# SUMMARY

Congenital hyperinsulinism of infancy is a pathology characterized by a profound hypoglicaemia related to a disregulated insulin secretion.; the frequency of the disease is 1:30000-50000 live birth world –wide, but in same isolated populations with high consanguinity like the Arabic Peninsula, frequency of 1:2500 is reported. CHI has two different histological presentations, which are characterized either by a diffuse insulin secretion or by an adenomatous iperplastic focal form of the pancreatic β- cells, underling different molecular mechanism at the basis of the disease..

At present it is well known that disfunctions of pancreatic β- cells K-ATP channel are a common cause of CHI. The two subunits of the channel are coded by two genes, KCNJ11 and ABCC8. Mutations in ABCC8 gene are responsible of the 50-60% of focal and diffuse CHI patients, otherwise mutations in KCNJ11 are responsible for 10-15% of cases..Other less frequent molecular mechanism responsible for CHI are glutamate dehydrogenase (GDH), glucokinase and short- chain 3 hydroxyacyl Coenzyme A dehydrogenase (SCHAD) deficiency. CHI derives from a heterogeneous genetic background and many other genes involved in the pancreatic β- cells metabolism are good candidate genes for this disease.

Current genome-wide high density mapping technologies have improved the number of feasible applications of SNPs genotyping in the identification of new disease associated loci using genotype calls not only for association studies involving single polymorphisms but, thanks to the contemporary analysis of copy number, also revealing the presence of long stretches of homozygosity and their classification as structural variations. These studies underline how this phenomenon is not normally due to uniparental disomy, that is a rare event associated to particular syndromes in which region of imprinted region or cancer predisposing genes are involved. This homozygosity stretches are generally located in regions of broad linkage disequilibrium and low frequency of recombination, and are considered autozygosity trait . One possible mechanism at the basis of autozygosity is the presence of a common ancestor which decreases the probability of recombination or the nature of particular chromosome trait allowing them to remain unchanged .

This methodological approach to the analysis of SNPs genotyping data can be called “homozygosity mapping”, which can be useful in the identification of disease associated locus/genes. Homozygosity mapping focuses on individuation of regions of autozygosity which are in common between affected individuals of the same family or between affected unrelated probands with the same disease which have high probability to harbor new mutations , and can underline regions that can lead to a new gene (or to a number of candidate loci) as responsible of the genetic basis of the disease.

**Summary**

In order to identify new disease susceptibility loci in a cohort of consanguineous and not consanguineous 34 CHI Italian families , we performed genome-wide homozygosity mapping using 250K NspI Gene Chip Affymetrix SNP microarrays and to reveal the number and the nature of homozygosity trait in common between probands to underline new regions containing candidate genes involved in the ethiopathogenesis of CHI. Further screening for mutations in candidate genes was performed and biochemical pathways involved in the disease has been analyzed.

We used CNAG (v3.0) software to describe homozygosity traits present in 34 families with CHI affected probands. This analysis evidenced that homozygosity traits of more than 1 Mb in length are present in 16 chromosome , they span between 3.4 and 79 Mb and they are not preferentially distributed .Analyzing homozygosity events inside every single family, we have evidenced their presence in more than one member , even if all families include only affected probands and parents and sibs are healthy subjects. Considering the number of homozygosity events for every family, it was reveled that consanguineous families bear a number of homozygous stretches much higher that not consanguineous ones. Following investigation of the parental origin of homozygous stretches has confirmed that they were not inherited and the contemporary analysis of the copy number with CNAG (v3.0) software did not reveled any copy number loss, concluding that they were not deletions. Mendelian errors check (all below 0.1%, threshold due to genotyping errors and not transmission error) and parental haplotypes analysis has excluded uniparental disomy (UPD) and they must be considered as autozygosity traits.

To further characterize homozygosity traits in CHI probands, we used dCHIP software to underline homozygosity regions in common between at least two patients. On Chromosome 11 we found five patients sharing a common homozygous trait on cytoband 11p15, where the two major CHI candidate genes are present. In two of these probands, where mutational screening of ABCC8 and KCNJ11 genes was completed , the causative mutation was revealed.

Other relevant homozigosity region in common between 4 patients are present on chromosome 6 (p21.33-p21.31) and on chromosome 10 (q26.13- q26.2), while 18 common traits between 3 patients were revealed One of the numerous traits in common between two patient on chromosome 1 showed a homozygous region in 1p31, one patien bearing a long contiguous trait of 25 Mb. This common region contains medium-chain acyl-CoA dehydrogenase (MCAD) gene, involved in organic acid metabolism deficiency, and in that proband an homozygous synonymous mutation P195P in exon 8 of MCAD gene was identified (587 T >G); her parents and brother evidenced the same etherozygous nucleotidic variation.

On chromosome 4 one probands showed the longest homozygosity stretch (from 4q13.2 to 4q31.21), which ranged 79,01 Mb and included 312 known genes in the region. This stretch contains the causative short- chain 3 hydroxyacyl Coenzyme A dehydrogenase gene (HADH), which is genotyped

on the microarray by 7 SNPs covering 45 kb. Direct sequencing of the coding region of HADH revealed a homozygous C to T transition in exon 6 causing a premature stop of the synthetized protein at codon 236 (706C>T), resulting in the nonsense mutation R236X .Mutation analysis of exon 6 of the HADH gene in the parents of the proband showed they were heterozygous carriers of the mutation.

We further investigated if SCHAD mRNA expression in various tissue and in patient's blood lymphocytes and we found that RNA bearing the mutation is not suppressed by RNA nonsense-mediated decay. Homology modeling of the protein bearing the R236X mutation revealed that it leads to a truncated protein lacking the most part of the C- terminal domain, particularly relevant to the closing of the active site cleft. This is the fourth described mutation in HADHSC gene in a CHI patient, indicating the importance of β-oxidation disregulation in this disease.

A preliminary bioinformatic pathways analysis in CHI probands underlined that more then one pathway is involved in the disease, in particular type I diabetes mellitus, glycolysis and gluconeogenesis, fatty acid metabolism and amino acids metabolism. Gene informations in these pathways clearly evidenced medium-chain acyl-CoA dehydrogenase (MCAD) gene has a major candidate, together with three other important genes of the fatty acid metabolism as alcohol dehydrogenase, long chain acil-CoA synthetase and cytocromeP450.

Homozygosity mapping in 34 CHI families has confirmed a relevant presence of autozygosity , due to presence of high inbreeding in at least 5 analysed families and to the restricted geographical provenience. Autozygosity excludes that deletions or UPD are at the molecular basis of the disease.

Even if each patient has an homozygosity profile strictly peculiar, there are common region shared between at least two probands. These regions are good candidates for the discovery of new disease associated genes/loci. The discovery that CHI probands showing homozygosity traits on chromosome 11, in particular on 11p15.1, bear a mutation in known causatives genes and the recent finding of a not jet described stop mutation in HADH gene on chromosome 4 confirm the power of SNPs genome-wide genotyping technologies to underline loci with high mutational probability and to elucidate the molecular basis of autosomic recessive diseases.

To our knowledge this is the first study in which molecular mechanism of CHI disease was investigated with genome-wide homozygosity mapping method in consanguineous and not consanguineous CHI families. This methodology was highly successful in revealing the recessive disease causing mutation in four of the 35 patients; two of these mutations were in ABCC8 and KCNJ11 known causative genes , one was a novel mutation in HADH gene which could describe the fourth case of SCHAD-HI, and the last one was a putative splicing variant of MCAD gene that could indicate the first described case of MCAD-HI.

Furthermore, bioinformatics pathway analysis underlined how a huge number of genes in autozygosity traits shared by our 35 CHI probands covered the two major metabolic pathways of

glycolisis/gluconeogenesis and fatty acid metabolism, indicating how it would be important for a correct diagnosis to consider CHI has as metabolic disease and to describe these patients with a metabolic profile which includes at least organic acids investigations and acilcarnitines characterizations, as suggested by Hussain K. Thanks to the increasing number of described SCHAD-HI, our study also indicated that mutations of HADH gene should be sought in hyperinsulinemic patients in whom diffuse form of CHI and autosomal recessive inheritance can be presumed, when KCNJ11 and ABCC8 genes mutational screening is negative.

Genome-wide homozygosity mapping in the present study has clearly evidenced 1p31.2-31.1, 4q27-28.1, 6p22.3-22.31, 11p15.4-15.2, and 18q12.2 chromosomal loci to be the most related to CHI disease and to better redefine that regions and to single out potential associated genes , we are planning to use SNPs genotypes calls generated from these 34 CHI family microarrays data to perform family based transmission disequilibrium analysis . On the other hand , we will increase the number of CHI families enrolled in the study in order to evidence by homozygosity mapping ROHs which are strictly peculiar of the disease and not common in the normal population.

# INTRODUCTION

# 1.Congenital Hyperinsulinism of Infancy (CHI)

## 1.1 History and terminology

Hyperinsulinism (HI) was first described in 1954 by Mac Quarri as "idiopathic hypoglycemia of infancy".

HI has subsequently been referred to by many names, including leucine-sensitive hypoglycemia, islet disregulation syndrome, persistent hyperinsulinemic hypoglycemia of infancy, and nesidioblastosi. Although the term nesidioblastosis continues to be present in literature, it has been recognized that nesidioblastosis is normal feature of the pancreas in early infancy , describing neo differentiation of pancreatic Langerhans Islets from duct pancreatic epithelium (Rahier, Guiot et al. 2000) and should not be used to refer to lesion associated with HI.We prefer to denote the condition as hyperinsulinism for simplicity and to avoid confusion, and to refer to the group of permanent inborn condition as Congenital Hyperinsulinism (CHI).

