protein) (FGF homologous factor receptor). [Source:Uniprot/SWISSPROT;Acc:Q8N441]	4	1	1
54361	ENSG00000162552	Protein Wnt-4 precursor. [Source:Uniprot/SWISSPROT;Acc:P56705]	1	1	1
54431	ENSG00000077232	DnaJ homolog subfamily C member 10 precursor (ER-resident protein ERdj5) (Macrothioredoxin) (MTHr).	2	1	

		[Source:Uniprot/SWISSPROT;Acc:Q8IXB1]			
5445	ENSG00000105854	Serum paraoxonase/arylesterase 2 (EC 3.1.1.2) (EC 3.1.8.1) (PON 2) (Serum arylalkylphosphatase 2) (A-esterase 2) (Aromatic esterase 2). [Source:Uniprot/SWISSPROT;Acc:Q15165]	7		1
5446	ENSG00000105852	Serum paraoxonase/lactonase 3 (EC 3.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q15166]	7	1	1
54658	ENSG00000167165	UDP-glucuronosyltransferase 1-8 precursor (EC 2.4.1.17) (UDP-glucuronosyltransferase 1A8) (UDPGT) (UGT1*8) (UGT1-08) (UGT1.8) (UGT- 1H) (UGT1H). [Source:Uniprot/SWISSPROT;Acc:Q9HAW9]	2	1	1
54659	ENSG00000167165	UDP-glucuronosyltransferase 1-8 precursor (EC 2.4.1.17) (UDP-glucuronosyltransferase 1A8) (UDPGT) (UGT1*8) (UGT1-08) (UGT1.8) (UGT- 1H) (UGT1H). [Source:Uniprot/SWISSPROT;Acc:Q9HAW9]	2	1	1
54663	ENSG00000133316	WD repeat protein 74 (NOP seven-associated protein 1). [Source:Uniprot/SWISSPROT;Acc:Q6RFH5]	11		
54665	ENSG00000081019	round spermatid basic protein 1 [Source:RefSeq_peptide;Acc:NP_060834]	1		
54810	ENSG00000137960	PDZ domain-containing protein GIPC2. [Source:Uniprot/SWISSPROT;Acc:Q8TF65]	1		

54887	ENSG00000065060	UHRF1-binding protein 1 (Ubiquitin-like containing PHD and RING finger domains 1-binding protein 1) (ICBP90-binding protein 1). [Source:Uniprot/SWISSPROT;Acc:Q6BDS2]	6		
550	ENSG00000115307	Ancient ubiquitous protein 1 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9Y679]	2		
55079	ENSG00000153266	zinc finger protein 312 [Source:RefSeq_peptide;Acc:NP_060478]	3		
55102	ENSG00000066739	Uncharacterized protein C14orf103. [Source:Uniprot/SWISSPROT;Acc:Q96BY7]	14		
55112	ENSG00000126870	WD repeat protein 60. [Source:Uniprot/SWISSPROT;Acc:Q8WVS4]	7		
552	ENSG00000166148	Vasopressin V1a receptor (V1aR) (Vascular/hepatic-type arginine vasopressin receptor) (Antidiuretic hormone receptor 1a) (AVPR V1a). [Source:Uniprot/SWISSPROT;Acc:P37288]	12	1	
55251	ENSG00000203880	Protein-L-isoaspartate O-methyltransferase domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q9NV79]	20		
55314	ENSG00000164124	transmembrane protein 144 (TMEM144), mRNA [Source:RefSeq_dna;Acc:NM_018342]	4	1	
55315	ENSG00000198246	Equilibrative nucleoside transporter 3 (Solute carrier family 29 member 3). [Source:Uniprot/SWISSPROT;Acc:Q9BZD2]	10	1	

55326	ENSG00000155189	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51) (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid acyltransferase epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate O-acyltransferase 5). [Source:Uniprot/SWISSPROT;Acc:Q9NUQ2]	8	1	1
554	ENSG00000126895	Vasopressin V2 receptor (Renal-type arginine vasopressin receptor) (Antidiuretic hormone receptor) (AVPR V2). [Source:Uniprot/SWISSPROT;Acc:P30518]	X	1	
5540	ENSG00000204174	Neuropeptide Y receptor type 4 (NPY4-R) (Pancreatic polypeptide receptor 1) (PP1). [Source:Uniprot/SWISSPROT;Acc:P50391]	10	1	
5545	ENSG00000121335	Basic salivary proline-rich protein 4 allele S precursor (Salivary proline-rich protein Po) (Parotid o protein) [Contains: Protein N1; Glycosylated protein A]. [Source:Uniprot/SWISSPROT;Acc:P10163]	12		1
55510	ENSG00000080007	Probable ATP-dependent RNA helicase DDX43 (EC 3.6.1.-) (DEAD box protein 43) (DEAD box protein HAGE) (Helical antigen). [Source:Uniprot/SWISSPROT;Acc:Q9NXZ2]	6		
55515	ENSG00000072182	Amiloride-sensitive cation channel 4 (Amiloride-sensitive cation channel 4, pituitary) (Acid-sensing ion channel 4). [Source:Uniprot/SWISSPROT;Acc:Q96FT7]	2		

55529	ENSG00000155099	Transmembrane protein 55A (EC 3.1.3.-) (Type II phosphatidylinositol 4,5-bisphosphate 4-phosphatase) (PtdIns-4,5-P2 4-Ptase II). [Source:Uniprot/SWISSPROT;Acc:Q8N4L2]	8	1	
55584	ENSG00000174343	Neuronal acetylcholine receptor subunit alpha-9 precursor (Nicotinic acetylcholine receptor subunit alpha 9) (NACHR alpha 9). [Source:Uniprot/SWISSPROT;Acc:Q9UGM1]	4	1	1
55643	ENSG00000133243	BTB/POZ domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q9BX70]	19		
55667	ENSG00000137145	DENN domain-containing protein 4C. [Source:Uniprot/SWISSPROT;Acc:Q5VZ89]	9		
55676	ENSG00000152683	solute carrier family 30 (zinc transporter), member 6 [Source:RefSeq_peptide;Acc:NP_060434]	2	1	
55690	ENSG00000175115	Phosphofurin acidic cluster sorting protein 1 (PACS-1). [Source:Uniprot/SWISSPROT;Acc:Q6VY07]	11		
55704	ENSG00000115355	Hook-related protein 1 [Source:RefSeq_peptide;Acc:NP_060554]	2		
55740	ENSG00000154380	Protein enabled homolog. [Source:Uniprot/SWISSPROT;Acc:Q8N8S7]	1		
55761	ENSG00000052841	Tetratricopeptide repeat protein 17 (TPR repeat protein 17). [Source:Uniprot/SWISSPROT;Acc:Q96AE7]	11		1

55814	ENSG00000145734	transcription factor-like nuclear regulator [Source:RefSeq_peptide;Acc:NP_060899]	5		
55840	ENSG00000145088	ELL-associated factor 2 (Testosterone-regulated apoptosis inducer and tumor suppressor protein). [Source:Uniprot/SWISSPROT;Acc:Q96CJ1]	3		
55847	ENSG00000122873	Zinc finger CDGSH domain-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9NZ45]	10	1	
55852	ENSG00000136478	Testis-expressed sequence 2 protein. [Source:Uniprot/SWISSPROT;Acc:Q8IWB9]	17	1	
55862	ENSG00000093144	Enoyl-CoA hydratase domain-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9NTX5]	6		
55892	ENSG00000085274	Myoneurin (Zinc finger and BTB domain-containing protein 31). [Source:Uniprot/SWISSPROT;Acc:Q9NPC7]	3		
5608	ENSG00000108984	Dual specificity mitogen-activated protein kinase kinase 6 (EC 2.7.12.2) (MAP kinase kinase 6) (MAPKK 6) (MAPK/ERK kinase 6) (SAPKK3). [Source:Uniprot/SWISSPROT;Acc:P52564]	17		
5663	ENSG00000080815	Presenilin-1 (EC 3.4.23.-) (PS-1) (Protein S182) [Contains: Presenilin-1 NTF subunit; Presenilin-1 CTF subunit; Presenilin-1 CTF12 (PS1-CTF12)]. [Source:Uniprot/SWISSPROT;Acc:P49768]	14	1	

56659	ENSG00000152315	Potassium channel subfamily K member 13 (Tandem pore domain halothane-inhibited potassium channel 1) (THIK-1). [Source:Uniprot/SWISSPROT;Acc:Q9HB14]	14	1	1
56674	ENSG00000175348	Transmembrane protein 9B precursor. [Source:Uniprot/SWISSPROT;Acc:Q9NQ34]	11	1	1
56751	ENSG00000125492	BarH-like 1 homeobox protein. [Source:Uniprot/SWISSPROT;Acc:Q9BZE3]	9		
56850	ENSG00000068400	GRIP1-associated protein 1 (GRASP-1). [Source:Uniprot/SWISSPROT;Acc:Q4V328]	X		
56900	ENSG00000162647	Uncharacterized protein C1orf119 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9NRX6]	1	1	1
56912	ENSG00000118096	UPF0360 protein C11orf60. [Source:Uniprot/SWISSPROT;Acc:Q9NQC8]	11		
56924	ENSG00000137843	Serine/threonine-protein kinase PAK 6 (EC 2.7.11.1) (p21-activated kinase 6) (PAK-6) (PAK-5). [Source:Uniprot/SWISSPROT;Acc:Q9NQU5]	15		
56994	ENSG00000111666	choline phosphotransferase 1 [Source:RefSeq_peptide;Acc:NP_064629]	12	1	
56995	ENSG00000130338	Tubby-related protein 4 (Tubby-like protein 4) (Tubby superfamily protein). [Source:Uniprot/SWISSPROT;Acc:Q9NRJ4]	6		
56999	ENSG00000163638	ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).	3		1

		[Source:Uniprot/SWISSPROT;Acc:Q9P2N4]			
57060	ENSG00000090097	Poly(rC)-binding protein 4 (Alpha-CP4). [Source:Uniprot/SWISSPROT;Acc:P57723]	3		
57120	ENSG00000047932	Golgi-associated PDZ and coiled-coil motif-containing protein (PDZ protein interacting specifically with TC10) (PIST) (CFTR-associated ligand) (Fused in glioblastoma). [Source:Uniprot/SWISSPROT;Acc:Q9HD26]	6		
5716	ENSG00000101843	26S proteasome non-ATPase regulatory subunit 10 (26S proteasome regulatory subunit p28) (Gankyrin). [Source:Uniprot/SWISSPROT;Acc:O75832]	X		
57231	ENSG00000135317	Sorting nexin-14. [Source:Uniprot/SWISSPROT;Acc:Q9Y5W7]	6	1	
57460	ENSG00000111110	Protein phosphatase 1H (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q9ULR3]	12		
57464	ENSG00000128578	Protein FAM40B. [Source:Uniprot/SWISSPROT;Acc:Q9ULQ0]	7		
57504	ENSG00000057935	Metastasis-associated protein MTA3. [Source:Uniprot/SWISSPROT;Acc:Q9BTC8]	2		
57513	ENSG00000177303	Caskin-2. [Source:Uniprot/SWISSPROT;Acc:Q8WXE0]	17		
57552	ENSG00000144959	Arylacetamide deacetylase-like 1 (EC 3.1.1.-).	3	1	

		[Source:Uniprot/SWISSPROT;Acc:Q6PIU2]			
57563	ENSG00000145332	Kelch-like protein 8. [Source:Uniprot/SWISSPROT;Acc:Q9P2G9]	4		
57569	ENSG00000137727	Rho GTPase-activating protein 20 (Rho-type GTPase-activating protein 20). [Source:Uniprot/SWISSPROT;Acc:Q9P2F6]	11		
57644	ENSG00000078814	myosin, heavy polypeptide 7B, cardiac muscle, beta [Source:RefSeq_peptide;Acc:NP_065935]	20		
57645	ENSG00000143157	Pogo transposable element with KRAB domain (LST003/SLTP003). [Source:Uniprot/SWISSPROT;Acc:Q9P215]	1		
57699	ENSG00000124772	Copine-5 (Copine V). [Source:Uniprot/SWISSPROT;Acc:Q9HCH3]	6		
57716	ENSG00000105227	Periaxin. [Source:Uniprot/SWISSPROT;Acc:Q9BXM0]	19		
57718	ENSG00000119698	HEAT-like repeat-containing protein isoform 1 [Source:RefSeq_peptide;Acc:NP_478144]	14		
57732	ENSG00000159733	Zinc finger FYVE domain-containing protein 28. [Source:Uniprot/SWISSPROT;Acc:Q9HCC9]	4		
5866	ENSG00000167994	RAB3A interacting protein (rabin3)-like 1 [Source:RefSeq_peptide;Acc:NP_037533]	11		
5934	ENSG00000103479	Retinoblastoma-like protein 2 (130 kDa retinoblastoma-associated protein) (PRB2) (P130) (RBR-2). [Source:Uniprot/SWISSPROT;Acc:Q08999]	16		

60681	ENSG00000141756	FK506-binding protein 10 precursor (EC 5.2.1.8) (Peptidyl-prolyl cis- trans isomerase) (PPIase) (Rotamase) (65 kDa FK506-binding protein) (FKBP65) (Immunophilin FKBP65). [Source:Uniprot/SWISSPROT;Acc:Q96AY3]	17		1
6102	ENSG00000102218	Protein XRP2. [Source:Uniprot/SWISSPROT;Acc:O75695]	X		
6195	ENSG00000117676	Ribosomal protein S6 kinase alpha-1 (EC 2.7.11.1) (S6K-alpha 1) (90 kDa ribosomal protein S6 kinase 1) (p90-RSK 1) (Ribosomal S6 kinase 1) (RSK-1) (pp90RSK1) (p90S6K) (MAP kinase-activated protein kinase 1a) (MAPKAPK1A). [Source:Uniprot/SWISSPROT;Acc:Q15418]	1		
6242	ENSG00000114993	Rhotekin. [Source:Uniprot/SWISSPROT;Acc:Q9BST9]	2		
6337	ENSG00000111319	Amiloride-sensitive sodium channel subunit alpha (Epithelial Na ⁽⁺⁾ channel subunit alpha) (Alpha ENaC) (Nonvoltage-gated sodium channel 1 subunit alpha) (SCNEA) (Alpha NaCH). [Source:Uniprot/SWISSPROT;Acc:P37088]	12	1	
6338	ENSG00000168447	Amiloride-sensitive sodium channel subunit beta (Epithelial Na ⁽⁺⁾ channel subunit beta) (Beta ENaC) (Nonvoltage-gated sodium channel 1 subunit beta) (SCNEB) (Beta NaCH) (ENaCB). [Source:Uniprot/SWISSPROT;Acc:P51168]	16	1	
6339	ENSG00000162572	Amiloride-sensitive sodium channel subunit delta (Epithelial Na ⁽⁺⁾ channel subunit delta) (Delta ENaC) (Nonvoltage-gated sodium channel 1 subunit	1	1	

		delta) (SCNED) (Delta NaCH). [Source:Uniprot/SWISSPROT;Acc:P51172]			
6340	ENSG00000166828	Amiloride-sensitive sodium channel subunit gamma (Epithelial Na ⁽⁺⁾ channel subunit gamma) (Gamma ENaC) (Nonvoltage-gated sodium channel 1 subunit gamma) (SCNEG) (Gamma NaCH). [Source:Uniprot/SWISSPROT;Acc:P51170]	16	1	
63906	ENSG00000198746	G patch domain-containing protein 3. [Source:Uniprot/SWISSPROT;Acc:Q96176]	1		
63928	ENSG00000166869	Calcineurin B homologous protein 2 (Hepatocellular carcinoma- associated antigen 520). [Source:Uniprot/SWISSPROT;Acc:O43745]	16		
64077	ENSG00000107902	phospholysine phosphohistidine inorganic pyrophosphate phosphatase [Source:RefSeq_peptide;Acc:NP_071409]	10		
64106	ENSG00000148734	Neuropeptide FF receptor 1 (G-protein coupled receptor 147) (RFamide-related peptide receptor OT7T022). [Source:Uniprot/SWISSPROT;Acc:Q9GZQ6]	10	1	1
64114	ENSG00000135926	Transmembrane BAX inhibitor motif-containing protein 1 (RECS1 protein homolog). [Source:Uniprot/SWISSPROT;Acc:Q969X1]	2	1	