## 1.2 Pathology

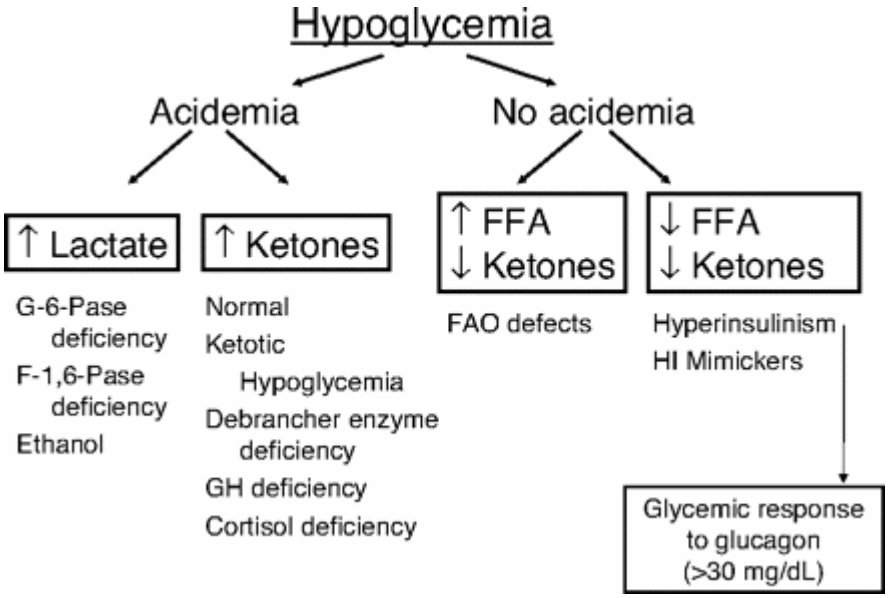
Congenital Hyperinsulinism (CHI) represent a group of clinically, genetically and morphologically heterogeneous disorders characterized by dysregulated insulin secretion and resulting in severe and persistent hypoglycemia. During the last 10 years, advances in molecular genetics and in the understanding of β-cell biochemical pathways have resulted in understanding of the pathophysiology of most form of CHI. Despite these advances, treatment of the infant with CHI continues to be one of the major challenges in pediatric endocrinology.

Worldwide, CHI occurs at a frequency of 1 in 30,000 to up to 50,000 live birth, but in some isolated population rates of 1 in 3220 have been reported. Higher rates of 1 in 2550 live birth have been reported in areas of high consanguinity such as the Arabian Peninsula (Otonkoski, Ammala et al. 1999).

## 1.3 Diagnosis

Infants with CHI present with severe and persistent hypoglycemia manifested by letargy, seizures, apnea and increased glucose requirements ( up to 20-30 mg/kg/min). Plasma insulin levels are inappropriately increased in the setting hypoglycemia ; however , clearly increased insulin levels are often not present at the time of hypoglycemia.

These might be due to the periodic release of insulin, to rapid hepatic clearance or to the activity of insulin degrading enzymes that are present in hemolysed samples. Therefore, the diagnosis of CHI must be frequently be based on evidence of excessive insulin action, such as suppression of plasma β-hydroxybutirate and free fatty acid levels.An inappropriate glycemic response to glucagon (>30 mg/dl) at the time of hypoglycemia is consistent with excess insulin action and is useful for confirming the diagnosis.Additional laboratory test for specific forms of CHI include plasma ammonia levels and a plasma acyl- carnitine profile (butyryl carnitine) and urine organic acid (3-hydroxyglutarate).(Palladino, Bennett et al. 2008)

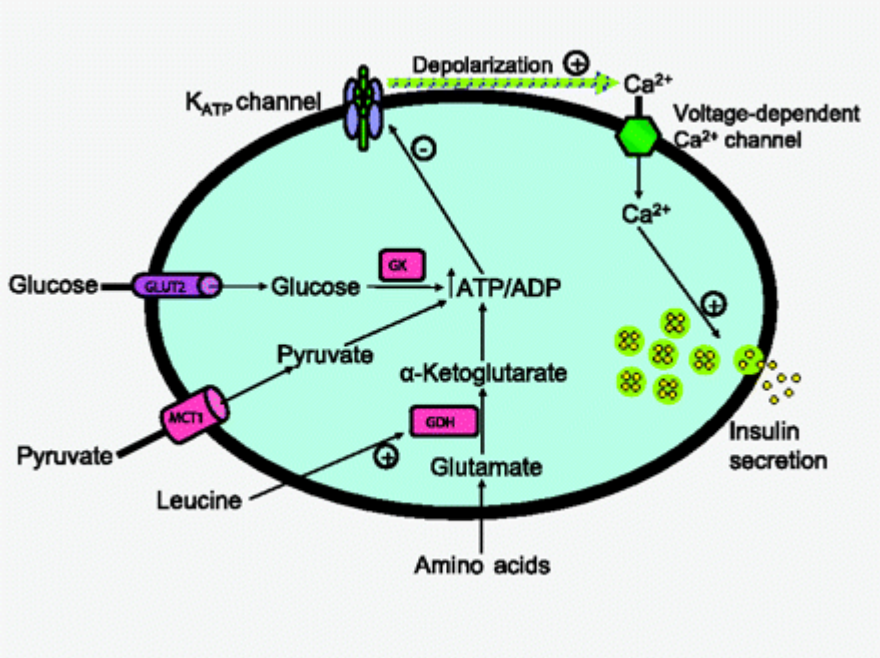



Results from the critical blood sample obtained at the time of hypoglycemia serve as the basis for distinguishing 4 categories of disease: impairment of gluconeogenesis, normal and abnormal forms of ketotic hypoglycemia, defects in fatty acid oxidation and ketogenesis, and impairment of lipolysis and ketogenesis.G-6-Pase, glucose-6-phophatase; F-1,6-Pase,fructose-1,6-biphosphatase; GH, growth hormone; FFA, free fatty acids; FAO, fatty acid oxidation.

| Genetic form | Gene | Chromosome | Inheritance | Clinical features | Treatment |
|---|---|---|---|---|---|
| $K_{ATP}$-HI | *ABCC8* *KCNJ11* | 11p15 | Diffuse: AR Focal: loss of heterozygosity with paternal mutation | Severe hypoglycemia; unresponsive to medical therapy | Pancreatectomy or "conservative" therapy with octreotide and continuous feedings |
| Dominant $K_{ATP}$-HI | *ABCC8* *KCNJ11* | 11p15 | AD | Milder hypoglycemia; responsive to diazoxide | Diazoxide |
| GDH-HI (HI/HA) | *GLUD-1* | 10q | AD | Fasting and postprandial hypoglycemia; less severe than $K_{ATP}$-HI; protein sensitivity; asymptomatic hyperammonemia | Diazoxide |
| GK-HI | *GCK* | 7p | AD | Variable phenotype: can range from easy to manage with medical therapy to very difficult to control | Diazoxide; pancreatectomy |
| SCHAD-HI | *HADH* | 4q | AR | Mild to severe hypoglycemia; abnormal acylcarnitine profile | Diazoxide |
| MCT1 (EIHI) | *SLC16A1* | 1p | AD | Exercise-induced hypoglycemia, especially anaerobic | Carbohydrate intake during exercise; limit exercise |

AR, autosomal recessive; AD, autosomal dominant.

Classification of genetic forms of congenital hyperinsulinism.




Mechanisms of insulin secretion in pancreatic β-cells.

Increase in ATP:ADP ratio inhibits the $K_{ATP}$ channel, resulting in closure of the channel, depolarization of the membrane, influx of calcium, and release of insulin. Insulin secretion is stimulated by glucose oxidation via GK and by leucine stimulation of glutamate oxidation via GDH. Abnormally increased pyruvate levels in the β-cell will stimulate insulin secretion. GLUT2, glucose transporter 2.

## 1.4 Clinical presentation

In CHI a blood sample taken at the time of hypoglycaemia will show an inappropriately raised serum insulin level with low serum fatty acid levels and ketone bodies .The low serum fatty acid levels and ketone bodies reflect the metabolic 'footprint' of insulin action. Unregulated insulin secretion increases the glucose consumption by insulin-sensitive tissues, such as muscle, adipose tissue, and liver, while simultaneously suppressing hepatic glucose production (both glycogenolysis and gluconeogenesis), lipolysis, and ketogenesis. Due to this metabolic 'footprint' of insulin action, the intravenous glucose infusion rate required to maintain normoglycaemia is increased (>8 mg/kg/min, normal 4-6 mg/kg/min). There is no correlation between the serum insulin level and the severity of hypoglycaemia. A 'normal' insulin level for normoglycaemia is usually inappropriate in the presence of hypoglycaemia, especially taken in the context of a high glucose requirement to maintain normoglycaemia .The serum lactate level may be elevated in some forms of CHI and the serum ammonia concentration must be measured in all patients presenting with CHI because of the association with hyperinsulinism/hyperammoniaemia (HI/HA) syndrome.Urinary organic acid and acylcarnitine analysis should also be performed, since short-chain *L*-3-hydroxyacyl-CoA dehydrogenase (SCHAD) deficiency can cause CHI . The serum cortisol and glucagon counterregulatory hormonal responses may be blunted in CHI, but replacement therapy with glucocorticoids does not seem to affect the severity of the disease (Hussain 2008) .

## 1.5 Prognosis and outcome

Children with HI are at risk for neurodevelopmental disabilities and must be screened.
In a series of 90 patients with HI, severe mental retardation was found in 8 %, with less sever disability in 18%. Some infants with CHI are macrosomic which may reflect their exposure to perinatal hyperinsulinaemia, but the absence of macrosomia does not exclude CHI, as not all infants with CHI are macrosomic. Some patients with CHI may have mild facial dysmorphism such as a high forehead, a small nasal tip, and short columella with a square face, although the reason for this is unclear (de Lonlay, Cormier-Daire et al. 2002).
Psychomotor retardation was found to be more common in patients with neonatal hypoglycemia than in those with onset of hypoglycemia during infancy.

# 2.CHI Syndromic and Transient Forms

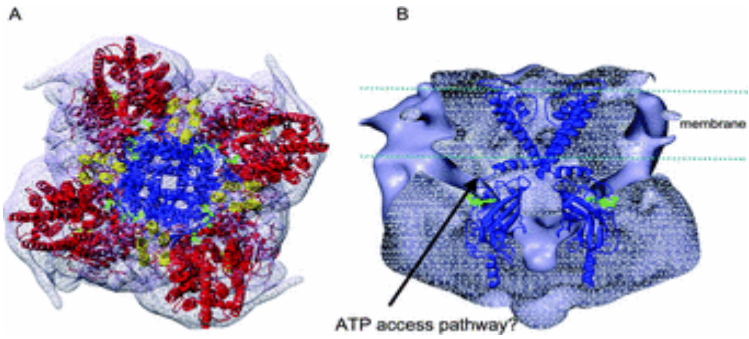
Beckwith-Wiedemann syndrome is the most common syndrome associated with CHI . Beckwith-Wiedemann syndrome is characterized by prenatal and/or postnatal overgrowth, macroglossia, anterior abdominal wall defects, organomegaly, hemihypertrophy, ear lobe creases, helical pits, and renal tract abnormalities. The incidence of HH in children with Beckwith-Wiedemann syndrome is about 50%.(Hussain, Cosgrove et al. 2005) This hypoglycemia can be transient, which, in the majority of infants, will be asymptomatic and resolve within the first few days of life. In about 5% of these children, the HH can be persistent and extend beyond the neonatal period, requiring either continuous feeding, medical therapy, or, in rare cases, partial pancreatectomy. Aetiology and the mechanisms responsible for transient forms of HH are unclear. Although the aetiology of these transient forms of HH is not thought to be genetic, recently mutations in the gene encoding the hepatic nuclear transcription factor-4α (*HNF4A* gene) have been reported. In these patients the birth weight of the heterozygote *HNF4A* mutation carriers was dramatically increased. However, it is not known how mutations in *HNF4A* gene cause transient HH.(Valayannopoulos, Romano et al. 2008).

# 3.CHI due to Channelopathies

CHI can be classified into 'channelopathies', where defects in the pancreatic β-cell $K_{ATP}$ channels lead to unregulated insulin secretion, or 'metabolopathies', with increased β-cell ATP formation or accumulation of intermediary metabolites, triggering insulin secretion (Ashcroft 2006).

The commonest genetic causes of CHI are autosomal recessive mutations in the genes *ABCC8* and *KCNJ11* (encoding the two subunits SUR1 and KIR6.2, respectively) of the pancreatic $K_{ATP}$ channels.Autosomal dominant mutations have also been described. $K_{ATP}$ channels play a pivotal role in transducing metabolic signals to electrical changes in membrane potential (De Leon and Stanley 2007). These mutations result in differing abnormalities of recombinant $K_{ATP}$ channels, including protein folding, protein synthesis defects, assembly and trafficking defects, and alterations in both nucleotide regulation and open-state frequency. Mutations that affect the regulation of the $K_{ATP}$ channels by altering their sensitivity to changes in ADP/ATP will also lead to an unregulated insulin secretion.

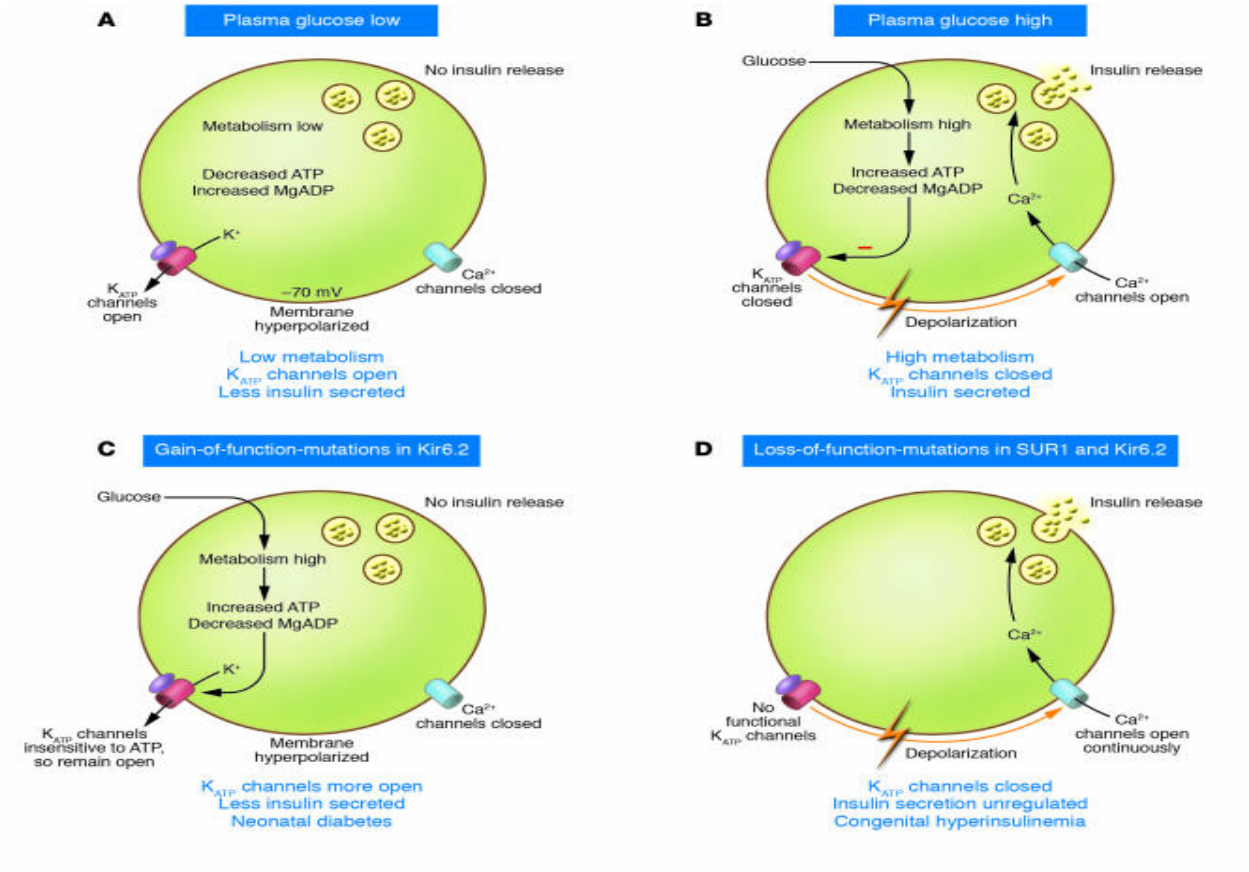
Several mutations have now been described that result in the loss of ADP-dependent gating properties of the channel. Loss of ADP-dependent gating results in the constitutive inhibition of $K_{ATP}$ channels by ATP. In more than 50% of the patients, screening has failed to define the genetic basis of CHI. In some populations, mutations in the *ABCC8* gene account for only about 20% of the cases of CHI suggesting that other genes may be involved (Giurgea, Bellanne-Chantelot et al. 2006).




Octameric structure of the $K_{ATP}$ channel

**(A)** Top view of the EM density map of the purified $K_{ATP}$ channel complex, with molecular models of Kir6.2 in blue, SUR1 minus the first five transmembrane domains (TMD0) in red, and TMD0 in yellow; inserted. ATP molecules are shown in green. Reproduced from [7], with permission. © 2005 European Molecular Biology Organization.

**(B)** Transverse sectional view of the EM density with a model of Kir6.2 (blue) inserted, showing clefts (arrowed) through which ATP (green) could access its binding site on Kir6.2. Reproduced from [7], with permission. © 2005 European Molecular Biology Organization.




The $K_{ATP}$ channel couples glucose metabolism to insulin secretion. Glucose enters the cell via the GLUT2 transporter, and via glycolytic and mitochondrial metabolism leads to an increase in $_{ATP}$ and a fall in MgADP in the immediate vicinity of the $K_{ATP}$ channel. This results in $K_{ATP}$ channel closure, membrane depolarization, opening of voltage-gated $Ca^{2+}$ channels, $Ca^{2+}$ influx, and exocytosis of insulin granules.

## *3.1 K-channel disease*

Historically, familial studies allowed location of a CHI candidate gene area on chromosome 11, and it happened that, in the interval 11p15.1, defined by conventional genetics, the two components of the β-cell K channel are present: *ABCC8* (coding for SUR1) and *KCNJ11* (coding for Kir6.2). In consequence, they were good candidate genes, and molecular biology studies showed that they are the

cause of the disease when mutated (Otonkoski, Ammala et al. 1999). Mutations of SUR1 are the most frequent. Mutations found in Israel and Finland are limited to 2 loci on the gene, whereas in France, there are more than 40 known mutations involving the same gene (Saudubray, de Lonlay et al. 2000).There are 2 hot spots for mutations on the 2 nucleotide-binding domain first (NBF) loops (de Lonlay, Fournet et al. 2002; de Lonlay, Touati et al. 2002; Saudubray, Nassogne et al. 2002) The first hypothesis was that the K channel plays a major role in cell polarization, being localized at the cell membrane, and $\beta$-cell electrophysiologic studies were carried out in vitro. Today, it is known that K channels are present mainly on the granule membrane inside the cell and are involved in granule docking and secretion. It has been demonstrated that SUR1 and Kir6.2 are assembled together on the granule membrane, allowing secretion, and are recycled at the cell membrane level. The same combination of SUR and Kir6.2, as is found in a K channel, is also present in some neurons of the myenteric plexus and the central nervous system. (Hussain and Aynsley-Green 2003)

Diffuse forms of hyperinsulinism are now, most of the time, linked to the presence of 2 mutations of the K-channel genes, one inherited from the father and the other from the mother, with a recessive inheritance. More rarely, the diseases are attributable to the presence of a dominant mutation (de Lonlay, Touati et al. 2002) (Saudubray, Nassogne et al. 2002).

Mutations of SUR1 are more common than those of Kir6.2. Focal forms of hyperinsulinism have a somatic K-channel mutation, generally inherited from the father, and in the focal area, LOH of the corresponding maternal chromosome 11 region is detectable by microsatellite marker analysis of DNA extracted from fresh tissue scraped from frozen sections. The defective chromosome area also contains the tumor-suppressor gene *H19*, which regulates parental imprinting, and the growth factor *IGF2* gene, which is subject to parental imprinting. The *KIP2* gene (*CDKN1C*), coding for P57 protein, is present in the same area, and P57 expression is lost in the focal lesion, whereas it is normal outside the islets of Langerhans. However, in a family with 2 parents heterozygous for the SUR-1 mutation, it is possible to have either a focal or a diffuse form of the disease, depending on the type of chromosome 11 inherited by the child. (Fournet, Mayaud et al. 2000; Fournet, Mayaud et al. 2001).