64135	ENSG00000115267	Interferon-induced helicase C domain-containing protein 1 (EC 3.6.1.-) (Interferon-induced with helicase C domain protein 1) (Helicase with 2 CARD domains) (Helicard) (Melanoma differentiation-associated protein 5) (MDA-5) (RNA helicase-DEAD box protein 1 [Source:Uniprot/SWISSPROT;Acc:Q9BYX4]	2		
641371	ENSG00000184227	Acyl-coenzyme A thioesterase 1 (EC 3.1.2.2) (Acyl-CoA thioesterase 1) (Inducible cytosolic acyl-coenzyme A thioester hydrolase) (Long chain acyl- CoA thioester hydrolase) (Long chain acyl-CoA hydrolase) (CTE-I) (CTE- Ib). [Source:Uniprot/SWISSPROT;Acc:Q86TX2]	14		
6425	ENSG00000120057	Secreted frizzled-related protein 5 precursor (sFRP-5) (Secreted apoptosis- related protein 3) (SARP-3) (Frizzled-related protein 1b) (FRP-1b). [Source:Uniprot/SWISSPROT;Acc:Q5T4F7]	10		1
64285	ENSG00000007384	rhomboid family 1 [Source:RefSeq_peptide;Acc:NP_071895]	16	1	
6429	ENSG00000116350	Splicing factor, arginine/serine-rich 4 (Pre-mRNA-splicing factor SRP75) (SRP001LB). [Source:Uniprot/SWISSPROT;Acc:Q08170]	1		
64375	ENSG00000123411	zinc finger protein, subfamily 1A, 4 [Source:RefSeq_peptide;Acc:NP_071910]	12		
643853	ENSG00000188167	similar to F40B5.2b (FLJ45032), mRNA	3	1	

		[Source:RefSeq_dna;Acc:NM_001039770]			
6441	ENSG00000133661	Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D) (Lung surfactant protein D). [Source:Uniprot/SWISSPROT;Acc:P35247]	10		1
6449	ENSG00000104969	Small glutamine-rich tetratricopeptide repeat-containing protein A (Vpu-binding protein) (UBP). [Source:Uniprot/SWISSPROT;Acc:O43765]	19		
645202	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
64577	ENSG00000118514	aldehyde dehydrogenase 8A1 isoform 2 [Source:RefSeq_peptide;Acc:NP_739577]	6		
64801	ENSG00000173409	Protein ARV1 (hARV1). [Source:Uniprot/SWISSPROT;Acc:Q9H2C2]	1	1	
64849	ENSG00000158296	Solute carrier family 13 member 3 (Sodium-dependent high-affinity dicarboxylate transporter 2) (Na(+)/dicarboxylate cotransporter 3) (NaDC-3) (hNaDC3). [Source:Uniprot/SWISSPROT;Acc:Q8WWT9]	20	1	1
65018	ENSG00000158828	Serine/threonine-protein kinase PINK1, mitochondrial precursor (EC 2.7.11.1) (PTEN-induced putative kinase protein 1) (BRPK). [Source:Uniprot/SWISSPROT;Acc:Q9BXM7]	1		
653	ENSG00000112175	Bone morphogenetic protein 5 precursor (BMP-5). [Source:Uniprot/SWISSPROT;Acc:P22003]	6		1

654346	ENSG00000171916	galectin 9 like [Source:RefSeq_peptide;Acc:NP_001035167]	17		
6547	ENSG00000100678	Sodium/calcium exchanger 3 precursor (Na(+)/Ca(2+)-exchange protein 3). [Source:Uniprot/SWISSPROT;Acc:P57103]	14	1	1
6549	ENSG00000115616	Sodium/hydrogen exchanger 2 (Na(+)/H(+) exchanger 2) (NHE-2) (Solute carrier family 9 member 2). [Source:Uniprot/SWISSPROT;Acc:Q9UBY0]	2	1	1
6553	ENSG00000135740	Sodium/hydrogen exchanger 5 (Na(+)/H(+) exchanger 5) (NHE-5) (Solute carrier family 9 member 5). [Source:Uniprot/SWISSPROT;Acc:Q14940]	16	1	1
6578	ENSG00000174640	Solute carrier organic anion transporter family member 2A1 (Solute carrier family 21 member 2) (Prostaglandin transporter) (PGT). [Source:Uniprot/SWISSPROT;Acc:Q92959]	3	1	
6654	ENSG00000115904	Son of sevenless homolog 1 (SOS-1). [Source:Uniprot/SWISSPROT;Acc:Q07889]	2		
6670	ENSG00000172845	Transcription factor Sp3 (SPR-2). [Source:Uniprot/SWISSPROT;Acc:Q02447]	2		
6716	ENSG00000049319	3-oxo-5-alpha-steroid 4-dehydrogenase 2 (EC 1.3.99.5) (Steroid 5- alpha-reductase 2) (SR type 2) (5 alpha-SR2) (Type II 5-alpha reductase). [Source:Uniprot/SWISSPROT;Acc:P31213]	2	1	1

6738	ENSG00000116747	60 kDa SS-A/Ro ribonucleoprotein (60 kDa Ro protein) (60 kDa ribonucleoprotein Ro) (RoRNP) (Ro 60 kDa autoantigen) (TROVE domain family member 2) (Sjogren syndrome type A antigen) (SS-A) (Sjogren syndrome antigen A2). [Source:Uniprot/SWISSPROT;Acc:P10155]	1		
6753	ENSG00000183473	Somatostatin receptor type 3 (SS3R) (SSR-28). [Source:Uniprot/SWISSPROT;Acc:P32745]	22	1	
6780	ENSG00000124214	Double-stranded RNA-binding protein Staufien homolog 1. [Source:Uniprot/SWISSPROT;Acc:O95793]	20		
6782	ENSG00000155304	Stress 70 protein chaperone microsome-associated 60 kDa protein precursor (Microsomal stress 70 protein ATPase core). [Source:Uniprot/SWISSPROT;Acc:P48723]	21	1	1
688	ENSG00000102554	Krueppel-like factor 5 (Intestinal-enriched krueppel-like factor) (Colon krueppel-like factor) (Transcription factor BTEB2) (Basic transcription element-binding protein 2) (BTE-binding protein 2) (GC box-binding protein 2). [Source:Uniprot/SWISSPROT;Acc:Q13887]	13		
6899	ENSG00000184058	T-box transcription factor TBX1 (T-box protein 1) (Testis-specific T- box protein). [Source:Uniprot/SWISSPROT;Acc:O43435]	22		

6929	ENSG00000071564	Transcription factor E2-alpha (Immunoglobulin enhancer-binding factor E12/E47) (Transcription factor 3) (TCF-3) (Immunoglobulin transcription factor 1) (Transcription factor ITF-1) (Kappa-E2-binding factor). [Source:Uniprot/SWISSPROT;Acc:P15923]	19		
7035	ENSG0000003436	Tissue factor pathway inhibitor precursor (TFPI) (Lipoprotein- associated coagulation inhibitor) (LACI) (Extrinsic pathway inhibitor) (EPI). [Source:Uniprot/SWISSPROT;Acc:P10646]	2		1
7180	ENSG00000124490	Cysteine-rich secretory protein 2 precursor (CRISP-2) (Testis-specific protein TPX-1). [Source:Uniprot/SWISSPROT;Acc:P16562]	6		1
7225	ENSG00000137672	Short transient receptor potential channel 6 (TrpC6). [Source:Uniprot/SWISSPROT;Acc:Q9Y210]	11	1	
726	ENSG00000149260	Calpain-5 (EC 3.4.22.-) (nCL-3) (htra-3). [Source:Uniprot/SWISSPROT;Acc:O15484]	11		
727726	ENSG00000204155	Poly(ADP-ribose) glycohydrolase (EC 3.2.1.143). [Source:Uniprot/SWISSPROT;Acc:Q86W56]	10		
727778	ENSG00000162647	Uncharacterized protein C1orf119 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9NRX6]	1	1	1
727832	ENSG00000197414	CDNA FLJ36131 fis, clone TEST12024586.	15		

		[Source:Uniprot/SPTREMBL;Acc:Q8N9W7]			
727910	ENSG00000185561		17	1	
727957	ENSG00000179832	CDNA FLJ35542 fis, clone SPLEN2002917. [Source:Uniprot/SPTREMBL;Acc:Q8NAD1]	8		
729272	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
730027	ENSG00000197414	CDNA FLJ36131 fis, clone TESTI2024586. [Source:Uniprot/SPTREMBL;Acc:Q8N9W7]	15		
7341	ENSG00000116030	Small ubiquitin-related modifier 1 precursor (SUMO-1) (Sentrin) (Ubiquitin-like protein SMT3C) (SMT3 homolog 3) (Ubiquitin-homology domain protein PIC1) (Ubiquitin-like protein UBL1) (GAP-modifying protein 1) (GMP1). [Source:Uniprot/SWISSPROT;Acc:P63165]	2		
7345	ENSG00000154277	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (EC 6.-.-.-) (UCH-L1) (Ubiquitin thioesterase L1) (Neuron cytoplasmic protein 9.5) (PGP 9.5) (PGP9.5). [Source:Uniprot/SWISSPROT;Acc:P09936]	4		
7384	ENSG00000010256	Ubiquinol-cytochrome-c reductase complex core protein 1, mitochondrial precursor (EC 1.10.2.2) (Core I protein). [Source:Uniprot/SWISSPROT;Acc:P31930]	3		

7410	ENSG00000160293	Protein vav-2. [Source:Uniprot/SWISSPROT;Acc:P52735]	9		
7439	ENSG00000167995	Bestrophin-1 (Vitelliform macular dystrophy protein 2) (TU15B). [Source:Uniprot/SWISSPROT;Acc:O76090]	11	1	
745	ENSG00000124920	C11orf9. [Source:Uniprot/SPTREMBL;Acc:Q9P1Q6]	11		
7456	ENSG00000115935	WAS/WASL interacting protein family member 1 (Wiskott-Aldrich syndrome protein-interacting protein) (WASP-interacting protein) (PRPL-2 protein). [Source:Uniprot/SWISSPROT;Acc:O43516]	2		
7546	ENSG00000043355	Zinc finger protein ZIC 2 (Zinc finger protein of the cerebellum 2). [Source:Uniprot/SWISSPROT;Acc:O95409]	13		
757	ENSG00000142188	Transmembrane protein 50B (HCV p7-transregulated protein 3). [Source:Uniprot/SWISSPROT;Acc:P56557]	21	1	
7849	ENSG00000125618	Paired box protein Pax-8. [Source:Uniprot/SWISSPROT;Acc:Q06710]	2		
7855	ENSG00000163251	Frizzled-5 precursor (Fz-5) (hFz5) (FzE5). [Source:Uniprot/SWISSPROT;Acc:Q13467]	2	1	1
78987	ENSG00000163703	Cysteine-rich with EGF-like domain protein 1 precursor. [Source:Uniprot/SWISSPROT;Acc:Q96HD1]	3	1	1
79192	ENSG00000170549	Iroquois-class homeodomain protein IRX-1 (Iroquois homeobox protein 1) (Homeodomain protein IRXA1). [Source:Uniprot/SWISSPROT;Acc:P78414]	5		

7942	ENSG00000112561	Transcription factor EB. [Source:Uniprot/SWISSPROT;Acc:P19484]	6		
79603	ENSG00000090661	LAG1 longevity assurance homolog 4. [Source:Uniprot/SWISSPROT;Acc:Q9HA82]	19	1	
79618	ENSG00000147421	Homeobox-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q6NT76]	8		
79632	ENSG00000111879	Uncharacterized protein C6orf60. [Source:Uniprot/SWISSPROT;Acc:Q8NB25]	6		
79659	ENSG00000187240	dynein, cytoplasmic, heavy polypeptide 2 [Source:RefSeq_peptide;Acc:NP_001073932]	11		
79699	ENSG00000162378	zyg-11 homolog B [Source:RefSeq_peptide;Acc:NP_078922]	1		
79714	ENSG00000164051	coiled-coil domain containing 51 [Source:RefSeq_peptide;Acc:NP_078937]	3	1	
79784	ENSG00000105357	Myosin-14 (Myosin heavy chain 14) (Myosin heavy chain, nonmuscle IIc) (Nonmuscle myosin heavy chain IIc) (NMHC II-C). [Source:Uniprot/SWISSPROT;Acc:Q7Z406]	19		
79847	ENSG00000138111	Transmembrane protein 180. [Source:Uniprot/SWISSPROT;Acc:Q14CX5]	10	1	1
79884	ENSG00000164114	Microtubule-associated protein 9 (Aster-associated protein). [Source:Uniprot/SWISSPROT;Acc:Q49MG5]	4		

80010	ENSG00000178966	Protein RMI1 homolog (75 kDa BLM-associated protein) (BLAP75) (FAAP75). [Source:Uniprot/SWISSPROT;Acc:Q9H9A7]	9		
80034	ENSG00000178662	TGF-beta-induced apoptosis protein 2 (TAIP-2). [Source:Uniprot/SWISSPROT;Acc:Q8WYN3]	2		
80173	ENSG00000096872	Intraflagellar transport 74 homolog (Coiled-coil domain-containing protein 2) (Capillary morphogenesis protein 1) (CMG-1). [Source:Uniprot/SWISSPROT;Acc:Q96LB3]	9		
80176	ENSG00000171621	SPRY domain-containing SOCS box protein 1 (SSB-1). [Source:Uniprot/SWISSPROT;Acc:Q96BD6]	1		
80725	ENSG00000017373	p130Cas-associated protein (p140Cap) (SNAP-25-interacting protein) (SNIP). [Source:Uniprot/SWISSPROT;Acc:Q9C0H9]	17		
80821	ENSG00000100523	Probable phospholipase DDHD1 (EC 3.1.1.-) (DDHD domain protein 1) (Phosphatidic acid-preferring phospholipase A1 homolog) (PA-PLA1). [Source:Uniprot/SWISSPROT;Acc:Q8NEL9]	14		
80853	ENSG00000006459	JmjC domain-containing histone demethylation protein 1D. [Source:Uniprot/SWISSPROT;Acc:Q6ZMT4]	7		
81029	ENSG00000111186	Protein Wnt-5b precursor. [Source:Uniprot/SWISSPROT;Acc:Q9H1J7]	12		1
81553	ENSG00000197872	Protein FAM49A. [Source:Uniprot/SWISSPROT;Acc:Q9H0Q0]	2		