ATP-sensitive potassium (KATP) channels regulate the flux of K ions across cell membranes and thereby link cell metabolism to electrical activity. Metabolic inhibition opens KATP channels, and the resulting efflux of K1 ions causes membrane hyperpolarization and the suppression of electrical activity. Conversely, increased metabolism closes the KATP channels, resulting in membrane depolarization that stimulates electrical activity. This electrical activity can then trigger muscle contraction and the release of hormones or neurotransmitters(Hussain and Cosgrove 2005). The importance of KATP channels in insulin secretion was established over 20 years ago. These channels

couple glucose metabolism to membrane electrical activity and insulin release in pancreatic betacells.(Hussain and Aynsley-Green 2003)

When blood glucose levels rise, the resulting increase in glucose metabolism results in a change in the ratio of cytosolic nucleotides (ADP/ATP), which causes closure of the KATP channel and leads to membrane depolarization. This subsequently activates voltage-dependent calcium channels, leading to an influx of calcium. This increase in intracellular calcium is then the trigger for insulin granule exocytosis .The beta-cell KATP channel can be pharmacologically regulated by sulfonylurea drugs, which work by binding to and closing the KATP channel, and are widely used to treat type 2 diabetes (T2DM). Conversely, a second class of drugs, such as diazoxide, work by opening the channel (K channel openers).(de Lonlay, Giurgea et al. 2004; de Lonlay, Giurgea et al. 2004).

The beta-cell KATP channel consists of two essential subunits: 1) Kir6.2, which is the pore-forming unit and belongs to the inwardly rectifying potassium channel family, and 2) sulfonylurea receptor 1 (SUR1), which belongs to the ATP-binding cassette (ABC) transporter family. The channel is a octameric complex of 4 Kir6.2 and 4 SUR1 subunits.The binding of ATP to Kir6.2 results in KATP channel closure, while sulfonylurea drugs, potassium channel openers, and magnesium nucleotides bind to SUR1 via its two cytosolic nucleotide binding domains. Given the central role of the KATP channel in insulin secretion, it is not surprising that mutations in the genes that encode the subunits of this channel can result in both hypo- and hyperglycemia (Gloyn, Pearson et al. 2004).

The human SUR1 cDNA contains a single open reading frame that encodes for 1,581 amino acids with a molecular weight of 177 KDa (GenBank NM_000352.2). The gene (ABCC8; MIM 600509) consists of 39 exons and spans more than 100 kb of genomic DNA. ABCC8 has an alternatively spliced exon 17, which incorporates an additional amino acid (GenBank L78208, L78224). Homology analysis has shown that SUR1 is a member of the ABC superfamily, which includes cystic fibrosis transconductance regulators (CFTRs) and multidrug-resistance proteins (MDPs). Membrane topology has shown that the SUR1 consists of two nucleotide binding domains (NBDs) that control channel activity through their interaction with cytocyolic nucleotides Both NBDs contain two amino acid sequence motifs (termed ‘‘Walker A’’ and ‘‘Walker B’’), which are phosphate binding loops that form intimate contact with the phosphates of the nucleotide triphosphates that bind to the NBDs. NBDs are also found in other members of the ABC superfamily. Kir6.2 was cloned the same year as SUR1, and the gene (KCNJ11; MIM] 600937) consists of a single exon encoding a 390-amino acid protein (GenBank NM_000525.2). Interestingly, it is only 4.5 Kb from the ABCC8 gene on chromosome 11p15.1. Kir6.2 consists of two transmembrane regions and a pore-forming subunit. (Ashcroft 2005)

# 4. CHI due to Metabolopathies

*4.1 Glutamate Dehydrogenase (GDH).*

Activating mutations in GDH are the second commonest cause of CHI. Activating mutations in GDH underlie the molecular basis of the HI/HA syndrome and may explain the 'leucine-sensitive' hypoglycaemia described in previous years. The HI/HA syndrome is caused by missense mutations of GDH that reduce the sensitivity of the enzyme to allosteric inhibition by the high-energy phosphates GTP and ATP. GDH is allosterically activated by leucine and inhibited by GTP. Mutations which cause loss of inhibition by GTP cause leucine to increase the oxidation of glutamate, thereby raising the ratio of ATP/ADP in the pancreatic 3-cell. The increased ratio of ATP/ADP then triggers closure of the $K_{ATP}$ channel, opening the voltage-gated calcium channel, raising cytosolic calcium, and triggering the release of insulin. Patients with the HI/HA syndrome can present with hypoglycaemia either in the neonatal period or later on in childhood. These patients also have a mildly elevated plasma ammonia concentration which appears to be asymptomatic. Patients show no signs of lethargy or headaches, typical of other forms of hyperammonaemia. The mechanism of the hyperammonaemia is still unclear at present. These patients typically demonstrate protein-induced HH (leucine sensitivity), but also have fasting hypoglycaemia.Children with the HI/HA syndrome have an unusual frequency of absence-type seizures These children have an EEG pattern of generalized epilepsy that resembles the seizures associated with mutations of plasma membrane ion channels. It is unlikely that this seizure pattern is a manifestation of ammonia toxicity. (Bahi-Buisson, El Sabbagh et al. 2008)

*4.2 Glucokinase (GCK).*

GCK is the rate-limiting step in the metabolism of glucose and acts as the cellular sensor of glucose concentrations. Activation of GCK lowers the threshold for glucose-stimulated insulin secretion ('resetting' of the glucose-stimulated insulin release threshold), thus causing hypoglycaemia. The first activating mutation in the *GCK* gene was Val455Met which was a single-base change, resulting in the substitution of methionine for valine at codon 455. When expressed in vitro, the Val455Met mutation increased the affinity of GCK for glucose. Several other patients with activating mutations in the *GCK* gene have now been reported all responsive to medical therapy with diazoxide(Christesen, Tribble et al. 2008; van de Bunt, Edghill et al. 2008). However a case of severe HH due to a 'de novo' mutation in *GCK* gene (Y214C) was reported which failed to respond to medical therapy. Functional studies of this mutant showed a sixfold increase in its affinity for glucose, and histology of the resected pancreas in this patient revealed abnormally large and hyperfunctional islets. It is unclear, why this patient failed to respond to diazoxide, one possibility being that the dose of diazoxide was insufficient.

4.3 *Short- chain 3 hydroxyacyl Coenzyme A dehydrogenase (SCHAD) Deficiency*.

Human SCHAD deficiency is an inherited defect in mitochondrial fatty acid oxidation. Various clinical presentations of this metabolic disorder have been described, including hypoglycemia, hepatoencephalopathy, myopathy or cardiomyopathy, and even sudden death. Urine organic acid analysis showed highly elevated levels of 3-hydroxybutyrate, and 3-hydroxydicarbonylic acids of chain lengths $C_6$-$C_{14}$, with relatively larger amounts of shorter chain length intermediates. Enzyme studies based on the reduction of 3-ketoacyl-CoAs revealed, in most cases, a drastic reduction of SCHAD activity towards acetoacetyl-CoA in liver, but not in skeletal muscle. However, an abnormal enzymatic activity may not be detected by assays with homogenates of fibroblasts, because the HAD activity in skin fibroblasts is about one-tenth of that in liver. In humans, SCHAD levels in liver and heart appear to be much higher than those in skeletal muscle. The forms of SCHAD present in liver and heart proved to be identical. However, a second isoform of SCHAD has now been cloned from skeletal muscle (GenBank accession number AF001903). Moreover, it appears that skeletal muscle may contain SCHAD isoforms that differ both quantitatively and qualitatively from liver isoforms (He and Yang 2007).

The alternative splicing machinery, which is tissue-specific and development-dependent, is probably responsible for differences between liver and skeletal muscle SCHADs. Alternative splicing is involved in many human diseases, and most likely underlies different types of SCHAD deficiency, e.g. muscle-type vs. liver-type. These metabolic diseases, with high phenotypic variations, appear to be of autosomal recessive inheritance, as heterozygosity was demonstrated in first-degree relatives. The nine exons of the *HADHSC* gene cover 45 kb of DNA (extracted from the NCBI database, the Human Gene View of the Sanger Institute). Because the coding sequence accounts for only about 2% of the gene, mutations in introns as well as in the regulatory region of this gene may occur in patients with SCHAD deficiency (van Hove, Hansen et al. 2006).

SCHAD, encoded by the *HADHSC* gene, is an intramitochondrial enzyme that catalyzes the penultimate reaction in the β-oxidation of fatty acids, the $NAD^+$-dependent dehydrogenation of 3-hydroxyacyl-CoA to the corresponding 3-ketoacyl-CoA. So far 3 patients with mutations in the *HADHSC* gene and HH have been reported. The clinical presentation can be heterogeneous, either with mild late-onset hypoglycaemia or severe neonatal hypoglycaemia. The acylcarnitine profile in all reported patients has demonstrated raised hydroxybutyrylcarnitine levels, and urinary organic acids showed raised 3-hydroxyglutarate concentrations with decreased expression and function of SCHAD.

The mechanism of how a defect in the *HADHSC* gene leads to dysregulated insulin secretion is unclear at present (Hardy, Hohmeier et al. 2007). In the pancreas the acivity of SCHAD is highest in the Islet of Langerhans and more specifically in the pancreatic β-cells, suggesting that fatty acid oxidation may have a role in regulating insulin secretion. Fatty acids increase insulin secretion by affecting the concentrations of long-chain fatty acyl derivatives as a result of the inhibitory effect of citrate and malonyl-CoA on the rate-controlling carnitine palmitoyltransferase-1 and by subsequent stimulation of G – protein couple receptor triggering the activation of L- type Ca++ channels , but it is still unclear how defects in SCHAD lead to unregulated insulin secretion. Interestingly, Foxa2 *(HNF3)* has recently been shown to be involved in regulating the expression of the *HADHSC* gene, with studies in Foxa2-deficient -cells showing a threefold downregulation of *HADHSC* gene transcripts along with the ability of Foxa2 to bind to and activate this gene (Yang, He et al. 2005).