81611	ENSG00000143401	Acidic leucine-rich nuclear phosphoprotein 32 family member E (LANP- like protein) (LANP-L). [Source:Uniprot/SWISSPROT;Acc:Q9BTT0]	1		
81855	ENSG00000107819	Sideroflexin-3. [Source:Uniprot/SWISSPROT;Acc:Q9BWM7]	10	1	
8228	ENSG00000006757	Patatin-like phospholipase domain-containing protein 4 (Protein GS2). [Source:Uniprot/SWISSPROT;Acc:P41247]	X		1
823	ENSG00000014216	Calpain-1 catalytic subunit (EC 3.4.22.52) (Calpain-1 large subunit) (Calcium-activated neutral proteinase 1) (CANP 1) (Calpain mu-type) (muCANP) (Micromolar-calpain). [Source:Uniprot/SWISSPROT;Acc:P07384]	11		
8291	ENSG00000135636	Dysferlin (Dystrophy-associated fer-1-like protein) (Fer-1-like protein 1). [Source:Uniprot/SWISSPROT;Acc:O75923]	2	1	
83440	ENSG00000159322	ADP-dependent glucokinase (EC 2.7.1.147) (ADPGK) (ADP-GK) (RbBP-35). [Source:Uniprot/SWISSPROT;Acc:Q9BRR6]	15		1
83452	ENSG00000172007	Ras-related protein Rab-33B. [Source:Uniprot/SWISSPROT;Acc:Q9H082]	4		
835	ENSG00000106144	Caspase-2 precursor (EC 3.4.22.55) (CASP-2) (ICH-1 protease) (ICH- 1L/1S) [Contains: Caspase-2 subunit p18; Caspase-2 subunit p13; Caspase-2 subunit p12]. [Source:Uniprot/SWISSPROT;Acc:P42575]	7		
8364	ENSG00000197837	Histone H4. [Source:Uniprot/SWISSPROT;Acc:P62805]	12		

83658	ENSG00000125971	Dynein light chain roadblock-type 1 (Dynein light chain 2A, cytoplasmic) (Dynein-associated protein Km23) (Bithoraxoid-like protein) (BLP). [Source:Uniprot/SWISSPROT;Acc:Q9NP97]	20		
8366	ENSG00000197837	Histone H4. [Source:Uniprot/SWISSPROT;Acc:P62805]	12		
83860	ENSG00000165632	Transcription initiation factor TFIID subunit 3 (TBP-associated factor 3) (Transcription initiation factor TFIID 140 kDa subunit) (140 kDa TATA box-binding protein-associated factor) (TAF140) (TAFII140). [Source:Uniprot/SWISSPROT;Acc:Q5VWG9]	10		
83874	ENSG00000099992	TBC1 domain family member 10A (EBP50-PDX interactor of 64 kDa) (EPI64 protein). [Source:Uniprot/SWISSPROT;Acc:Q9BXI6]	22		
83959	ENSG00000088836	Sodium bicarbonate transporter-like protein 11 (Sodium borate cotransporter 1) (NaBC1) (Bicarbonate transporter-related protein 1) (Solute carrier family 4 member 11). [Source:Uniprot/SWISSPROT;Acc:Q8NBS3]	20	1	
84109	ENSG00000186867	Orexigenic neuropeptide QRFP receptor (G-protein coupled receptor 103) (SP9155) (AQ27). [Source:Uniprot/SWISSPROT;Acc:Q96P65]	4	1	
8411	ENSG00000102189	Early endosome antigen 1 (Endosome-associated protein p162) (Zinc finger FYVE domain-containing protein 2). [Source:Uniprot/SWISSPROT;Acc:Q15075]	12		

84239	ENSG00000127249	ATPase type 13A4 [Source:RefSeq_peptide;Acc:NP_115655]	3	1	
84255	ENSG00000157800	solute carrier family 37 (glycerol-3-phosphate transporter), member 3 isoform 2 [Source:RefSeq_peptide;Acc:NP_115671]	7	1	1
84272	ENSG00000119820	Protein YIPF4 (YIP1 family member 4). [Source:Uniprot/SWISSPROT;Acc:Q9BSR8]	2	1	
84293	ENSG00000122378	Uncharacterized protein C10orf58 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9BRX8]	10	1	
84328	ENSG00000162441	Protein LZIC (Leucine zipper and ICAT homologous domain-containing protein) (Leucine zipper and CTNNBIP1 domain-containing protein). [Source:Uniprot/SWISSPROT;Acc:Q8WZA0]	1		
844	ENSG00000143318	Calsequestrin-1 precursor (Calsequestrin, skeletal muscle isoform) (Calmitin). [Source:Uniprot/SWISSPROT;Acc:P31415]	1		1
8440	ENSG00000071051	Cytoplasmic protein NCK2 (NCK adaptor protein 2) (SH2/SH3 adaptor protein NCK-beta) (Nck-2). [Source:Uniprot/SWISSPROT;Acc:O43639]	2		
84457	ENSG00000165443	phytanoyl-CoA hydroxylase interacting protein-like [Source:RefSeq_peptide;Acc:NP_115815]	10		

84557	ENSG00000101460	Microtubule-associated proteins 1A/1B light chain 3A precursor (MAP1A/MAP1B LC3 A) (MAP1A/1B light chain 3 A) (MAP1 light chain 3- like protein 1) (Microtubule-associated protein 1 light chain 3 alpha) (Autophagy-related protein LC3 A) (Autophagy-related [Source:Uniprot/SWISSPROT;Acc:Q9H492]	20		
84570	ENSG00000188517	Collagen alpha-1(XXV) chain (CLAC-P) (Alzheimer disease amyloid- associated protein) (AMY) [Contains: Collagen-like Alzheimer amyloid plaque component (CLAC)]. [Source:Uniprot/SWISSPROT;Acc:Q9BXS0]	4	1	
84649	ENSG00000062282	Diacylglycerol O-acyltransferase 2 (EC 2.3.1.20) (Diglyceride acyltransferase 2). [Source:Uniprot/SWISSPROT;Acc:Q96PD7]	11	1	
84693	ENSG00000124150	Methylmalonyl-CoA epimerase, mitochondrial precursor (EC 5.1.99.1) (DL- methylmalonyl-CoA racemase). [Source:Uniprot/SWISSPROT;Acc:Q96PE7]	1		
84747	ENSG00000175970		12		
84803	ENSG00000138678	lysophosphatidic acid acyltransferase theta [Source:RefSeq_peptide;Acc:NP_116106]	4	1	1
8525	ENSG00000149091	Diacylglycerol kinase zeta (EC 2.7.1.107) (Diglyceride kinase zeta) (DGK- zeta) (DAG kinase zeta). [Source:Uniprot/SWISSPROT;Acc:Q13574]	11		

8531	ENSG00000060138	DNA-binding protein A (Cold shock domain-containing protein A) (Single-strand DNA-binding protein NF-GMB). [Source:Uniprot/SWISSPROT;Acc:P16989]	12		
85358	ENSG00000099882	SH3 and multiple ankyrin repeat domains protein 3 (Shank3) (Proline- rich synapse-associated protein 2) (ProSAP2) (Fragment). [Source:Uniprot/SWISSPROT;Acc:Q9BYB0]	22		
85462	ENSG00000137460		4		
8556	ENSG00000079335	Dual specificity protein phosphatase CDC14A (EC 3.1.3.48) (EC 3.1.3.16) (CDC14 cell division cycle 14 homolog A). [Source:Uniprot/SWISSPROT;Acc:Q9UNH5]	1		
8601	ENSG00000147509	Regulator of G-protein signaling 20 (RGS20) (Regulator of Gz-selective protein signaling 1) (Gz-selective GTPase-activating protein) (G(z)GAP). [Source:Uniprot/SWISSPROT;Acc:O76081]	8		
8630	ENSG00000025423	hydroxysteroid (17-beta) dehydrogenase 6 [Source:RefSeq_peptide;Acc:NP_003716]	12		
8790	ENSG00000116783	Serine/threonine-protein kinase TNNI3K (EC 2.7.11.1) (TNNI3- interacting kinase) (Cardiac ankyrin repeat kinase). [Source:Uniprot/SWISSPROT;Acc:Q59H18]	1		

88455	ENSG00000076513	Ankyrin repeat domain-containing protein 13A (Protein KE03). [Source:Uniprot/SWISSPROT;Acc:Q8IZ07]	12		
8861	ENSG00000198728	LIM domain-binding protein 1 (Carboxyl-terminal LIM domain-binding protein 2) (CLIM-2) (LIM domain-binding factor CLIM2). [Source:Uniprot/SWISSPROT;Acc:Q86U70]	10		
889	ENSG00000001631	Krev interaction trapped protein 1 (Krev interaction trapped 1) (Cerebral cavernous malformations 1 protein). [Source:Uniprot/SWISSPROT;Acc:O00522]	7		
8897	ENSG00000100330	Myotubularin-related protein 3 (EC 3.1.3.48) (FYVE domain-containing dual specificity protein phosphatase 1) (FYVE-DSP1) (Zinc finger FYVE domain-containing protein 10). [Source:Uniprot/SWISSPROT;Acc:Q13615]	22		
8908	ENSG00000056998	Glycogenin-2 (EC 2.4.1.186) (GN-2) (GN2). [Source:Uniprot/SWISSPROT;Acc:O15488]	X		
8916	ENSG00000138641	Probable E3 ubiquitin-protein ligase HERC3 (EC 6.3.2.-) (HECT domain and RCC1-like domain-containing protein 3). [Source:Uniprot/SWISSPROT;Acc:Q15034]	4		
8939	ENSG00000107164	Far upstream element-binding protein 3 (FUSE-binding protein 3). [Source:Uniprot/SWISSPROT;Acc:Q96I24]	9		

89795	ENSG00000067798	neuron navigator 3 [Source:RefSeq_peptide;Acc:NP_055718]	12		
89846	ENSG00000127084	FYVE, RhoGEF and PH domain-containing protein 3 (Zinc finger FYVE domain-containing protein 5). [Source:Uniprot/SWISSPROT;Acc:Q5JSP0]	9		
89853	ENSG00000196814	Protein FAM125B. [Source:Uniprot/SWISSPROT;Acc:Q9H7P6]	9		
89884	ENSG00000121454	LIM/homeobox protein Lhx4. [Source:Uniprot/SWISSPROT;Acc:Q969G2]	1		
90161	ENSG00000171004	Heparan-sulfate 6-O-sulfotransferase 2 (EC 2.8.2.-) (HS6ST-2). [Source:Uniprot/SWISSPROT;Acc:Q96MM7]	X	1	1
9023	ENSG00000138135	Cholesterol 25-hydroxylase (EC 1.14.99.38) (Cholesterol 25-monooxygenase) (h25OH). [Source:Uniprot/SWISSPROT;Acc:O95992]	10	1	
90249	ENSG00000113763	Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1). [Source:Uniprot/SWISSPROT;Acc:Q6ZN44]	5	1	1
9033	ENSG00000107593	Polycystic kidney disease 2-like 1 protein (Polycystin-L) (Polycystin- 2 homolog). [Source:Uniprot/SWISSPROT;Acc:Q9P0L9]	10	1	
90668	ENSG00000186648	Uncharacterized protein C14orf121. [Source:Uniprot/SPTREMBL;Acc:Q8ND23]	14		
9091	ENSG00000007541	Phosphatidylinositol N-acetylglucosaminyltransferase subunit Q (EC 2.4.1.198) (Phosphatidylinositol-glycan biosynthesis class Q protein) (PIG-Q) (N-acetylglucosamyl transferase component GPI1).	16	1	

		[Source:Uniprot/SWISSPROT;Acc:Q9BRB3]			
9153	ENSG00000137860	Sodium/nucleoside cotransporter 2 (Na(+)/nucleoside cotransporter 2) (Sodium-coupled nucleoside transporter 2) (Concentrative nucleoside transporter 2) (CNT 2) (hCNT2) (Sodium/purine nucleoside co- transporter) (SPNT). [Source:Uniprot/SWISSPROT;Acc:O43868]	15	1	
9187	ENSG00000074621	Sodium/potassium/calcium exchanger 1 (Na(+)/K(+)/Ca(2+)-exchange protein 1) (Retinal rod Na-Ca+K exchanger). [Source:Uniprot/SWISSPROT;Acc:O60721]	15	1	1
91947	ENSG00000140450	Arrestin domain-containing protein 4. [Source:Uniprot/SWISSPROT;Acc:Q8NCT1]	15		
9245	ENSG00000140297	glucosaminyl (N-acetyl) transferase 3, mucin type [Source:RefSeq_peptide;Acc:NP_004742]	15	1	1
9265	ENSG00000008256	Cytohesin-3 (PH, SEC7 and coiled-coil domain-containing protein 3) (ARF nucleotide-binding site opener 3) (Protein ARNO3) (General receptor of phosphoinositides 1) (Grp1). [Source:Uniprot/SWISSPROT;Acc:O43739]	7		
93	ENSG00000114739	Activin receptor type-2B precursor (EC 2.7.11.30) (Activin receptor type IIB) (ACTR-IIB). [Source:Uniprot/SWISSPROT;Acc:Q13705]	3	1	1

9312	ENSG00000182674	Potassium voltage-gated channel subfamily B member 2 (Voltage-gated potassium channel subunit Kv2.2). [Source:Uniprot/SWISSPROT;Acc:Q92953]	8	1	
9334	ENSG00000158470	Beta-1,4-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,4-GalTase 5) (Beta4Gal-T5) (b4Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-1,4-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,4-galactosyltransferase 5) (Beta-1,4-GalT II). [Source:Uniprot/SWISSPROT;Acc:O43286]	20	1	1
93408	ENSG00000106436	myosin light chain 2, precursor lymphocyte-specific [Source:RefSeq_peptide;Acc:NP_612412]	7		
9364	ENSG00000157869	Ras-related protein Rab-28 (Rab-26). [Source:Uniprot/SWISSPROT;Acc:P51157]	4		
9368	ENSG00000109062	Ezrin-radixin-moesin-binding phosphoprotein 50 (EBP50) (Na ⁺)/H ⁺ exchange regulatory cofactor NHE-RF) (NHERF-1) (Regulatory cofactor of Na ⁺)/H ⁺ exchanger) (Sodium-hydrogen exchanger regulatory factor 1) (Solute carrier family 9 isoform 3 regulatory f [Source:Uniprot/SWISSPROT;Acc:O14745]	17		
93953	ENSG00000147174	ACRC protein [Source:RefSeq_peptide;Acc:NP_443189]	X		

9397	ENSG00000152465	Glycylpeptide N-tetradecanoyltransferase 2 (EC 2.3.1.97) (Peptide N-myristoyltransferase 2) (Myristoyl-CoA:protein N-myristoyltransferase 2) (NMT 2) (Type II N-myristoyltransferase). [Source:Uniprot/SWISSPROT;Acc:O60551]	10		
9399	ENSG00000067221	Stomatin-like protein 1 (SLP-1) (Stomatin-related protein) (STORP) (EPB72-like 1) (UNC24 homolog). [Source:Uniprot/SWISSPROT;Acc:Q9UBI4]	15	1	
9411	ENSG00000137962	PTPL1-associated RhoGAP 1 [Source:RefSeq_peptide;Acc:NP_004806]	1		
9475	ENSG00000134318	Rho-associated protein kinase 2 (EC 2.7.11.1) (Rho-associated, coiled-coil-containing protein kinase 2) (p164 ROCK-2) (Rho kinase 2). [Source:Uniprot/SWISSPROT;Acc:O75116]	2		
948	ENSG00000135218	Platelet glycoprotein 4 (Platelet glycoprotein IV) (GPIV) (Glycoprotein IIIb) (GPIIIB) (Leukocyte differentiation antigen CD36) (CD36 antigen) (PAS IV) (PAS-4 protein) (Platelet collagen receptor) (Fatty acid translocase) (FAT) (Thrombospondin receptor). [Source:Uniprot/SWISSPROT;Acc:P16671]	7	1	
950	ENSG00000138760	Lysosome membrane protein 2 (Lysosome membrane protein II) (LIMP II) (Scavenger receptor class B member 2) (85 kDa lysosomal membrane sialoglycoprotein) (LGP85) (CD36 antigen-like 2). [Source:Uniprot/SWISSPROT;Acc:Q14108]	4	1	1

9647	ENSG00000100034	Ca(2+)/calmodulin-dependent protein kinase phosphatase (EC 3.1.3.16) (CaM-kinase phosphatase) (CaMKPase) (Partner of PIX 2) (hFEM-2) (Protein phosphatase 1F). [Source:Uniprot/SWISSPROT;Acc:P49593]	22		
9749	ENSG00000112419	Phosphatase and actin regulator 2. [Source:Uniprot/SWISSPROT;Acc:O75167]	6		
9767	ENSG00000102221	Protein Jade-3 (PHD finger protein 16). [Source:Uniprot/SWISSPROT;Acc:Q92613]	X		
9797	ENSG00000157014	TatD DNase domain-containing deoxyribonuclease 2 (EC 3.1.21.-). [Source:Uniprot/SWISSPROT;Acc:Q93075]	3		
9807	ENSG00000176095	Inositol hexaphosphate kinase 1 (EC 2.7.4.21) (InsP6 kinase 1) (Inositol hexakisphosphate kinase 1). [Source:Uniprot/SWISSPROT;Acc:Q92551]	3		
9863	ENSG00000187391	Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2 (Membrane-associated guanylate kinase inverted 2) (MAGI-2) (Atrophin-1-interacting protein 1) (Atrophin-1-interacting protein A). [Source:Uniprot/SWISSPROT;Acc:Q86UL8]	7		
9902	ENSG00000011028	Macrophage mannose receptor 2 precursor (Urokinase receptor-associated protein) (Endocytic receptor 180) (CD280 antigen). [Source:Uniprot/SWISSPROT;Acc:Q9UBG0]	17	1	1

9953	ENSG00000125430	Heparan sulfate glucosamine 3-O-sulfotransferase 3B1 (EC 2.8.2.30) (Heparan sulfate D-glucosaminyl 3-O-sulfotransferase 3B1) (Heparan sulfate 3-O-sulfotransferase 3B1) (h3-OST-3B). [Source:Uniprot/SWISSPROT;Acc:Q9Y662]	17	1	
9989	ENSG00000154845	Serine/threonine-protein phosphatase 4 regulatory subunit 1. [Source:Uniprot/SWISSPROT;Acc:Q8TF05]	18		
The folloing genes do not have Ensembl counterparts.					
727954	LOC388630	similar to C05G5.5 [<i>Homo sapiens</i>]	1p3 3		
10715	LASS1	LAG1 longevity assurance homolog 1; longevity assurance (LAG1, <i>S. cerevisiae</i>) homolog 1; longevity assurance gene 1; upstream of GDF1	19p 12		
137902	PXDNL	cardiac peroxidase; peroxidasin homolog-like	8q1 1.2 2		
23358	USP24	deubiquitinating enzyme 24; ubiquitin carboxyl-terminal	1p3		

		hydrolase 24; ubiquitin specific protease 24; ubiquitin thiolesterase 24; ubiquitin-specific processing protease 24	2.3		
56927	GPR108	G protein-coupled receptor 108 [<i>Homo sapiens</i>]	19p 13. 3		
729395	LOC729395	hypothetical protein LOC729395 [<i>Homo sapiens</i>]	8q2 4.2 2		
388630	LOC727954	hypothetical protein LOC727954 [<i>Homo sapiens</i>]	17q 24. 3		

3.5.2 Table for gene list generated using InParanoid

Table A.2. Gene list generated using InParanoid

B12 genes were labelled with green blocks. Gene for cblD was labelled with a yellow block. Chr = Chromosome; TM= Transmembrane Domain; SD= Singal Domain

Ensembl Gene ID	Description	Chr	TM	SD
ENSG00000006740	Rho GTPase-activating protein RICH2 (Rho-type GTPase-activating protein RICH2) (RhoGAP interacting with CIP4 homologs protein 2) (RICH-2) (NPC-A-10). [Source:Uniprot/SWISSPROT;Acc:Q17R89]	17		
ENSG00000008256	Cytohesin-3 (PH, SEC7 and coiled-coil domain-containing protein 3) (ARF nucleotide-binding site opener 3) (Protein ARNO3) (General receptor of phosphoinositides 1) (Grp1). [Source:Uniprot/SWISSPROT;Acc:O43739]	7		
ENSG00000012048	Breast cancer type 1 susceptibility protein (RING finger protein 53). [Source:Uniprot/SWISSPROT;Acc:P38398]	17		
ENSG00000013561	E3 ubiquitin-protein ligase RNF14 (EC 6.3.2.-) (RING finger protein 14) (Androgen receptor-associated protein 54) (Triad2 protein) (HFB30). [Source:Uniprot/SWISSPROT;Acc:Q9UBS8]	5		

ENSG00000032514	DNA excision repair protein ERCC-6 (EC 3.6.1.-) (ATP-dependent helicase ERCC6) (Cockayne syndrome protein CSB). [Source:Uniprot/SWISSPROT;Acc:Q03468]	10		
ENSG00000057468	MutS protein homolog 4. [Source:Uniprot/SWISSPROT;Acc:O15457]	1		
ENSG00000062598	Engulfment and cell motility protein 2 (CED-12 homolog A) (hCED-12A). [Source:Uniprot/SWISSPROT;Acc:Q96JJ3]	20		
ENSG00000063241	Isochorismatase domain-containing protein 2, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q96AB3]	19		
ENSG00000068878	Proteasome activator complex subunit 4 (Proteasome activator PA200). [Source:Uniprot/SWISSPROT;Acc:Q14997]	2		
ENSG00000069345	DnaJ homolog subfamily A member 2 (HIRA-interacting protein 4) (Cell cycle progression restoration gene 3 protein) (Dnj3) (Renal carcinoma antigen NY-REN-14). [Source:Uniprot/SWISSPROT;Acc:O60884]	16		
ENSG00000069998	Cat eye syndrome critical region protein 5 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9BXW7]	22		
ENSG00000072210	Fatty aldehyde dehydrogenase (EC 1.2.1.3) (Aldehyde dehydrogenase, microsomal) (Aldehyde dehydrogenase family 3 member A2) (Aldehyde dehydrogenase 10). [Source:Uniprot/SWISSPROT;Acc:P51648]	17	1	

ENSG00000075151	Eukaryotic translation initiation factor 4 gamma 3 (eIF-4-gamma 3) (eIF-4G 3) (eIF4G 3) (eIF-4-gamma II) (eIF4GII). [Source:Uniprot/SWISSPROT;Acc:O43432]	1		1
ENSG00000075651	Phospholipase D1 (EC 3.1.4.4) (PLD 1) (Choline phosphatase 1) (Phosphatidylcholine-hydrolyzing phospholipase D1) (hPLD1). [Source:Uniprot/SWISSPROT;Acc:Q13393]	3		
ENSG00000075975	Makorin-2 (RING finger protein 62). [Source:Uniprot/SWISSPROT;Acc:Q9H000]	3		
ENSG00000076248	Uracil-DNA glycosylase (EC 3.2.2.-) (UDG). [Source:Uniprot/SWISSPROT;Acc:P13051]	12		
ENSG00000077264	Serine/threonine-protein kinase PAK 3 (EC 2.7.11.1) (p21-activated kinase 3) (PAK-3) (Beta-PAK) (Oligophrenin-3). [Source:Uniprot/SWISSPROT;Acc:O75914]	X		
ENSG00000077522	Alpha-actinin-2 (Alpha actinin skeletal muscle isoform 2) (F-actin cross-linking protein). [Source:Uniprot/SWISSPROT;Acc:P35609]	1		
ENSG00000080345	Telomere-associated protein RIF1 (Rap1-interacting factor 1 homolog). [Source:Uniprot/SWISSPROT;Acc:Q5UIP0]	2		
ENSG00000080709	Small conductance calcium-activated potassium channel protein 2 (SK2). [Source:Uniprot/SWISSPROT;Acc:Q9H2S1]	5	1	
ENSG00000080815	Presenilin-1 (EC 3.4.23.-) (PS-1) (Protein S182) [Contains: Presenilin-1 NTF subunit; Presenilin-1 CTF subunit; Presenilin-1 CTF12 (PS1-CTF12)]. [Source:Uniprot/SWISSPROT;Acc:P49768]	14	1	

ENSG00000080986	Kinetochore protein Hec1 (HsHec1) (Kinetochore-associated protein 2) (Highly expressed in cancer protein) (Retinoblastoma-associated protein HEC). [Source:Uniprot/SWISSPROT;Acc:O14777]	18		
ENSG00000084110	Histidine ammonia-lyase (EC 4.3.1.3) (Histidase). [Source:Uniprot/SWISSPROT;Acc:P42357]	12		
ENSG00000085998	Protein O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase 1 (EC 2.4.1.-) (POMGnT1) (UDP-GlcNAc:alpha-D-mannoside beta-1,2-N- acetylglucosaminyltransferase I.2) (GnT I.2). [Source:Uniprot/SWISSPROT;Acc:Q8WZA1]	1	1	
ENSG00000086015	Microtubule-associated serine/threonine-protein kinase 2 (EC 2.7.11.1). [Source:Uniprot/SWISSPROT;Acc:Q6P0Q8]	1		
ENSG00000088386	Oligopeptide transporter, small intestine isoform (Peptide transporter 1) (Intestinal H(+)/peptide cotransporter) (Solute carrier family 15 member 1). [Source:Uniprot/SWISSPROT;Acc:P46059]	13	1	
ENSG00000088451	dTDP-D-glucose 4,6-dehydratase (EC 4.2.1.46). [Source:Uniprot/SWISSPROT;Acc:O95455]	13		
ENSG00000088812	Attractin precursor (Mahogany homolog) (DPPT-L). [Source:Uniprot/SWISSPROT;Acc:O75882]	20	1	
ENSG00000088836	Sodium bicarbonate transporter-like protein 11 (Sodium borate cotransporter 1) (NaBC1) (Bicarbonate transporter-related protein 1) (Solute carrier family 4 member 11). [Source:Uniprot/SWISSPROT;Acc:Q8NBS3]	20	1	
ENSG00000090686	Ubiquitin carboxyl-terminal hydrolase 48 (EC 3.1.2.15) (Ubiquitin thioesterase 48) (Ubiquitin-specific-processing protease 48) (Deubiquitinating enzyme 48). [Source:Uniprot/SWISSPROT;Acc:Q86UV5]	1	1	

ENSG00000091137	Pendrin (Sodium-independent chloride/iodide transporter) (Solute carrier family 26 member 4). [Source:Uniprot/SWISSPROT;Acc:O43511]	7	1	
ENSG00000096474	Protein G7d. [Source:Uniprot/SWISSPROT;Acc:Q5SSQ6]	c6 CO X		1
ENSG00000100034	Ca(2+)/calmodulin-dependent protein kinase phosphatase (EC 3.1.3.16) (CaM-kinase phosphatase) (CaMKPase) (Partner of PIX 2) (hFEM-2) (Protein phosphatase 1F). [Source:Uniprot/SWISSPROT;Acc:P49593]	22		
ENSG00000100201	Probable ATP-dependent RNA helicase DDX17 (EC 3.6.1.-) (DEAD box protein 17) (RNA-dependent helicase p72) (DEAD box protein p72). [Source:Uniprot/SWISSPROT;Acc:Q92841]	22		1
ENSG00000100209	Co-chaperone protein HscB, mitochondrial precursor (Hsc20) (DnaJ homolog subfamily C member 20). [Source:Uniprot/SWISSPROT;Acc:Q8IWL3]	22		
ENSG00000100258	transmembrane protein 112B (TMEM112B), mRNA [Source:RefSeq_dna;Acc:NM_033200]	22	1	1
ENSG00000100908	UPF0172 protein C14orf122. [Source:Uniprot/SWISSPROT;Acc:Q9Y3B6]	14		
ENSG00000101160	Cathepsin Z precursor (EC 3.4.22.-) (Cathepsin X) (Cathepsin P). [Source:Uniprot/SWISSPROT;Acc:Q9UBR2]	20	1	1

ENSG00000101182	Proteasome subunit alpha type 7 (EC 3.4.25.1) (Proteasome subunit RC6- 1) (Proteasome subunit XAPC7). [Source:Uniprot/SWISSPROT;Acc:O14818]	20		
ENSG00000101417	Peroxisomal membrane protein 4 (24 kDa peroxisomal intrinsic membrane protein). [Source:Uniprot/SWISSPROT;Acc:Q9Y6I8]	20		
ENSG00000101843	26S proteasome non-ATPase regulatory subunit 10 (26S proteasome regulatory subunit p28) (Gankyrin). [Source:Uniprot/SWISSPROT;Acc:O75832]	X		
ENSG00000102897	LYR motif-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:O43325]	16		
ENSG00000103089	fatty acid 2-hydroxylase [Source:RefSeq_peptide;Acc:NP_077282]	16	1	
ENSG00000103150	Malonyl-CoA decarboxylase, mitochondrial precursor (EC 4.1.1.9) (MCD). [Source:Uniprot/SWISSPROT;Acc:O95822]	16		
ENSG00000103226	Nodal modulator 3 precursor (pM5 protein 3). [Source:Uniprot/SWISSPROT;Acc:P69849]	16		1
ENSG00000103642	Serine beta-lactamase-like protein LACTB, mitochondrial precursor (EC 3.4.-.-). [Source:Uniprot/SWISSPROT;Acc:P83111]	15		
ENSG00000104442	Armadillo repeat-containing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9NVT9]	8		
ENSG00000104626	3' histone mRNA exonuclease 1 (EC 3.1.-.-) (3'-5' exonuclease ERI1) (Eri-1 homolog) (Histone mRNA 3' end-specific exoribonuclease) (Protein 3'hExo) (HEXO). [Source:Uniprot/SWISSPROT;Acc:Q8IV48]	8		