SCHAD exhibits a two-domain topology, with the N-terminal domain of the monomer (residues 12-201) having a --- fold similar to $NAD(P)^+$-binding enzymes. It consists of a core eight-stranded -sheet flanked by -helices. The first six strands of the sheet are in a parallel conformation as observed in a typical Rossmann fold. The final two strands are also parallel but run in the opposite direction relative to the first six strands. An interesting feature of the N-terminal domain is the large helix-turn-helix "tail" which connects strands 2 and 3 of the central -sheet. This motif extends from the $NAD^+$-binding domain and contains numerous charged amino acid side chains. The amino acid sequence in this fingerlike protrusion is apparent in the 2 and 3 regions as given in Figure 3.The C-terminal domain (residues 207-302) consists primarily of -helices and is involved in subunit dimerization and, presumably, binding of acyl-CoA substrates. It contains a bundle of 5 -helices and a small $3_{10}$-helix (14). The orientation of these helices relative to one another is critical as judged by the number of conserved glycine residues flanking -helical structural elements (F. In particular, glycine residues at positions 225, 241, 245, and 254 may provide the flexibility necessary for proper packing and/or formation of the -helices. Three highly conserved aspartic acid residues are found in this domain. Asp 233 is located in the middle of helix 11, Asp 251 at the end of helix 12, and Asp 256 at the beginning of helix 13. A carboxylate oxygen from Asp 233 and Asp 251 hydrogen bonds to the hydroxyl group of Tyr 299 and Tyr 301, respectively, while Asp 256 forms a salt bridge with Lys 293. These interactions apparently stabilize the hairpin-like motif found in the C-terminal tail. This unusual secondary structure is possible due to the presence of several conserved glycine residues, Gly 291, Gly 295, and Gly 297.The two domains are connected by a short linker region (residues 202-206) which contains a highly conserved PGF sequence (Barycki, O'Brien et al. 1999).

Subunit dimerization is mediated primarily by hydrophobic interactions between the C-terminal domains. Helix α10 (residues 207-223) of one subunit runs nearly antiparallel to helix 10 of the other subunit to form the core of the dimer interface.Helix 11, corresponding to residues 229-240, makes numerous hydrophobic interactions with the final two β-strands, strands 7 and 8, of the N-terminal domain of the opposing monomer. Additional hydrophobic contacts are provided by the small helical domain, residues 273-275, of each subunit. The dimer interface is maintained by the presence of four hydrogen bonds and two salt bridges (Barycki, O'Brien et al. 2000).

*4.3.1First HADH mutation Description*

Sequence analysis of genomic DNA of the patient, revealed a homozygous C→T point mutation in exon 7 of *SCHAD.*. The mutated nucleotide corresponds to nucleotide 773 of the coding sequence (taking the A of the translation start codon as nt 1). Her parents are heterozygous for the C773T mutation . The The C773T mutation leads to replacement of Pro258 (Pro246 of the mature SCHAD protein) by Leu. Pro258 is completely conserved in SCHAD sequences from different species. The short-chain HAD activity measured in patient' fibroblasts was significantly lower than that in control. To assess the effect of the P258L substitution on the activity of the protein, the wild-type and the mutated protein were synthesized in vitro from plasmids containing the wild-type SCHAD coding sequence and the SCHAD coding sequence harboring the C773T point mutation, respectively (Clayton, Eaton et al. 2001).

The absence of amplification artifacts in the wild-type construct and of mutations other than the C773T point mutation in the patient construct was ascertained by sequencing of the SCHAD expression plasmids in two directions. Expression of the wild-type SCHAD protein in a reticulocyte lysate system yielded a protein with an apparent molecular mass of approximately 35 kDa, which is in agreement with the calculated molecular mass of the wild-type protein of 34.3 kDa. The in vitro expressed mutated protein reacted strongly with the anti-SCHAD antibody and had an apparent molecular mass 0.5 kDa less than that of the wild-type protein. Immunoblot analysis showed that comparable amounts of protein were obtained for the wild-type and mutated SCHAD protein, whereas the vector alone did not yield a positive band of similar molecular mass. SCHAD activity was assayed in three separate experiments. The wild-type SCHAD exhibited a distinct but variable activity, whereas the mutated SCHAD did not yield activity above background in all experiments performed, indicating that the P258L substitution abolishes enzyme activity of in vitro expressed SCHAD (Clayton, Eaton et al. 2001).

### *4.3.2 Second HADH mutation Description*

A deletion of the six base pairs CAGGTC was discovered at the start of *HADHSC* exon 5. As expected from the microsatellite marker analysis , the proband and his affected sister were homozygous and the parents were heterozygous for the mutation. The six-base pair deletion removes the acceptor splice site adjacent to exon 5 of *HADHSC*. To examine the molecular consequence, mRNA from fibroblasts obtained from the two patients and their parents was analysed and showed that exon 5 is skipped during the mRNA splicing process, so that exon 4 is coupled directly onto exon 6. This leads to an in-frame deletion of 90 nucleotides in the mature mRNA, resulting in a protein product predicted to lack 30 amino acids. The impact of the mutation on protein function was examined by determining SCHAD activity in fibroblast cultures and in the two affected subjects and the parents, enzyme activity was <10% and <60% of control values, respectively . For comparison, long-chain 3-hydroxyacyl-CoA dehydrogenase activity was assayed, and abnormal activity ratios were found for both siblings, but not their parents. A plasma analysis of the distribution of fatty acyl groups conjugated to the transport molecule carnitine showed a normal pattern, except for 3-hydroxybutyryl(C4)-carnitine . This substance is barely detectable in normal control subjects, but reached a plasma concentration around 1 µmol/l in both patients. Finally, metabolic screening using gas chromatography-mass spectrometry on urine samples from the family was performed. One strong peak, corresponding to 3-hydroxyglutaric acid, was observed in both patients, but not in their parents or healthy siblings (Molven, Matre et al. 2004).

### *.4.3.3 Third HADH mutation Description*

The patient was found to be homozygous for an IVS6-2 a→g mutation and both parents are heterozygous for the IVS6-2 a→g mutation. The mutation was not found in 100 control subjects. Electrophoresis of the DNA fragments from the restriction fragment length analysis indicated that the IVS6-2 a→g mutation, in combination with NCO1 digestion, yields two bands of 59 and 17 nucleotides compared with 76 nucleotides in the control (undigested PCR product).Amplification of SCHAD cDNA from patient's fibroblast mRNA yielded no detectable products of exon 27, and a decreased amount of product when regions of exon 2 or exon 4 sequences were amplified. Reamplification of exon 7 by half-nested PCR yielded various PCR products from the patient's mRNA, including 269 nt, which could not be detected in control fibroblast. On reamplification and sequencing, the 270 nt product was seen to consist of at least three different sequences. The normal SCHAD cDNA sequence could be deduced from the mix of sequences, indicating that at least some normal SCHAD mRNA can be formed in patient fibroblasts (Hussain, Clayton et al. 2005).

# 5. Exercise-Induced HH.

In exercise-induced HH, strenuous physical exercise leads to inappropriate insulin release from β-cells causing postexercise hypoglycaemia. These patients show an increased insulin secretion in response to intravenous pyruvate administration in comparison with control patients . The molecular mechanism/s of how exogenous pyruvate triggers an inappropriate insulin secretion in these patients is/are still unclear (Hussain 2007).

# 6.CHI Anatomopathology

There are 2 main anatomical forms fo CHI : one focal and the other diffuse .

Many radiology investigations have been done to try to distinguish between the 2 anatomical forms: transhepatic catheterization of the pancreatic veins is dependent on radiologist training but defines the anatomical location well for the surgeon, whereas calcium injection into the 3 pancreatic arteries (splenic, gastroduodenal, superior mesenteric), followed by blood sampling showing an insulin increase, is less reliable. The new reliable and noninvasive method is positron emission tomography (PET) scanning with injection of [$^{18}$F]-l–dihydroxyphenylalanine (DOPA), which produces a single bright spot in the focal form. However, the anatomical localization for the surgeon is poor, which is why a combined PET/computed tomography image is strongly recommended (Hussain, Seppanen et al. 2006; Ribeiro, Boddaert et al. 2007; Ribeiro, Boddaert et al. 2007; Kapoor, Gilbert et al. 2008).

Surgery is indicated only for disease resistant to medical treatment when there is a high degree of suspicion that the focal form is present. Surgery is the only source of reliable pathologic data (Valayannopoulos, Vaxillaire et al. 2007). In the beginning, the focal and diffuse forms were operated on with equal frequency, whereas today, focal forms are the most frequent indication for surgery because they are cured by removal with less pancreatic insufficiency and no diabetes (Hardy, Hernandez-Pampaloni et al. 2007; Hardy, Hernandez-Pampaloni et al. 2007). Focal disease has not recurred postoperatively with a follow-up of more than 20 years, demonstrating that the disease is limited. Some older children have been able to attend prestigious schools, showing that they have no brain damage. Focal forms are always limited to a small area (5 mm). For example, when the pattern

observed in 2 or more of the biopsy specimens is small round islets with little cytoplasm or nuclei touching each other in rows with the same small regular size, crowded or at rest, the picture favors a focal form of the disease somewhere else. By contrast, the tissue in the focal form is composed of either ramified or branched areas of clear endocrine cells, frequently surrounded by a thin rim of acini at the periphery (Delonlay, Simon et al. 2007). This picture can involve multiple adjacent lobules with mainly abnormal $\beta$ cells occupying their centers. Sometimes, the lesion appears as a polyp at the pancreas surface, which is easy for the surgeon to find, or it may be deeper inside the pancreas, making it difficult to locate and resect with safe margins to avoid postoperative relapse secondary to incomplete removal. More rarely, the abnormal foci are extrapancreatic. The endocrine cells themselves have abundant cytoplasm with nuclei far apart and of irregular size and, here and there, nuclei of irregular angulated shape that are more than 3 to 4 times the size of nearby acinar nuclei taken as controls .