ENSG00000104687	Glutathione reductase, mitochondrial precursor (EC 1.8.1.7) (GR) (GRase). [Source:Uniprot/SWISSPROT;Acc:P00390]	8		1
ENSG00000104969	Small glutamine-rich tetratricopeptide repeat-containing protein A (Vpu-binding protein) (UBP). [Source:Uniprot/SWISSPROT;Acc:O43765]	19		
ENSG00000105135	ilvB (bacterial acetolactate synthase)-like isoform 1 [Source:RefSeq_peptide;Acc:NP_006835]	19	1	1
ENSG00000105983	Limb region 1 protein homolog (Differentiation-related gene 14 protein). [Source:Uniprot/SWISSPROT;Acc:Q8WVP7]	7	1	1
ENSG00000106638	Transducin beta-like 2 protein (WS beta-transducin repeats protein) (WS-betaTRP) (Williams-Beuren syndrome chromosome region 13 protein). [Source:Uniprot/SWISSPROT;Acc:Q9Y4P3]	7		
ENSG00000107201	Probable ATP-dependent RNA helicase DDX58 (EC 3.6.1.-) (DEAD-box protein 58) (Retinoic acid-inducible gene 1 protein) (RIG-1) (RIG-I). [Source:Uniprot/SWISSPROT;Acc:O95786]	9		
ENSG00000107537	Phytanoyl-CoA dioxygenase, peroxisomal precursor (EC 1.14.11.18) (Phytanoyl-CoA alpha-hydroxylase) (PhyH) (Phytanic acid oxidase). [Source:Uniprot/SWISSPROT;Acc:O14832]	10		
ENSG00000108239	TBC1 domain family member 12. [Source:Uniprot/SWISSPROT;Acc:O60347]	10		
ENSG00000108848	Cisplatin resistance-associated overexpressed protein (cAMP regulatory element-associated protein 1) (CRE-associated protein 1) (CREAP-1) (Luc7A) (Okadaic acid-inducible phosphoprotein OA48-18). [Source:Uniprot/SWISSPROT;Acc:O95232]	17		

ENSG00000109854	Oxidoreductase HTATIP2 (EC 1.1.1.-) (HIV-1 TAT-interactive protein 2) (30 kDa HIV-1 TAT-interacting protein). [Source:Uniprot/SWISSPROT;Acc:Q9BUP3]	11		
ENSG00000110315	RING finger protein 141 (Zinc finger protein 230). [Source:Uniprot/SWISSPROT;Acc:Q8WVD5]	11		
ENSG00000110455	1-aminocyclopropane-1-carboxylate synthase [Source:RefSeq_peptide;Acc:NP_115981]	11		
ENSG00000110660	solute carrier family 35, member F2 [Source:RefSeq_peptide;Acc:NP_059985]	11	1	
ENSG00000110880	Coronin-1C (Coronin-3) (hCRNN4). [Source:Uniprot/SWISSPROT;Acc:Q9ULV4]	12		
ENSG00000111727	Host cell factor 2 (HCF-2) (C2 factor). [Source:Uniprot/SWISSPROT;Acc:Q9Y5Z7]	12		
ENSG00000111885	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA (EC 3.2.1.113) (Processing alpha-1,2-mannosidase IA) (Alpha-1,2-mannosidase IA) (Mannosidase alpha class 1A member 1) (Man(9)-alpha-mannosidase) (Man9-mannosidase). [Source:Uniprot/SWISSPROT;Acc:P33908]	6	1	
ENSG00000112077	Rhesus blood group-associated glycoprotein (Rhesus blood group- associated ammonia channel) (Erythrocyte plasma membrane 50 kDa glycoprotein) (Rh50A) (CD241 antigen). [Source:Uniprot/SWISSPROT;Acc:Q02094]	6	1	1
ENSG00000112306	40S ribosomal protein S12. [Source:Uniprot/SWISSPROT;Acc:P25398]	6		
ENSG00000112992	NAD(P) transhydrogenase, mitochondrial precursor (EC 1.6.1.2) (Pyridine nucleotide transhydrogenase) (Nicotinamide nucleotide transhydrogenase). [Source:Uniprot/SWISSPROT;Acc:Q13423]	5	1	

ENSG00000113369	Arrestin domain-containing protein 3. [Source:Uniprot/SWISSPROT;Acc:Q96B67]	5		
ENSG00000113552	Glucosamine-6-phosphate isomerase (EC 3.5.99.6) (Glucosamine-6- phosphate deaminase) (GNPDA) (GlcN6P deaminase) (Oscillin). [Source:Uniprot/SWISSPROT;Acc:P46926]	5		
ENSG00000114054	Propionyl-CoA carboxylase beta chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit beta) (Propanoyl-CoA:carbon dioxide ligase subunit beta). [Source:Uniprot/SWISSPROT;Acc:P05166]	3		1
ENSG00000114166	Histone acetyltransferase PCAF (EC 2.3.1.48) (P300/CBP-associated factor) (P/CAF) (Histone acetylase PCAF). [Source:Uniprot/SWISSPROT;Acc:Q92831]	3		
ENSG00000115307	Ancient ubiquitous protein 1 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9Y679]	2		
ENSG00000115919	Kynureninase (EC 3.7.1.3) (L-kynurenine hydrolase). [Source:Uniprot/SWISSPROT;Acc:Q16719]	2		
ENSG00000116478	Histone deacetylase 1 (HD1). [Source:Uniprot/SWISSPROT;Acc:Q13547]	1		
ENSG00000116791	Quinone oxidoreductase (EC 1.6.5.5) (NADPH:quinone reductase) (Zeta- crystallin). [Source:Uniprot/SWISSPROT;Acc:Q08257]	1		
ENSG00000117758	Syntaxin-12. [Source:Uniprot/SWISSPROT;Acc:Q86Y82]	1	1	
ENSG00000118514	aldehyde dehydrogenase 8A1 isoform 2 [Source:RefSeq_peptide;Acc:NP_739577]	6		
ENSG00000118515	Serine/threonine-protein kinase Sgk1 (EC 2.7.11.1) (Serum/glucocorticoid-regulated kinase 1). [Source:Uniprot/SWISSPROT;Acc:O00141]	6		

ENSG00000118777	ATP-binding cassette sub-family G member 2 (Placenta-specific ATP- binding cassette transporter) (Breast cancer resistance protein) (Mitoxantrone resistance-associated protein) (CD338 antigen) (CDw338). [Source:Uniprot/SWISSPROT;Acc:Q9UNQ0]	4	1	
ENSG00000119688	ATP-binding cassette sub-family D member 4 (Peroxisomal membrane protein 69) (PMP69) (Peroxisomal membrane protein 1-like) (PXMP1-L) (P70R). [Source:Uniprot/SWISSPROT;Acc:O14678]	14	1	
ENSG00000119953	Survival of motor neuron-related-splicing factor 30 (SMN-related protein) (30 kDa splicing factor SMNrp) (Survival motor neuron domain- containing protein 1). [Source:Uniprot/SWISSPROT;Acc:O75940]	10		
ENSG00000120008	Bromodomain and WD repeat domain-containing protein 2 (WD repeat protein 11). [Source:Uniprot/SWISSPROT;Acc:Q9BZH6]	10		
ENSG00000121749	TBC1 domain family member 15. [Source:Uniprot/SWISSPROT;Acc:Q8TC07]	12		
ENSG00000123136	ATP-dependent RNA helicase DDX39 (EC 3.6.1.-) (DEAD box protein 39) (Nuclear RNA helicase URH49). [Source:Uniprot/SWISSPROT;Acc:O00148]	19		
ENSG00000123992	Aspartyl aminopeptidase (EC 3.4.11.21). [Source:Uniprot/SWISSPROT;Acc:Q9ULA0]	2		
ENSG00000124370	Methylmalonyl-CoA epimerase, mitochondrial precursor (EC 5.1.99.1) (DL-methylmalonyl-CoA racemase). [Source:Uniprot/SWISSPROT;Acc:Q96PE7]	2		

ENSG00000124496	Transcriptional-regulating factor 1 (Transcriptional-regulating protein 132) (Zinc finger transcription factor TReP-132) (Zinc finger protein rapa) (Breast cancer anti-estrogen resistance 2). [Source:Uniprot/SWISSPROT;Acc:Q96PN7]	6		
ENSG00000125246	citrate lyase beta like [Source:RefSeq_peptide;Acc:NP_996531]	13		1
ENSG00000127080	Inositol-pentakisphosphate 2-kinase (EC 2.7.1.158) (Inositol- 1,3,4,5,6-pentakisphosphate 2-kinase) (Ins(1,3,4,5,6)P5 2-kinase) (InsP5 2-kinase) (IPK1 homolog). [Source:Uniprot/SWISSPROT;Acc:Q9H8X2]	9		
ENSG00000127314	Ras-related protein Rap-1b precursor (GTP-binding protein smg p21B). [Source:Uniprot/SWISSPROT;Acc:P61224]	12		
ENSG00000127946	Huntingtin-interacting protein 1 (HIP-1). [Source:Uniprot/SWISSPROT;Acc:O00291]	7		
ENSG00000128923	family with sequence similarity 63, member B (FAM63B), transcript variant 1, mRNA [Source:RefSeq_dna;Acc:NM_001040450]	15		
ENSG00000129282	rRNA methyltransferase 1, mitochondrial precursor (EC 2.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q6IN84]	17		
ENSG00000129353	Choline transporter-like protein 2 (Solute carrier family 44 member 2). [Source:Uniprot/SWISSPROT;Acc:Q8IWA5]	19	1	

ENSG00000130699	Transcription initiation factor TFIID subunit 4 (TBP-associated factor 4) (Transcription initiation factor TFIID 135 kDa subunit) (TAF(II)135) (TAFII-135) (TAFII135) (TAFII-130) (TAFII130). [Source:Uniprot/SWISSPROT;Acc:O00268]	20		
ENSG00000130741	Eukaryotic translation initiation factor 2 subunit 3 (Eukaryotic translation initiation factor 2 subunit gamma) (eIF-2-gamma). [Source:Uniprot/SWISSPROT;Acc:P41091]	X		
ENSG00000131791	5'-AMP-activated protein kinase subunit beta-2 (AMPK beta-2 chain). [Source:Uniprot/SWISSPROT;Acc:O43741]	1		
ENSG00000132323	Integrin-linked kinase-associated serine/threonine phosphatase 2C (EC 3.1.3.16) (ILKAP). [Source:Uniprot/SWISSPROT;Acc:Q9H0C8]	2		
ENSG00000132773	Target of EGR1 protein 1. [Source:Uniprot/SWISSPROT;Acc:Q96GM8]	1		
ENSG00000132780	Nuclear autoantigenic sperm protein (NASP). [Source:Uniprot/SWISSPROT;Acc:P49321]	1		
ENSG00000132953	Exportin-4 (Exp4). [Source:Uniprot/SWISSPROT;Acc:Q9C0E2]	13		1
ENSG00000133243	BTB/POZ domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q9BX70]	19		
ENSG00000133315	Protein LRP16. [Source:Uniprot/SWISSPROT;Acc:Q9BQ69]	11		
ENSG00000133424	Glycosyltransferase-like protein LARGE1 (EC 2.4.-.-) (Acetylglucosaminyltransferase-like 1A). [Source:Uniprot/SWISSPROT;Acc:O95461]	22	1	1

ENSG00000133597	Uncharacterized aarF domain-containing protein kinase 2 (EC 2.7.-.-). [Source:Uniprot/SWISSPROT;Acc:Q7Z695]	7		
ENSG00000133872	Transmembrane protein 66 precursor (Protein FOAP-7) (HBV X- transactivated gene 3 protein). [Source:Uniprot/SWISSPROT;Acc:Q96BY9]	8	1	1
ENSG00000134242	Tyrosine-protein phosphatase non-receptor type 22 (EC 3.1.3.48) (Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP) (Lymphoid phosphatase) (LyP). [Source:Uniprot/SWISSPROT;Acc:Q9Y2R2]	1		
ENSG00000134716	Cytochrome P450 2J2 (EC 1.14.14.1) (CYP11J2) (Arachidonic acid epoxygenase). [Source:Uniprot/SWISSPROT;Acc:P51589]	1	1	1
ENSG00000134824	fatty acid desaturase 2 [Source:RefSeq_peptide;Acc:NP_004256]	11	1	
ENSG00000134996	Osteoclast-stimulating factor 1. [Source:Uniprot/SWISSPROT;Acc:Q92882]	9		
ENSG00000135241	intracellular membrane-associated calcium-independent phospholipase A2 gamma [Source:RefSeq_peptide;Acc:NP_056538]	7		
ENSG00000135720	Cytoplasmic dynein 1 light intermediate chain 2 (Dynein light intermediate chain 2, cytosolic) (LIC53/55) (LIC-2). [Source:Uniprot/SWISSPROT;Acc:O43237]	16		
ENSG00000136490	LIM domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q9BT23]	17		

ENSG00000136492	Fanconi anemia group J protein (EC 3.6.1.-) (ATP-dependent RNA helicase BRIP1) (Protein FACJ) (BRCA1-interacting protein C-terminal helicase 1) (BRCA1-interacting protein 1) (BRCA1-associated C-terminal helicase 1). [Source:Uniprot/SWISSPROT;Acc:Q9BX63]	17		
ENSG00000136527	Arginine/serine-rich-splicing factor 10 (Transformer-2-beta) (HTRA2- beta) (Transformer 2 protein homolog). [Source:Uniprot/SWISSPROT;Acc:P62995]	3		
ENSG00000137628	CDNA FLJ10787 fis, clone NT2RP4000481, weakly similar to ATP-DEPENDENT RNA HELICASE DOB1. (Fragment). [Source:Uniprot/SPTREMBL;Acc:Q9NVE3]	4		
ENSG00000137770	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase like 2 [Source:RefSeq_peptide;Acc:NP_057480]	15		
ENSG00000137976	Deoxyribonuclease-2-beta precursor (EC 3.1.22.1) (Deoxyribonuclease II beta) (DNase II beta) (DNase2-like acid DNase) (DNase II-like acid DNase) (Endonuclease DLAD). [Source:Uniprot/SWISSPROT;Acc:Q8WZ79]	1	1	1
ENSG00000138449	Solute carrier family 40 member 1 (Ferroportin-1) (Iron-regulated transporter 1). [Source:Uniprot/SWISSPROT;Acc:Q9NP59]	2	1	
ENSG00000138613	Gamma-secretase subunit APH-1B (APH-1b) (Aph-1beta) (Presenilin- stabilization factor-like). [Source:Uniprot/SWISSPROT;Acc:Q8WW43]	15	1	1

ENSG00000138760	Lysosome membrane protein 2 (Lysosome membrane protein II) (LIMP II) (Scavenger receptor class B member 2) (85 kDa lysosomal membrane sialoglycoprotein) (LGP85) (CD36 antigen-like 2). [Source:Uniprot/SWISSPROT;Acc:Q14108]	4	1	1
ENSG00000138796	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial precursor (EC 1.1.1.35) (Short chain 3-hydroxyacyl-CoA dehydrogenase) (HCDH) (Medium and short chain L-3-hydroxyacyl-coenzyme A dehydrogenase). [Source:Uniprot/SWISSPROT;Acc:Q16836]	4		
ENSG00000138829	Fibrillin-2 precursor. [Source:Uniprot/SWISSPROT;Acc:P35556]	5	1	1
ENSG00000138869	CHCH domain-containing protein C22orf16, mitochondrial precursor (Protein N27C7-4). [Source:Uniprot/SWISSPROT;Acc:Q8WYQ3]	22		
ENSG00000139133	Alpha-1,2-glucosyltransferase ALG10-A (EC 2.4.1.-) (Alpha-2- glucosyltransferase ALG10-A) (Asparagine-linked glycosylation protein 10 homolog A). [Source:Uniprot/SWISSPROT;Acc:Q5BKT4]	12	1	1
ENSG00000139344	Probable imidazolonepropionase (EC 3.5.2.7) (Amidohydrolase domain- containing protein 1). [Source:Uniprot/SWISSPROT;Acc:Q96NU7]	12		
ENSG00000139428	Cob(I)yrinic acid a,c-diamide adenosyltransferase, mitochondrial precursor (EC 2.5.1.17) (Cob(I)alamin adenosyltransferase) (Methylmalonic aciduria type B protein). [Source:Uniprot/SWISSPROT;Acc:Q96EY8]	7		