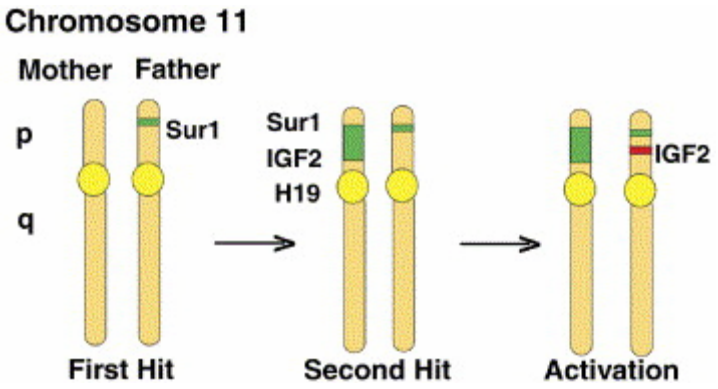
Diffuse form" means disease involving the whole pancreas, with variable involvement of islets. The pattern observed is the same on the 3 biopsy specimens with a preserved islet pattern containing very active $\beta$ cells with abundant cytoplasm and highly abnormal nuclei 3 to 4 times the size of acinar nuclei and having an irregular angulated shape. The abnormal nuclei are present in some of the islets only but can be seen in all 3 sampled areas, proving that the disease involves the whole pancreas, which is why it is called diffuse (Delonlay, Simon et al. 2007) .

*6.1 Focal form*

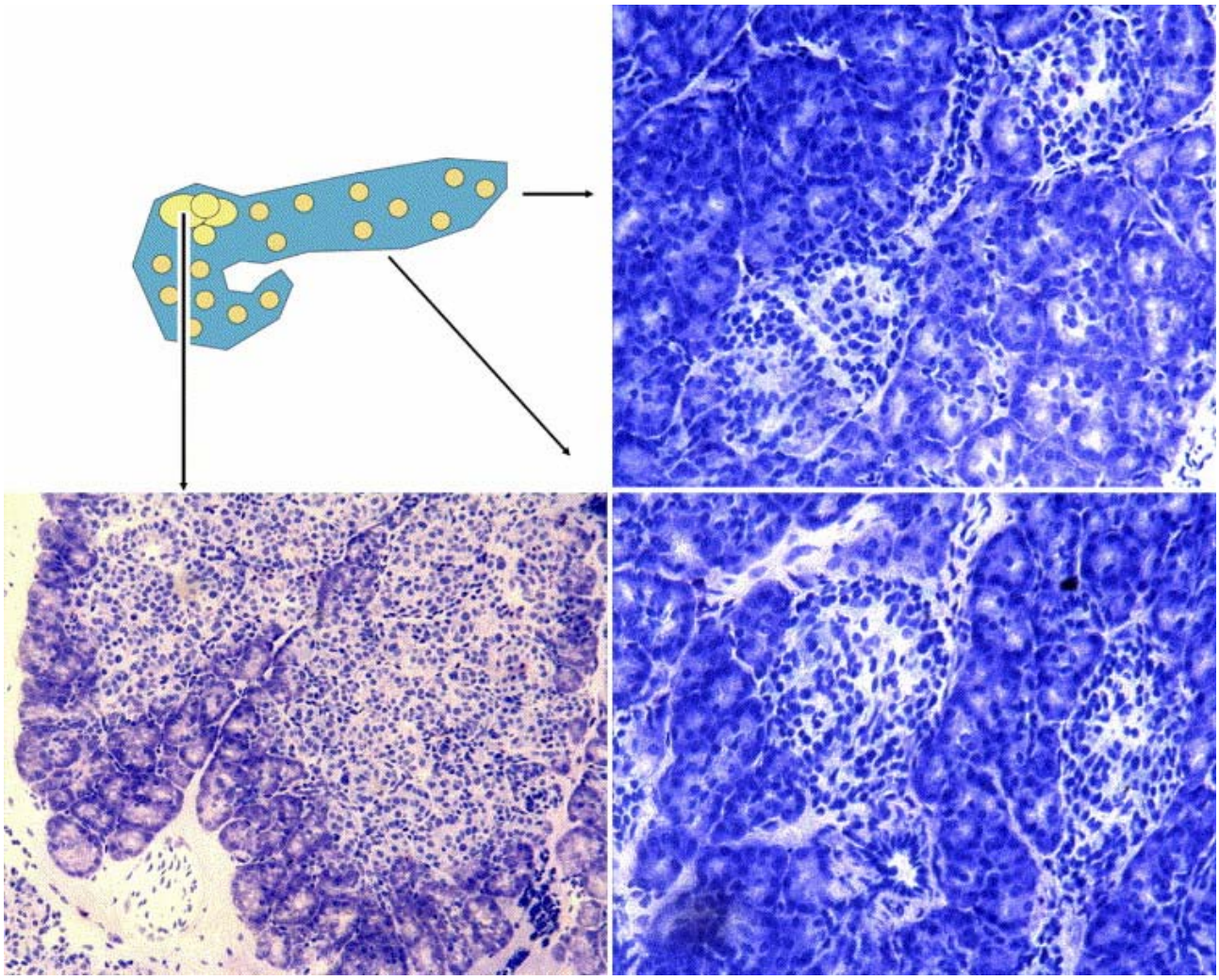
In HES (hematein eosin safran), the endocrine pancreas is always difficult to analyze, and Giemsa staining shows a pattern close to that obtained with toluidine blue on frozen section [22], [23], [24] and . The foci are composed of large endocrine cells with dispersed abnormal nuclei. Most of the time, a focus contains multiple abnormal adjacent lobules frequently intermingled with or surrounded by acinar foci. Somatostatin detection by immunohistochemistry reveals a second endocrine-cell population within the foci. That is, the focal form is not composed of $\beta$ cells only. It is poorly delimited, but an organoid pattern remains, suggesting an abnormal developmental process more than a tumor because it does not invade or push at the margins, and there is no pseudocapsule. The area of abnormal pancreatic development is multilobular and can have some satellite lobules in the nearby pancreas that necessitate intraoperative margin analysis to ensure complete excision and avoid recurrence (Damaj, le Lorch et al. 2008).

*6.2 Diffuse form*

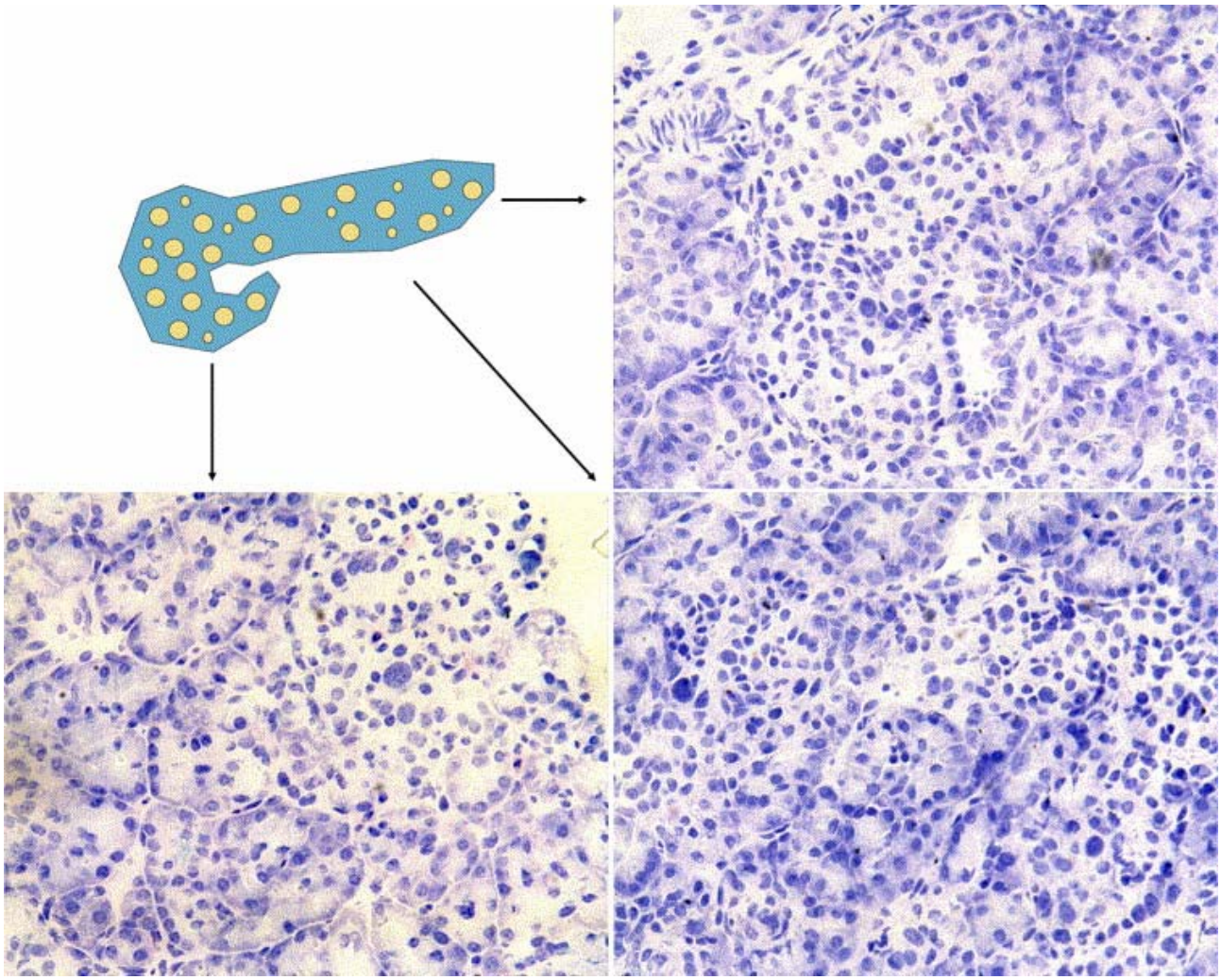
Voluminous abnormal extralobular islets are often present, associated with abnormal intralobular islets of hyperfunctional $\beta$ cells, some having huge irregular nuclei. Frequently, intralobular islets are less voluminous with fewer vessels but are more numerous. Independent endocrine-cell foci are also frequent and have highly abnormal nuclei. Their $\beta$-cell nature is confirmed by immunohistochemical detection of proinsulin. The pattern is similar in all the specimens taken at pancreatectomy. The disease involves the whole pancreas with variations in intensity from one islet to the other (de Lonlay, Simon-Carre et al. 2006).




Schematic drawing of multistep events leading to focal form of hyperinsulinism. First step is a mutation, inherited from the father, on chromosome 11p15.5. Second step is an LOH of the corresponding region on maternal chromosome shown by microsatellite study. Third hypothetical step is *H19* deletion on maternal chromosome, removing parental imprinting of *IGF2* on paternal chromosome and causing cell multiplication**.**



Representation of frozen sections obtained from 3 pancreas sites during surgery for a focal form of hyperinsulinism in the head. Specimens taken from tail and body show islets of Langerhans at rest with little cytoplasm, leading to crowded nuclei, whereas the focal form aspect from the head is that of multilobular involvement with local signs of *β*-cell hyperactivity: abundant cytoplasm and big abnormal nuclei (toluidine blue stain; original magnifications ×200 and ×100).



Representation of frozen sections obtained from 3 pancreas sites during surgery for diffuse form of hyperinsulinism. On each biopsy, there is 1 islet showing hyperfunctional signs with abundant cytoplasm and irregular nuclei more than 4 times the size of the acinar nuclei nearby taken as internal control. The disease involves the whole pancreas and, in consequence, can be called diffuse, which does not mean that all islets are involved in the same manner (toluidine blue stain; original magnification ×200).

# 7. Fatty Acid Signaling in the ß-Cell and Insulin Secretion

The FA supply to the ß-cell can be from exogenous sources such as plasma FFAs and lipoproteins or endogenous sources such as intracellular TG and phospholipid stores. It is worth noting that islet tissue expresses lipoprotein lipase . The islet, therefore, can access plasma TG as a source of FFAs, such that the FFA concentration in the immediate vicinity of ß-cells is likely to be higher than that measured in plasma. The actual concentration of FFAs that ß-cells are normally exposed to, however, is not known. The importance of regulated access to endogenous lipids for normal ß-cell function is increasingly being realized, with particular interest currently being directed to the lipase enzymes that may be involved.

## *7.1 Nutrient-secretion coupling: triggering and amplification pathways*

Islet ß-cell glucose metabolism is essential for the coupling of glucose sensing to insulin release. It is well accepted that its metabolism through pyruvate to acetyl-CoA with subsequent mitochondrial oxidation increases the ATP/ADP ratio, which results in closure of ATP-sensitive $K^+$ ($K_{ATP}$) channels, depolarization of the plasma membrane, opening of voltage-dependent $Ca^{2+}$ channels, and $Ca^{2+}$ triggering of insulin granule exocytosis. This pathway, often termed the $K_{ATP}$ channel–dependent pathway, is considered to be the major triggering pathway for GSIS). In addition, pyruvate from glucose can be channeled via pyruvate carboxylase into the anaplerosis pathway, which can affect insulin secretion by increasing the levels of metabolism-derived signaling molecules such as NADPH from the malate-pyruvate shuttle, citrate cataplerosis, glutamate, and lipid signaling molecules from the malonyl-CoA/LC-CoA pathway. These latter pathways can be considered to be amplification pathways. Of particular interest is the role of the malonyl-CoA/LC-CoA pathway on lipid signaling.(Palladino, Bennett et al. 2008)

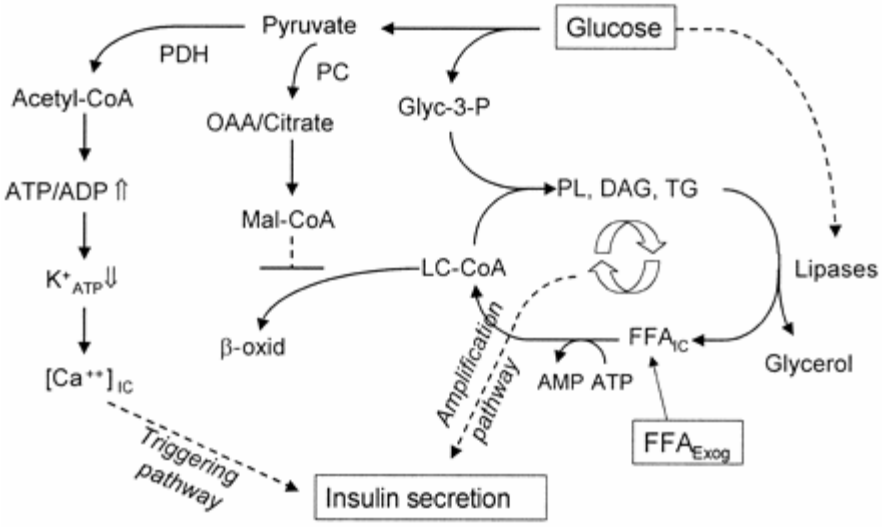
## *7.2 Malonyl-Coa/LC-CoA pathway of lipid signaling*

The malonyl-CoA/LC-CoA model of GSIS predicts that malonyl-CoA, derived from glucose metabolism via anaplerosis/cataplerosis, inhibits FA oxidation by allosteric inhibition of carnitine palmitoyl-transferase (CPT)-1, thereby increasing the availability of LC-CoA for lipid signaling to cellular processes involved in exocytosis. In this model, the effectiveness of malonyl-CoA to promote insulin release depends on both a cytosolic $Ca^{2+}$ rise caused by the $K_{ATP}$-dependent pathway and the prevailing availability of FAs to the ß-cell. Malonyl-CoA levels are also influenced by the activity of AMPK. AMPK senses cellular energy status and is activated by an increase in the AMP/ATP ratio

brought on by fuel deprivation, fasting or exercise. AMPK activates cellular energy production (e.g., glucose oxidation and FA oxidation) and reduces energy consumption (e.g., FA synthesis and esterification). Malonyl-CoA, on the other hand, is a "signal of plenty" and promotes nutrient storage, including FA esterification. AMPK phosphorylates acetyl-CoA carboxylase and malonyl-CoA decarboxylase, the enzymes that regulate malonyl-CoA synthesis and degradation, respectively, with the resultant effect of lowering malonyl-CoA). In addition to conditions of food deprivation and exercise, AMPK can be activated by adipokines, including adiponectin and leptin, and pharmacological agents such as metformin and thiazolidinediones. The evidence supportive of a role for malonyl-CoA/LC-CoA signaling in the regulation of insulin secretion is substantial. Malonyl-CoA levels have been documented to increase in response to glucose, and this precedes GSIS in ß-cells. Elevated glucose causes inhibition of FA oxidation and stimulates FA esterification processes (Palladino, Bennett et al. 2008).

*7.3 TG/FFA cycling and insulin secretion*

In addition to FA oxidation and esterification, lipolysis is the third major pathway of intracellular FA partitioning. Lipolysis of intracellular TG refers to the hydrolytic removal of the fatty-acyl chains from the glycerol backbone by lipase enzymes. Of particular interest to this discussion on FA signaling is the recent finding that both FA esterification and lipolysis processes are glucose-responsive in the ß-cell . Glucose increased lipolysis, as determined by glycerol release, in islets from both wild-type and hormone-sensitive lipase (HSL) null mice. Unclear, however, is the mechanism by which TG/FFA cycling, including lipolysis, contributes to FFA amplification of GSIS. Elevated glucose, particularly in the presence of exogenous FFAs, will result in increased levels of all lipid moieties within the cycle, including LC-CoA, DAG, phospholipids, and FFAs (Shanik, Xu et al. 2008).




Interactions between glucose and fatty acid metabolism in nutrient-secretion coupling. This model illustrates the synergistic interaction between glucose and FA metabolism and the generation of lipid signaling molecules that augment GSIS. Glucose gives rise to pyruvate, which, when channeled through pyruvate dehydrogenase (PDH), contributes to induction of insulin secretion via ATP production and the $K_{ATP}$-dependent triggering pathway. Pyruvate alternatively can be channeled via pyruvate carboxylase (PC) into the anaplerosis/cataplerosis pathway, which contributes to increases in cytosolic oxaloacetate (OAA) and citrate. Glucose increases malonyl-CoA (Mal-CoA), which blocks FA oxidation by inhibiting CPT-1. Inhibition of FA oxidation allows LC-CoA esters to accumulate in the cytosol. LC-CoAs are formed from FFAs supplied externally or produced by the lipolysis of endogenous TG. LC-CoAs can be esterified with glycerol-3-phosphate (Glyc-3-P) to form complex lipids such as TG, DAG, and phospholipids (PL). Glucose also activates lipolysis, which favors TG/FFA cycling from endogenous lipids. Exogenous FFAs ($FFA_{Exog}$ contribute by amplifying the levels of cycle intermediates (LC-CoAs, TG, DAG, PL, and intracellular FFAs [$FFA_{IC}$ themselves), all of which may be implicated in lipid signaling effector pathways. ß-oxid, ß-oxidation; $[Ca^{++}]_{IC}$, intracellular calcium.