ENSG00000140043	Zinc-binding alcohol dehydrogenase domain-containing protein 1 (EC 1.-.-.-). [Source:Uniprot/SWISSPROT;Acc:Q8N8N7]	14		
ENSG00000140612	Signal peptidase complex catalytic subunit SEC11A (EC 3.4.-.-) (SEC11 homolog A) (SEC11-like protein 1) (Microsomal signal peptidase 18 kDa subunit) (SPase 18 kDa subunit) (SPC18) (Endopeptidase SP18). [Source:Uniprot/SWISSPROT;Acc:P67812]	15		
ENSG00000141084	RAN binding protein 10 [Source:RefSeq_peptide;Acc:NP_065901]	16		
ENSG00000141552	Anaphase-promoting complex subunit 11 (APC11) (Cyclosome subunit 11) (Hepatocellular carcinoma-associated RING finger protein). [Source:Uniprot/SWISSPROT;Acc:Q9NYG5]	17		
ENSG00000142188	Transmembrane protein 50B (HCV p7-transregulated protein 3). [Source:Uniprot/SWISSPROT;Acc:P56557]	21	1	
ENSG00000143499	SET and MYND domain-containing protein 2 (HSKM-B). [Source:Uniprot/SWISSPROT;Acc:Q9NRG4]	1		
ENSG00000143748	Nuclear valosin-containing protein-like (Nuclear VCP-like protein) (NVLp). [Source:Uniprot/SWISSPROT;Acc:O15381]	1		
ENSG00000144040	Sideroflexin-5. [Source:Uniprot/SWISSPROT;Acc:Q8TD22]	2	1	
ENSG00000144231	DNA-directed RNA polymerase II 16 kDa polypeptide (EC 2.7.7.6) (RPB4). [Source:Uniprot/SWISSPROT;Acc:O15514]	2		

ENSG00000144959	Arylacetamide deacetylase-like 1 (EC 3.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q6PIU2]	3	1	
ENSG00000145740	zinc transporter ZTL1 isoform 1 [Source:RefSeq_peptide;Acc:NP_075053]	5	1	
ENSG00000146021	Kelch-like protein 3. [Source:Uniprot/SWISSPROT;Acc:Q9UH77]	5		
ENSG00000146085	Methylmalonyl-CoA mutase, mitochondrial precursor (EC 5.4.99.2) (MCM) (Methylmalonyl-CoA isomerase). [Source:Uniprot/SWISSPROT;Acc:P22033]	6		
ENSG00000146151	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase-like 1 isoform b [Source:RefSeq_peptide;Acc:NP_001035865]	6		
ENSG00000146216	Tau-tubulin kinase 1 (EC 2.7.11.1) (Brain-derived tau kinase). [Source:Uniprot/SWISSPROT;Acc:Q5TCY1]	6		
ENSG00000146409	OTTHUMP00000017233. [Source:Uniprot/SPTREMBL;Acc:Q5TBJ2]	6	1	
ENSG00000147010	SH3 domain-containing kinase-binding protein 1 (Cbl-interacting protein of 85 kDa) (Human Src-family kinase-binding protein 1) (HSB-1) (CD2-binding protein 3) (CD2BP3). [Source:Uniprot/SWISSPROT;Acc:Q96B97]	X		
ENSG00000147251	dedicator of cytokinesis 11 [Source:RefSeq_peptide;Acc:NP_653259]	X		
ENSG00000147669	DNA-directed RNA polymerases I, II, and III 7.0 kDa polypeptide (EC 2.7.7.6) (ABC10-alpha) (RPB7.0) (RPB10alpha) (RPABC4). [Source:Uniprot/SWISSPROT;Acc:P53803]	8		

ENSG00000147955	Sigma 1-type opioid receptor (Sigma1R) (Sigma1-receptor) (hSigmaR1) (SIG-1R) (SR31747-binding protein) (SR-BP) (Aging-associated gene 8 protein). [Source:Uniprot/SWISSPROT;Acc:Q99720]	9	1	1
ENSG00000148057	gluconokinase-like protein [Source:RefSeq_peptide;Acc:NP_001001551]	9		
ENSG00000148450	Methionine-R-sulfoxide reductase B2 (EC 1.8.4.-). [Source:Uniprot/SWISSPROT;Acc:Q9Y3D2]	10		1
ENSG00000149476	Dihydroxyacetone kinase (EC 2.7.1.29) (Glycerone kinase) (DHA kinase). [Source:Uniprot/SWISSPROT;Acc:Q3LXA3]	11		
ENSG00000151229	Proton myo-inositol cotransporter (H(+)-myo-inositol cotransporter) (Hmit). [Source:Uniprot/SWISSPROT;Acc:Q96QE2]	12	1	
ENSG00000151498	Acyl-CoA dehydrogenase family member 8, mitochondrial precursor (EC 1.3.99.-) (ACAD-8) (Isobutyryl-CoA dehydrogenase) (Activator- recruited cofactor 42 kDa component) (ARC42). [Source:Uniprot/SWISSPROT;Acc:Q9UKU7]	11		
ENSG00000151503	Condensin-II complex subunit D3 (Non-SMC condensin II complex subunit D3) (hCAP-D3). [Source:Uniprot/SWISSPROT;Acc:P42695]	11		
ENSG00000151611	Methylmalonic aciduria type A protein, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q8IVH4]	4		

ENSG00000151729	ADP/ATP translocase 1 (Adenine nucleotide translocator 1) (ANT 1) (ADP,ATP carrier protein 1) (Solute carrier family 25 member 4) (ADP,ATP carrier protein, heart/skeletal muscle isoform T1). [Source:Uniprot/SWISSPROT;Acc:P12235]	4	1	
ENSG00000153086	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase (EC 4.1.1.45). [Source:Uniprot/SWISSPROT;Acc:Q8TDX5]	2	1	
ENSG00000154059	Impact homolog [Source:RefSeq_peptide;Acc:NP_060909]	18		
ENSG00000154845	Serine/threonine-protein phosphatase 4 regulatory subunit 1. [Source:Uniprot/SWISSPROT;Acc:Q8TF05]	18		
ENSG00000155111	Cell division cycle 2-like protein kinase 6 (EC 2.7.11.22) (CDC2- related protein kinase 6) (Death-preventing kinase) (Cyclin-dependent kinase 11). [Source:Uniprot/SWISSPROT;Acc:Q9BWU1]	6		
ENSG00000155158	C9orf52 protein (Chromosome 9 open reading frame 52). [Source:Uniprot/SPTREMBL;Acc:Q8IXZ6]	9		1
ENSG00000155368	Acyl-CoA-binding protein (ACBP) (Diazepam-binding inhibitor) (DBI) (Endozepine) (EP). [Source:Uniprot/SWISSPROT;Acc:P07108]	2		
ENSG00000155660	Protein disulfide-isomerase A4 precursor (EC 5.3.4.1) (Protein ERp-72) (ERp72). [Source:Uniprot/SWISSPROT;Acc:P13667]	7	1	1
ENSG00000155827	E3 ubiquitin-protein ligase BRE1A (EC 6.3.2.-) (BRE1-A) (hBRE1) (RING finger protein 20). [Source:Uniprot/SWISSPROT;Acc:Q5VTR2]	9		

ENSG00000155980	Kinesin heavy chain isoform 5A (Neuronal kinesin heavy chain) (NKHC) (Kinesin heavy chain neuron-specific 1). [Source:Uniprot/SWISSPROT;Acc:Q12840]	12		
ENSG00000156802	ATPase family AAA domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q6PL18]	8		
ENSG00000157014	TatD DNase domain-containing deoxyribonuclease 2 (EC 3.1.21.-). [Source:Uniprot/SWISSPROT;Acc:Q93075]	3		
ENSG00000157181	Uncharacterized protein C1orf27 (Odorant response abnormal 4 protein) (LAG1-interacting protein). [Source:Uniprot/SPTREMBL;Acc:Q5SWX8]	1	1	
ENSG00000157353	L-fucose kinase (EC 2.7.1.52) (Fucokinase). [Source:Uniprot/SWISSPROT;Acc:Q8N0W3]	16		
ENSG00000157483	Myosin-Ie (Myosin-Ie). [Source:Uniprot/SWISSPROT;Acc:Q12965]	15		
ENSG00000158315	Rhomboid-related protein 2 (EC 3.4.21.105) (RRP2) (Rhomboid-like protein 2). [Source:Uniprot/SWISSPROT;Acc:Q9NX52]	1	1	
ENSG00000159199	ATP synthase lipid-binding protein, mitochondrial precursor (EC 3.6.3.14) (ATP synthase proteolipid P1) (ATPase protein 9) (ATPase subunit C). [Source:Uniprot/SWISSPROT;Acc:P05496]	17	1	
ENSG00000159461	Autocrine motility factor receptor, isoform 2 (EC 6.3.2.-) (AMF receptor) (gp78). [Source:Uniprot/SWISSPROT;Acc:Q9UKV5]	16		
ENSG00000159650	Probable urocanate hydratase (EC 4.2.1.49) (Urocanase) (Imidazolonepropionate hydrolase). [Source:Uniprot/SWISSPROT;Acc:Q96N76]	3		

ENSG00000159733	Zinc finger FYVE domain-containing protein 28. [Source:Uniprot/SWISSPROT;Acc:Q9HCC9]	4		
ENSG00000160179	ATP-binding cassette sub-family G member 1 (White protein homolog) (ATP-binding cassette transporter 8). [Source:Uniprot/SWISSPROT;Acc:P45844]	21	1	
ENSG00000161048	N-acyl-phosphatidylethanolamine-hydrolyzing phospholipase D [Source:RefSeq_peptide;Acc:NP_945341]	7		
ENSG00000162441	Protein LZIC (Leucine zipper and ICAT homologous domain-containing protein) (Leucine zipper and CTNNBIP1 domain-containing protein). [Source:Uniprot/SWISSPROT;Acc:Q8WZA0]	1		
ENSG00000162542	transmembrane and coiled-coil domains 4 [Source:RefSeq_peptide;Acc:NP_859070]	1	1	
ENSG00000162591	Multiple epidermal growth factor-like domains 6 precursor (EGF-like domain-containing protein 3) (Multiple EGF-like domain protein 3). [Source:Uniprot/SWISSPROT;Acc:O75095]	1	1	1
ENSG00000162819	Uncharacterized protein C1orf58. [Source:Uniprot/SPTREMBL;Acc:Q5VW32]	1		
ENSG00000162882	3-hydroxyanthranilate 3,4-dioxygenase (EC 1.13.11.6) (3-HAO) (3- hydroxyanthranilic acid dioxygenase) (3-hydroxyanthranilate oxygenase). [Source:Uniprot/SWISSPROT;Acc:P46952]	2		
ENSG00000163050	Chaperone-activity of bc1 complex-like, mitochondrial precursor (Chaperone-ABC1-like) (aarF domain-containing protein kinase 3). [Source:Uniprot/SWISSPROT;Acc:Q8NI60]	1		
ENSG00000163170	BolA-like protein 3. [Source:Uniprot/SWISSPROT;Acc:Q53S33]	2		

ENSG00000163516	Ankyrin repeat and zinc finger domain-containing protein 1 (Zinc finger protein 744). [Source:Uniprot/SWISSPROT;Acc:Q9H8Y5]	2		
ENSG00000163624	Phosphatidate cytidylyltransferase 1 (EC 2.7.7.41) (CDP-diglyceride synthetase 1) (CDP-diglyceride pyrophosphorylase 1) (CDP- diacylglycerol synthase 1) (CDS 1) (CTP:phosphatidate cytidylyltransferase 1) (CDP-DAG synthase 1) (CDP-DG synthetase 1). [Source:Uniprot/SWISSPROT;Acc:Q92903]	4	1	
ENSG00000163945		4		
ENSG00000164062	Acylamino-acid-releasing enzyme (EC 3.4.19.1) (AARE) (Acyl-peptide hydrolase) (APH) (Acylaminoacyl-peptidase) (Oxidized protein hydrolase) (OPH) (DNF15S2 protein). [Source:Uniprot/SWISSPROT;Acc:P13798]	3		
ENSG00000164104	High mobility group protein B2 (High mobility group protein 2) (HMG- 2). [Source:Uniprot/SWISSPROT;Acc:P26583]	4		
ENSG00000164124	transmembrane protein 144 (TMEM144), mRNA [Source:RefSeq_dna;Acc:NM_018342]	4	1	
ENSG00000164535	Sn1-specific diacylglycerol lipase beta (EC 3.1.1.-) (DGL-beta) (KCCR13L). [Source:Uniprot/SWISSPROT;Acc:Q8NCG7]	7	1	1
ENSG00000164744	Sad1 and UNC84 domain containing 1 [Source:RefSeq_peptide;Acc:NP_689995]	7		

ENSG00000165029	ATP-binding cassette sub-family A member 1 (ATP-binding cassette transporter 1) (ATP-binding cassette 1) (ABC-1) (Cholesterol efflux regulatory protein). [Source:Uniprot/SWISSPROT;Acc:O95477]	9	1	1
ENSG00000165097	Flavin-containing amine oxidase domain-containing protein 1 (EC 1.-.-.-). [Source:Uniprot/SWISSPROT;Acc:Q8NB78]	6		
ENSG00000165494	Pre-mRNA cleavage complex 2 protein Pcf11 (Pre-mRNA cleavage complex II protein Pcf11) (Fragment). [Source:Uniprot/SWISSPROT;Acc:O94913]	11		
ENSG00000165526	RNA pseudouridylate synthase domain containing 4 [Source:RefSeq_peptide;Acc:NP_116184]	11		
ENSG00000165609	ADP-sugar pyrophosphatase (EC 3.6.1.13) (EC 3.6.1.-) (Nucleoside diphosphate-linked moiety X motif 5) (Nudix motif 5) (YSA1H). [Source:Uniprot/SWISSPROT;Acc:Q9UKK9]	10		
ENSG00000165644	Catechol-O-methyltransferase domain-containing protein 1 (EC 2.1.1.-). [Source:Uniprot/SWISSPROT;Acc:Q86VU5]	10	1	1
ENSG00000166128	Ras-related protein Rab-8B. [Source:Uniprot/SWISSPROT;Acc:Q92930]	15		
ENSG00000166530	Heat shock factor-binding protein 1 (Nasopharyngeal carcinoma- associated antigen 13) (NPC-A-13). [Source:Uniprot/SWISSPROT;Acc:O75506]	16		
ENSG00000166582	Proline-rich protein 6 (Nuclear protein p30). [Source:Uniprot/SWISSPROT;Acc:Q7Z7K6]	17		
ENSG00000166816	Probable D-lactate dehydrogenase, mitochondrial precursor (EC 1.1.2.4) (Lactate dehydrogenase D) (DLD). [Source:Uniprot/SWISSPROT;Acc:Q86WU2]	16		

ENSG00000167258	Cell division cycle 2-related protein kinase 7 (EC 2.7.11.22) (CDC2- related protein kinase 7) (Cdc2-related kinase, arginine/serine-rich) (CrkRS). [Source:Uniprot/SWISSPROT;Acc:Q9NYV4]	17		
ENSG00000167306	Myosin-Vb. [Source:Uniprot/SWISSPROT;Acc:Q9ULV0]	18		
ENSG00000167548	Myeloid/lymphoid or mixed-lineage leukemia protein 2 (ALL1-related protein). [Source:Uniprot/SWISSPROT;Acc:O14686]	12		
ENSG00000167619	Transmembrane protein 145. [Source:Uniprot/SWISSPROT;Acc:Q8NBT3]	19	1	1
ENSG00000167900	Thymidine kinase, cytosolic (EC 2.7.1.21). [Source:Uniprot/SWISSPROT;Acc:P04183]	17		
ENSG00000168061	SAC3 domain containing 1 [Source:RefSeq_peptide;Acc:NP_037431]	11		
ENSG00000168288	Uncharacterized protein C2orf25, mitochondrial precursor. [Source:Uniprot/SWISSPROT;Acc:Q9H3L0]	2		
ENSG00000169299	Phosphoglucomutase-2 (EC 5.4.2.2) (Glucose phosphomutase 2) (PGM 2). [Source:Uniprot/SWISSPROT;Acc:Q96G03]	4		
ENSG00000169925	Bromodomain-containing protein 3 (RING3-like protein). [Source:Uniprot/SWISSPROT;Acc:Q15059]	9		
ENSG00000170144	Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3). [Source:Uniprot/SWISSPROT;Acc:P51991]	2		
ENSG00000171928	Protein FAM18B. [Source:Uniprot/SWISSPROT;Acc:Q9NYZ1]	17	1	
ENSG00000172009	Thimet oligopeptidase (EC 3.4.24.15) (Endopeptidase 24.15) (MP78). [Source:Uniprot/SWISSPROT;Acc:P52888]	19		

ENSG00000172115	Cytochrome c. [Source:Uniprot/SWISSPROT;Acc:P99999]	7		
ENSG00000172869	Protein DmX-like 1 (X-like 1 protein). [Source:Uniprot/SWISSPROT;Acc:Q9Y485]	5		
ENSG00000172954	lysocardiolipin acyltransferase isoform 1 [Source:RefSeq_peptide;Acc:NP_872357]	2	1	
ENSG00000173473	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily C member 1 (SWI/SNF complex 155 kDa subunit) (BRG1-associated factor 155). [Source:Uniprot/SWISSPROT;Acc:Q92922]	3		
ENSG00000174851	Protein YIF1A (YIP1-interacting factor homolog A) (54TmP). [Source:Uniprot/SWISSPROT;Acc:O95070]	11	1	
ENSG00000175198	Propionyl-CoA carboxylase alpha chain, mitochondrial precursor (EC 6.4.1.3) (PCCase subunit alpha) (Propanoyl-CoA:carbon dioxide ligase subunit alpha). [Source:Uniprot/SWISSPROT;Acc:P05165]	13		
ENSG00000175575	WD repeat protein 71 (Protein G-16). [Source:Uniprot/SWISSPROT;Acc:Q9BRP4]	11		
ENSG00000175711	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1 [Source:RefSeq_peptide;Acc:NP_001009905]	17		
ENSG00000176623	Protein FAM82B. [Source:Uniprot/SWISSPROT;Acc:Q96DB5]	8		
ENSG00000176946	THAP domain-containing protein 4. [Source:Uniprot/SWISSPROT;Acc:Q8WY91]	2		
ENSG00000177963	Synembryn-A (Protein Ric-8A). [Source:Uniprot/SWISSPROT;Acc:Q9NPQ8]	11		
ENSG00000178075	GRAM domain containing 1C [Source:RefSeq_peptide;Acc:NP_060047]	3	1	

ENSG00000178381	AN1-type zinc finger protein 2A. [Source:Uniprot/SWISSPROT;Acc:Q8N6M9]	7		
ENSG00000179761	Peroxisomal sarcosine oxidase (EC 1.5.3.1) (PSO) (L-pipecolate oxidase) (EC 1.5.3.7) (L-pipecolic acid oxidase). [Source:Uniprot/SWISSPROT;Acc:Q9P0Z9]	17	1	
ENSG00000180817	Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho- hydrolase) (PPase). [Source:Uniprot/SWISSPROT;Acc:Q15181]	10		
ENSG00000180902	D-2-hydroxyglutarate dehydrogenase, mitochondrial precursor (EC 1.1.99.-). [Source:Uniprot/SWISSPROT;Acc:Q8N465]	2		1
ENSG00000181413	HESB like domain containing 2. [Source:Uniprot/SPTREMBL;Acc:Q5TBE2]	5		
ENSG00000182054	Isocitrate dehydrogenase [NADP], mitochondrial precursor (EC 1.1.1.42) (Oxalosuccinate decarboxylase) (IDH) (NADP(+)-specific ICDH) (IDP) (ICD-M). [Source:Uniprot/SWISSPROT;Acc:P48735]	15		
ENSG00000182544	Major facilitator superfamily domain-containing protein 5. [Source:Uniprot/SWISSPROT;Acc:Q6N075]	12	1	1
ENSG00000184575	Exportin-T (tRNA exportin) (Exportin(tRNA)). [Source:Uniprot/SWISSPROT;Acc:O43592]	12		
ENSG00000184840	Transmembrane emp24 domain-containing protein 9 precursor (Glycoprotein 25L2). [Source:Uniprot/SWISSPROT;Acc:Q9BVK6]	5	1	1
ENSG00000185009	AP-3 complex subunit mu-1 (Adapter-related protein complex 3 mu-1 subunit) (Mu-adaptin 3A) (AP-3 adapter complex mu3A subunit). [Source:Uniprot/SWISSPROT;Acc:Q9Y2T2]	10		

ENSG00000185068	TFIIH basal transcription factor complex TTD-A subunit (General transcription factor IIH polypeptide 5) (TFB5 ortholog). [Source:Uniprot/SWISSPROT;Acc:Q6ZYL4]	6		
ENSG00000187239	Formin-binding protein 1 (Formin-binding protein 17) (hFBP17). [Source:Uniprot/SWISSPROT;Acc:Q96RU3]	9		
ENSG00000187546	CDNA FLJ16237 fis, clone HCASM2002754 (Hypothetical protein FLJ16237). [Source:Uniprot/SPTREMBL;Acc:Q6ZNB7]	7	1	
ENSG00000188014	prematurely terminated mRNA decay factor-like [Source:RefSeq_peptide;Acc:NP_619643]	4		
ENSG00000188167	similar to F40B5.2b (FLJ45032), mRNA [Source:RefSeq_dna;Acc:NM_001039770]	3	1	
ENSG00000188493	CDNA FLJ41131 fis, clone BRACE2024627. [Source:Uniprot/SPTREMBL;Acc:Q6ZWG5]	19		
ENSG00000189201	Cytochrome c oxidase assembly protein COX19 (hCOX19). [Source:Uniprot/SWISSPROT;Acc:Q49B96]	7		
ENSG00000196642	Putative GTP-binding protein Parf (Partner of ARF). [Source:Uniprot/SWISSPROT;Acc:Q3YEC7]	9		
ENSG00000196839	Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase). [Source:Uniprot/SWISSPROT;Acc:P00813]	20		
ENSG00000197408	Cytochrome P450 2B6 (EC 1.14.14.1) (CYPIIB6) (P450 IIB1). [Source:Uniprot/SWISSPROT;Acc:P20813]	19		1
ENSG00000197915	Repetin. [Source:Uniprot/SWISSPROT;Acc:Q6XPR3]	1		

ENSG00000198242	60S ribosomal protein L23a. [Source:Uniprot/SWISSPROT;Acc:P62750]	17		
ENSG00000198648	STE20/SPS1-related proline-alanine-rich protein kinase (EC 2.7.11.1) (Ste-20-related kinase) (Serine/threonine-protein kinase 39) (DCHT). [Source:Uniprot/SWISSPROT;Acc:Q9UEW8]	2		
ENSG00000198692	Eukaryotic translation initiation factor 1A, Y-chromosomal (eIF-1A Y isoform) (eIF-4C). [Source:Uniprot/SWISSPROT;Acc:O14602]	Y		
ENSG00000198727	Cytochrome b. [Source:Uniprot/SWISSPROT;Acc:P00156]	MT	1	
ENSG00000198786	NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3) (NADH dehydrogenase subunit 5). [Source:Uniprot/SWISSPROT;Acc:P03915]	MT	1	1
ENSG00000198804	Cytochrome c oxidase subunit 1 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I). [Source:Uniprot/SWISSPROT;Acc:P00395]	MT	1	1
ENSG00000198833	Ubiquitin-conjugating enzyme E2 J1 (EC 6.3.2.19) (Non-canonical ubiquitin-conjugating enzyme 1) (NCUBE1) (Yeast ubiquitin-conjugating enzyme UBC6 homolog E) (HSUBC6e). [Source:Uniprot/SWISSPROT;Acc:Q9Y385]	6	1	
ENSG00000198886	NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3) (NADH dehydrogenase subunit 4). [Source:Uniprot/SWISSPROT;Acc:P03905]	MT	1	1
ENSG00000198888	NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3) (NADH dehydrogenase subunit 1). [Source:Uniprot/SWISSPROT;Acc:P03886]	MT	1	1

ENSG00000198899	ATP synthase a chain (EC 3.6.3.14) (ATPase protein 6). [Source:Uniprot/SWISSPROT;Acc:P00846]			MT	1	1
These three are in Dec 2005 biomart version but not found in the june 2007 version						
Chromosome Name	Ensembl Gene ID	Ensembl Peptide ID	External Gene ID	Description		
1_NT_077967	ENSG00000188333	ENSP00000343742	NP_653296.1			
2_NT_079503	ENSG00000188762	ENSP00000340640	THAP4	THAP domain protein 4. [Source:Uniprot/SWISSPROT;Acc:Q8WY91]		
14	ENSG00000189170	ENSP00000340467				

References

1. Minot,G.R., Murphy,W.P. (1983) Landmark article (JAMA 1926). Treatment of pernicious anemia by a special diet. By George R. Minot and William P. Murphy. *JAMA*, **250**, 3328-3335.
2. RICKES,E.L., BRINK,N.G., KONIUSZY,F.R., WOOD,T.R., FOLKERS,K.A.R.L. (1948) Crystalline Vitamin B12. *Science*, **107**, 396-397.
3. Smith EL (1948) Purification of anti-pernicious anemia factors from liver. *Nature*, **161**, 9.
4. Hodgkin,D., Kamper,J., Mackay,M., Pickworth,J., Trueblood,K., White,J. (1956) Structure of Vitamin B-12. *Nature*, **173**, 64.
5. LENHERT,P.G., HODGKIN,D.C. (1961) Structure of the 5,6-dimethyl-benzimidazolylcobamide coenzyme. *Nature*, **192**, 937-938.
6. EDWARD L.RICKES, NORMAN G.BRINK, NORMAN G.BRINK, THOMAS R.WOOD, KARL FOLKERS (1948) Comparative data on vitamin B₁₂ from liver and from a new source *Streptomyces griseus*. *Science*, 634-635.
7. SHORB,M.S. (1948) Activity of Vitamin B12 for the Growth of *Lactobacillus lactis*. *Science*, **107**, 397-398.
8. Ribichich,K.F., Georg,R.C., Gomes,S.L. (2006) Comparative EST analysis provides insights into the basal aquatic fungus *Blastocladiella emersonii*. *BMC Genomics*, **7**, 177.
9. Raux,E., Schubert,H.L., Warren,M.J. (2000) Biosynthesis of cobalamin (vitamin B12): a bacterial conundrum. *Cell Mol Life Sci*, **57**, 1880-1893.
10. Rodionov,D.A., Vitreschak,A.G., Mironov,A.A., Gelfand,M.S. (2003) Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J. Biol. Chem.*, **278**, 41148-41159.
11. Buckel,W., Golding,B.T. (2006) Radical enzymes in anaerobes. *Annu. Rev. Microbiol.*, **60**, 27-49.
12. Drennan,C.L., Matthews,R.G., Ludwig,M.L. (1994) Cobalamin-dependent methionine synthase: the structure of a methylcobalamin-binding fragment and

- implications for other B12-dependent enzymes. *Curr. Opin. Struct. Biol.*, **4**, 919-929.
13. Roth,J.R., Lawrence,J.G., Bobik,T.A. (1996) Cobalamin (coenzyme B12): synthesis and biological significance. *Annu. Rev. Microbiol.*, **50**, 137-181.
 14. Baik,H.W., Russell,R.M. (1999) Vitamin B12 deficiency in the elderly. *Annu Rev Nutr*, **19**, 357-377.
 15. He,Q., Madsen,M., Kilkenney,A., Gregory,B., Christensen,E.I., Vorum,H., Hojrup,P., Schaffer,A.A., Kirkness,E.F., Tanner,S.M., *et al.* (2005) Amnionless function is required for cubilin brush-border expression and intrinsic factor-cobalamin (vitamin B12) absorption in vivo. *Blood*, **106**, 1447-1453.
 16. Seetharam,B. (1999) RECEPTOR-MEDIATED ENDOCYTOSIS OF COBALAMIN (VITAMIN B12). *Annual Review of Nutrition*, **19**, 173-195.
 17. Bose,S., Komorowski,R., Seetharam,S., Gilfix,B., Rosenblatt,D.S., Seetharam,B. (1996) In Vitro and in Vivo Inactivation of Transcobalamin II Receptor by Its Antiserum. *J. Biol. Chem.*, **271**, 4195-4200.
 18. Hvas,A.M., Nexø,E. (2006) Diagnosis and treatment of vitamin B12 deficiency--an update. *Haematologica*, **91**, 1506-1512.
 19. Banerjee,R.V., Johnston,N.L., Sobeski,J.K., Datta,P., Matthews,R.G. (1989) Cloning and sequence analysis of the Escherichia coli methH gene encoding cobalamin-dependent methionine synthase and isolation of a tryptic fragment containing the cobalamin-binding domain. *J. Biol. Chem.*, **264**, 13888-13895.
 20. Drennan,C.L., Huang,S., Drummond,J.T., Matthews,R.G., Lidwig,M.L. (1994) How a protein binds B12: A 3.0 Å X-ray structure of B12-binding domains of methionine synthase. *Science*, **266**, 1669-1674.
 21. Li,Y.N., Gulati,S., Baker,P.J., Brody,L.C., Banerjee,R., Kruger,W.D. (1996) Cloning, mapping and RNA analysis of the human methionine synthase gene. *Hum. Mol. Genet.*, **5**, 1851-1858.
 22. Leclerc,D., Wilson,A., Dumas,R., Gafuik,C., Song,D., Watkins,D., Heng,H.H., Rommens,J.M., Scherer,S.W., Rosenblatt,D.S., Gravel,R.A. (1998) Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 3059-3064.
 23. Chen,L.H., Liu,M.L., Hwang,H.Y., Chen,L.S., Korenberg,J., Shane,B. (1997) Human methionine synthase. cDNA cloning, gene localization, and expression. *J. Biol. Chem.*, **272**, 3628-3634.

24. Fujii,K., Huennekens,F.M. (1974) Activation of methionine synthetase by a reduced triphosphopyridine nucleotide-dependent flavoprotein system. *J. Biol. Chem.*, **249**, 6745-6753.
25. Fujii,K., Galivan,J.H., Huennekens,F.M. (1977) Activation of methionine synthase: further characterization of flavoprotein system. *Arch. Biochem. Biophys.*, **178**, 662-670.
26. Leclerc,D., Odievre,M., Wu,Q., Wilson,A., Huizenga,J.J., Rozen,R., Scherer,S.W., Gravel,R.A. (1999) Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene. *Gene*, **240**, 75-88.
27. Kolhouse,J.F., Utley,C., Allen,R.H. (1980) Isolation and characterization of methylmalonyl-CoA mutase from human placenta. *J. Biol. Chem.*, **255**, 2708-2712.
28. Fenton,W.A., Hack,A.M., Willard,H.F., Gertler,A., Rosenberg,L.E. (1982) Purification and properties of methylmalonyl coenzyme A mutase from human liver. *Arch. Biochem. Biophys.*, **214**, 815-823.
29. Ledley,F.D., Lumetta,M., Nguyen,P.N., Kolhouse,J.F., Allen,R.H. (1988) Molecular cloning of L-methylmalonyl-CoA mutase: gene transfer and analysis of mut cell lines. *Proc. Natl. Acad. Sci. U. S. A.*, **85**, 3518-3521.
30. Jansen,R., Kalousek,F., Fenton,W.A., Rosenberg,L.E., Ledley,F.D. (1989) Cloning of full-length methylmalonyl-CoA mutase from a cDNA library using the polymerase chain reaction. *Genomics*, **4**, 198-205.
31. Cavicchi,C., Donati,M.A., Pasquini,E., Poggi,G.M., onisi-Vici,C., Parini,R., Zammarchi,E., Morrone,A. (2005) Mutational spectrum in ten Italian patients affected by methylmalonyl-CoA mutase deficiency. *J. Inherit. Metab Dis.*, **28**, 1175-1178.
32. Acquaviva,C., Benoist,J.F., Pereira,S., Callebaut,I., Koskas,T., Porquet,D., Elion,J. (2005) Molecular basis of methylmalonyl-CoA mutase apoenzyme defect in 40 European patients affected by mut(o) and mut- forms of methylmalonic acidemia: identification of 29 novel mutations in the MUT gene. *Hum. Mutat.*, **25**, 167-176.
33. Mancía,F., Smith,G.A., Evans,P.R. (1999) Crystal Structure of Substrate Complexes of Methylmalonyl-CoA Mutase. *Biochemistry*, **38**, 7999-8005.
34. Mancía,F., Keep,N.H., Nakagawa,A., Leadlay,P.F., McSweeney,S., Rasmussen,B., Bosecke,P., Diat,O., Evans,P.R. (1996) How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution. *Structure.*, **4**, 339-350.

35. Gravel,R.A., Mahoney,M.J., Ruddle,F.H., Rosenberg,L.E. (1975) Genetic complementation in heterokaryons of human fibroblasts defective in cobalamin metabolism. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 3181-3185.
36. Mahoney,M.J., Hart,A.C., Steen,V.D., Rosenberg,L.E. (1975) Methylmalonicacidemia: biochemical heterogeneity in defects of 5'-deoxyadenosylcobalamin synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **72**, 2799-2803.
37. Leal,N.A., Park,S.D., Kima,P.E., Bobik,T.A. (2003) Identification of the human and bovine ATP:Cob(I)alamin adenosyltransferase cDNAs based on complementation of a bacterial mutant. *J. Biol. Chem.*, **278**, 9227-9234.
38. Dobson,C.M., Wai,T., Kadir,H., Narang,M., Lerner-Ellis,J., Hudson,T.J., Rosenblatt,D.S., Gravel,R.A. (2002) Identification of the gene responsible for the cblB complementation group of vitamin B12-responsive methylmalonic acidemia. *Hum. Mol. Genet.*, **In press**.
39. Dobson,C.M., Wai,T., Leclerc,D., Wilson,A., Wu,X., Dore,C., Hudson,T.J., Rosenblatt,D.S., Gravel,R.A. (2002) Identification of the gene responsible for the cblA complementation group of vitamin B12-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proc. Natl. Acad. Sci. U. S. A.*, **In press**.
40. Lerner-Ellis,J.P., Gradinger,A.B., Watkins,D., Tirone,J.C., Villeneuve,A., Dobson,C.M., Montpetit,A., Lepage,P., Gravel,R.A., Rosenblatt,D.S. (2006) Mutation and biochemical analysis of patients belonging to the cblB complementation class of vitamin B12-dependent methylmalonic aciduria. *Mol. Genet. Metab*, **87**, 219-225.
41. Yang,X., Sakamoto,O., Matsubara,Y., Kure,S., Suzuki,Y., Aoki,Y., Suzuki,Y., Sakura,N., Takayanagi,M., Inuma,K., Ohura,T. (2004) Mutation analysis of the MMAA and MMAB genes in Japanese patients with vitamin B(12)-responsive methylmalonic acidemia: identification of a prevalent MMAA mutation. *Mol. Genet. Metab*, **82**, 329-333.
42. Martinez,M.A., Rincon,A., Desviat,L.R., Merinero,B., Ugarte,M., Perez,B. (2005) Genetic analysis of three genes causing isolated methylmalonic acidemia: identification of 21 novel allelic variants. *Mol. Genet. Metab*, **84**, 317-325.
43. Saridakis,V., Yakunin,A., Xu,X., Anandakumar,P., Pennycooke,M., Gu,J., Cheung,F., Lew,J.M., Sanishvili,R., Joachimiak,A., *et al.* (2004) The structural basis for methylmalonic aciduria. The crystal structure of archaeal ATP:cobalamin adenosyltransferase. *J. Biol. Chem.*, **279**, 23646-23653.
44. Zhang,J., Dobson,C.M., Wu,X., Lerner-Ellis,J., Rosenblatt,D.S., Gravel,R.A. (2006) Impact of cblB mutations on the function of ATP:cob(I)alamin

- adenosyltransferase in disorders of vitamin B12 metabolism. *Mol. Genet. Metab.*, **87**, 315-322.
45. Schubert, H.L., Hill, C.P. (2006) Structure of ATP-Bound Human ATP:Cobalamin Adenosyltransferase. *Biochemistry*, **45**, 15188-15196.
 46. Lerner-Ellis, J.P., Dobson, C.M., Wai, T., Watkins, D., Tirone, J.C., Leclerc, D., Dore, C., Lepage, P., Gravel, R.A., Rosenblatt, D.S. (2004) Mutations in the MMAA gene in patients with the cblA disorder of vitamin B12 metabolism. *Hum. Mutat.*, **24**, 509-516.
 47. Korotkova, N., Lidstrom, M.E. (2004) MeaB is a component of the methylmalonyl-CoA mutase complex required for protection of the enzyme from inactivation. *J. Biol. Chem.*, **279**, 13652-13658.
 48. Padovani, D., Banerjee, R. (2006) Assembly and protection of the radical enzyme, methylmalonyl-CoA mutase, by its chaperone. *Biochemistry*, **45**, 9300-9306.
 49. Padovani, D., Labunska, T., Banerjee, R. (2006) Energetics of interaction between the G-protein chaperone, MeaB, and B12-dependent methylmalonyl-CoA mutase. *J. Biol. Chem.*, **281**, 17838-17844.
 50. Atkinson, J.L., Paterson, A., Renaud, D., Clarke, J.T.R., Wilken, B., Bulman, D., Rommens, J.M. (2002) 1652.
 51. Lerner-Ellis, J.P., Tirone, J.C., Pawelek, P.D., Dore, C., Atkinson, J.L., Watkins, D., Morel, C.F., Fujiwara, T.M., Moras, E., Hosack, A.R., *et al.* (2006) Identification of the gene responsible for methylmalonic aciduria and homocystinuria, cblC type. *Nat. Genet.*, **38**, 93-100.
 52. Rosenblatt, D.S., Hosack, A., Matiaszuk, N.V., Cooper, B.A., Laframboise, R. (1985) Defect in vitamin B12 release from lysosomes: newly described inborn error of vitamin B12 metabolism. *Science*, **228**, 1319-1321.
 53. Watkins, D., Rosenblatt, D.S. (1986) Failure of lysosomal release of vitamin B12: a new complementation group causing methylmalonic aciduria (cblF). *Am. J. Hum. Genet.*, **39**, 404-408.
 54. Vassiliadis, A., Rosenblatt, D.S., Cooper, B.A., Bergeron, J.J. (1991) Lysosomal cobalamin accumulation in fibroblasts from a patient with an inborn error of cobalamin metabolism (cblF complementation group): visualization by electron microscope radioautography. *Exp. Cell Res.*, **195**, 295-302.
 55. Waggoner, D.J., Ueda, K., Mantia, C., Dowton, S.B. (1998) Methylmalonic aciduria (cblF): case report and response to therapy. *Am. J. Med. Genet.*, **79**, 373-375.

56. Goodman,S.I., Moe,P.G., Hammond,K.B., Mudd,S.H., Uhlendorf,B.W. (1970) Homocystinuria with methylmalonic aciduria: two cases in a sibship. *Biochem. Med.*, **4**, 500-515.
57. Willard,H.F., Mellman,I.S., Rosenberg,L.E. (1978) Genetic complementation among inherited deficiencies of methylmalonyl-CoA mutase activity: evidence for a new class of human cobalamin mutant. *Am. J. Hum. Genet.*, **30**, 1-13.
58. Suormala,T., Baumgartner,M.R., Coelho,D., Zavadakova,P., Kozich,V., Koch,H.G., Berghauser,M., Wraith,J.E., Burlina,A., Sewell,A., *et al.* (2004) The cblD defect causes either isolated or combined deficiency of methylcobalamin and adenosylcobalamin synthesis. *J. Biol. Chem.*, **279**, 42742-42749.
59. Cooper,B.A., Rosenblatt,D.S., Watkins,D. (1990) Methylmalonic aciduria due to a new defect in adenosylcobalamin accumulation by cells. *Am. J. Hematol.*, **34**, 115-120.
60. Watkins,D., Matiaszuk,N., Rosenblatt,D.S. (2000) Complementation studies in the cblA class of inborn error of cobalamin metabolism: evidence for interallelic complementation and for a new complementation class (cblH). *J. Med. Genet.*, **37**, 510-513.
61. Hill,H.Z., Goodman,S.I. (1974) Detection of inborn errors of metabolism. II. Defects in propionic acid metabolism. *Clin. Genet.*, **6**, 73-78.
62. Loyer,M., Leclerc,D., Gravel,R.A. (1995) Interallelic complementation of beta-subunit defects in fibroblasts of patients with propionyl-CoA carboxylase deficiency microinjected with mutant cDNA constructs. *Hum. Mol. Genet.*, **4**, 1035-1039.
63. Lubke,T., Marquardt,T., Etzioni,A., Hartmann,E., von,F.K., Korner,C. (2001) Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. *Nat. Genet.*, **28**, 73-76.
64. Stacey,D.W., Allfrey,V.G. (1976) Microinjection studies of duck globin messenger RNA translation in human and avian cells. *Cell*, **9**, 725-732.
65. McCormick,M. (1987) Sib selection. *Methods Enzymol.*, **151**, 445-449.
66. Lin,P.F., Brown,D.B., Murphy,P., Yamaizumi,M., Ruddle,F.H. (1987) Detection and characterization of specific mRNA by microinjection and complementation of mutant cells. *Methods Enzymol.*, **151**, 371-382.
67. Halarnkar,P.P., Blomquist,G.J. (1989) Comparative aspects of propionate metabolism. *Comp Biochem. Physiol B*, **92**, 227-231.

68. Chandler,R.J., Aswani,V., Tsai,M.S., Falk,M., Wehrli,N., Stabler,S., Allen,R., Sedensky,M., Kazazian,H.H., Venditti,C.P. (2006) Propionyl-CoA and adenosylcobalamin metabolism in *Caenorhabditis elegans*: evidence for a role of methylmalonyl-CoA epimerase in intermediary metabolism. *Mol. Genet. Metab*, **89**, 64-73.
69. Chandler,R.J., Venditti,C.P. (2005) Genetic and genomic systems to study methylmalonic acidemia. *Mol. Genet. Metab*, **86**, 34-43.
70. Wheeler,D.L., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Chetvernin,V., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S., *et al.* (2007) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **35**, D5-12.
71. Remm,M., Storm,C.E., Sonnhammer,E.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.*, **314**, 1041-1052.
72. O'Brien,K.P., Remm,M., Sonnhammer,E.L. (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.*, **33**, D476-D480.
73. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W., Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389-3402.
74. Watanabe,F., Nakano,Y., Tamura,Y., Yamanaka,H. (1991) Vitamin B12 metabolism in a photosynthesizing green alga, *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta*, **1075**, 36-41.
75. Miyamoto,E., Watanabe,F., Ebara,S., Takenaka,S., Takenaka,H., Yamaguchi,Y., Tanaka,N., Inui,H., Nakano,Y. (2001) Characterization of a vitamin B12 compound from unicellular coccolithophorid alga (*Pleurochrysis carterae*). *J. Agric. Food Chem.*, **49**, 3486-3489.
76. Croft,M.T., Lawrence,A.D., Raux-Deery,E., Warren,M.J., Smith,A.G. (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature*, **438**, 90-93.
77. Ribichich,K.F., Georg,R.C., Gomes,S.L. (2006) Comparative EST analysis provides insights into the basal aquatic fungus *Blastocladiella emersonii*. *BMC. Genomics*, **7**, 177.
78. PICKEN,J.C., Jr., BAURIEDEL,W.R. (1950) Comparative bioautography of vitamin B12 and related growth factors using *Euglena gracilis* and *Lactobacillus leichmannii*. *Proc. Soc. Exp. Biol. Med.*, **75**, 511-515.

79. ROSS,G.I. (1952) Vitamin B12 assay in body fluids using *Euglena gracilis*. *J. Clin. Pathol.*, **5**, 250-256.
80. Watanabe,F., Tamura,Y., Stupperich,E., Nakano,Y. (1993) Uptake of cobalamin by *Euglena* mitochondria. *J. Biochem. (Tokyo)*, **114**, 793-799.
81. Torrents,E., Trevisiol,C., Rotte,C., Hellman,U., Martin,W., Reichard,P. (2006) *Euglena gracilis* ribonucleotide reductase: the eukaryote class II enzyme and the possible antiquity of eukaryote B12 dependence. *J. Biol. Chem.*, **281**, 5604-5611.
82. Li,N., Seetharam,S., Seetharam,B. (1995) Genomic Structure of Human Transcobalamin II: Comparison to Human Intrinsic Factor and Transcobalamin I. *Biochemical and Biophysical research communications*, **208**, 756-764.
83. Roth,J.R., Lawrence,J.G., Bobik,T.A. (1996) Cobalamin (coenzyme B12): synthesis and biological significance. *Annu. Rev. Microbiol.*, **50**, 137-181.
84. Hu,Z.Z., Valencia,J.C., Huang,H., Chi,A., Shabanowitz,J., Hearing,V.J., Appella,E., Wu,C. (2007) Comparative Bioinformatics Analyses and Profiling of Lysosome-Related Organelle Proteomes. *Int. J. Mass Spectrom.*, **259**, 147-160.