Insulin vesicle exocytosis is a complex process involving many steps, including vesicle movement, docking, priming, and finally fusion with the plasma membrane. Some of these steps could be modulated by TG/FFA cycle intermediates acting as lipid signaling molecules. For example, LC-CoA can be used to acylate proteins, such as the synaptosomal-associated protein and synaptogamin, which can enhance their association with target membranes. Interestingly, LC-CoA has also been shown to activate lipases from islet tissue. As such, a mediator of glucose-induced lipolysis could be elevated LC-CoA themselves. DAG, the levels of which rise in the ß-cell in response to glucose, not only activate protein kinase C, which is implicated in insulin secretion. TG/FFA will also affect membrane glycerophospholipid metabolism, which could influence secretion via alteration of membrane physicochemical properties. The glycerophospholipids may also have more direct effects, It may be that the TG/FFA cycle is a means of targeting the delivery of FFAs, and perhaps specific FFAs such as arachidonic acid (AA), to a particular subcellular site within the ß-cell. Of relevance to this, HSL has

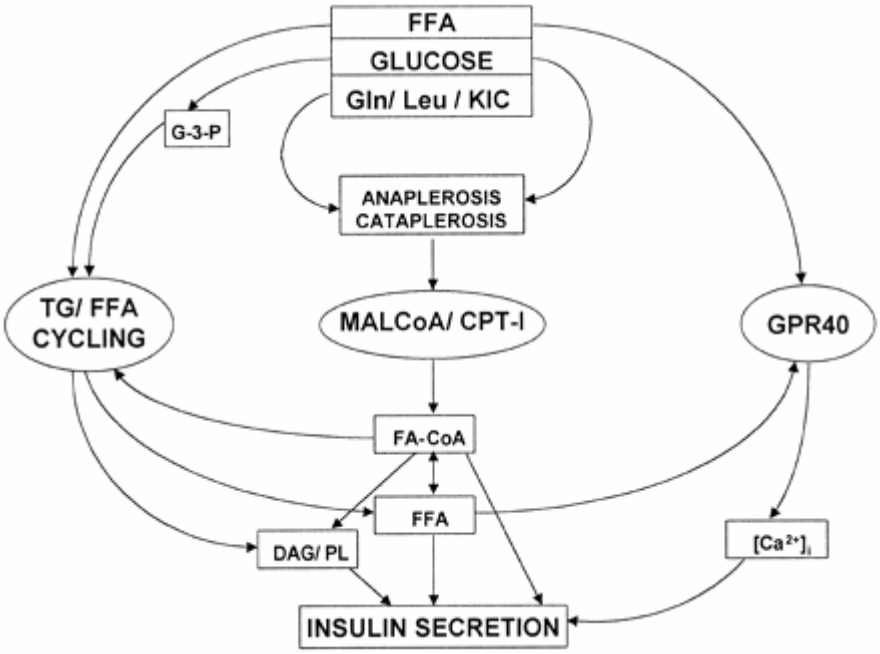
been shown to be colocalized with insulin vesicles. The TG/FFA cycle of ß-cells, therefore, appears to be a major and hitherto overlooked pathway for both glucose and FFA metabolism. There are two conceivable advantages for the ß-cell to use glucose in this way. First, glucose usage via the TG/FFA cycle provides an alternative pathway by which glucose metabolism may be coupled to insulin secretion, which, via the provision of lipid-signaling molecules, allows amplification of the triggering pathway for exocytosis. The second is related to the preservation of ß-cell mass in the face of fuel surfeit as occurs in obesity and diabetes. Thus, nutrient toxicity to cells occurs when the flux of metabolites through mitochondrial oxidation is high, since this causes the production of damaging superoxides and reactive oxygen species.

Thus, glucose flux through TG/FFA cycling is a means of nutrient-secretion coupling that bypasses the need for oxidation to cause secretion and allows for detoxification of glucose carbons in the form of glycerol that is released from the ß-cell.. Furthermore, use of ATP by the cycle will reduce the mitochondrial membrane potential and, in this way, lessen the rate of reactive oxygen species production (Nolan, Madiraju et al. 2006).

.

*7.4 Refined model of lipid signaling in nutrient-secretion coupling*

We now have a more advanced, but as yet incomplete understanding of lipid signaling in nutrient-secretion coupling for insulin secretion. FFAs are involved in amplification of GSIS that is triggered by the $K_{ATP}$-dependent pathway. This involves indirect signaling that requires intracellular FA metabolism, which implicates both the anaplerotic/malonyl-CoA pathway and intracellular TG/FFA cycling. The model needs to be extended, however, to include direct signaling via FFAR1/GPR40. Lipid signaling in the ß-cell involves three arms (trident model) that comprise signaling via glucose-derived malonyl-CoA, TG/FFA cycling including lipolysis, and FFAR1. In this view, signaling via the AMPK/malonyl-CoA/CPT-1 network and TG/FFA cycling are tightly linked processes. Glucose promotes activity in the TG/FFA cycle by elevating malonyl-CoA, which inhibits partitioning of LC-CoA to FA oxidation (via CPT-1 inhibition), such that LC-CoAs are more available for esterification processes. Glucose also provides the glycerol-3-phosphate necessary for FA esterification into complex lipids. Lastly, glucose activates lipase enzyme(s), possibly through a direct effect of LC-CoA on the activity of lipases, although confirmation of this and assessment of other potential mechanisms are needed. A reduction in AMPK activity promoted by fuel stimulation that would enhance ß-cell lipolysis is an alternative possibility in view that AMPK activation has been associated with reduced lipolysis in adipocytes. While glucose, by activating lipolysis, promotes TG/FFA cycling from endogenous lipid stores, the provision of exogenous FFA supply to the ß-cell will amplify the pathway, increasing the intracellular concentrations of the lipid signals LC-CoA, DAG, phospholipids, and unesterified FFAs (Nolan, Madiraju et al. 2006).



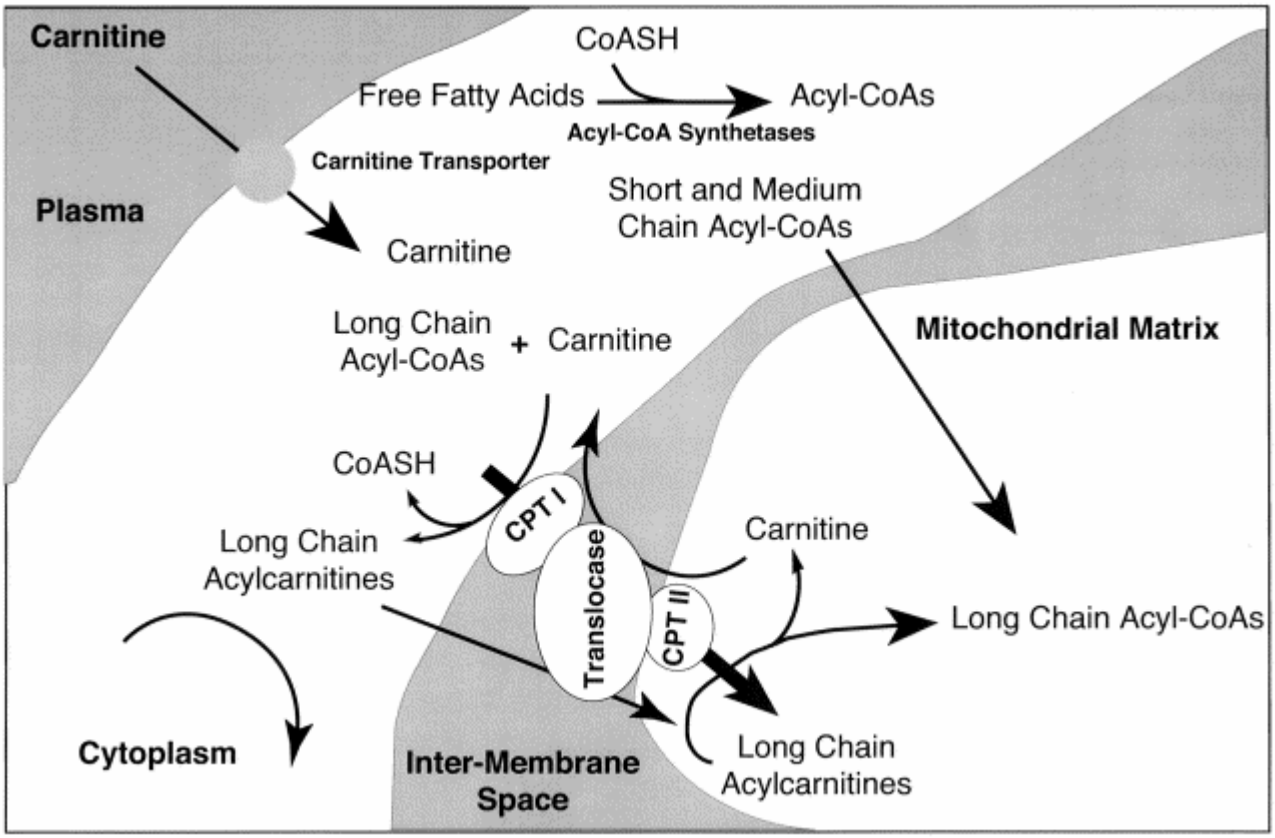

The trident model of ß-cell lipid signaling. Three interdependent arms of lipid signaling are proposed by which FFAs augment insulin secretion. First, glucose and other nutrient secretagogues such as glutamine (Gln), Leucine (Leu), and α-ketoisocaproate (KIC) contribute to anaplerosis, which allows cataplerotic efflux of citrate from mitochondria. This results in malonyl-CoA (MALCoA) formation, inhibition of CPT-1 activity, and FA oxidation, and accumulation of LC-CoAs that stimulate insulin secretion directly or by the formation of complex lipids such as DAG and various phospholipids (PL). The second arm involves glucose-responsive TG/FFA cycling, due to the effects of glucose to concomitantly promote FA esterification (provision of glycerol-3-phosphate [G-3-P] and Mal-CoA inhibition of FA oxidation) and lipolysis processes. This allows, particularly in the presence of exogenous FFAs, for the accumulation of cycle intermediates (LC-CoA, DAG, PL, and FFA) that have signaling roles. Third, exogenous FFAs activate the cell surface FA receptor, FFAR1/GPR40, which causes an increase in intracellular $Ca^{2+}$. FFAs formed from the hydrolysis of TG can cross the cell membrane and, together with exogenous FFAs, activate FFAR1. Thus, the "trident pathways" of lipid amplification intercommunicate and synergize to promote insulin secretion.

# 8.Defects of mitochondrial β-oxidation

β-Oxidation, which results in sequential cleavage of two carbon units from fatty acid, represent an important source of energy for the body during times of fasting and metabolic stress. Free fatty acids, release into the blood via catabolism of fat stores , or from dietary sources are metabolized into two intracellular compartments: the peroxisomes and the mitochondria. Mitochondria are complex structures bounded by two lipid bilayer membranes, containing their own genetic information on a circular chromosome which encodes 12 protein subunits for the oxidative phosphorylation complexes along with genes for mitochondrial specific ribosomal and tRNAs. Most proteins found in the mitochondria, however, are nuclear encoded in a larger precursor form containing information in an amino terminal signal peptide necessary for targeting the proteins to the mitochondria (Kompare and Rizzo 2008).

Mitochondrial fatty acid oxidation is predominantly responsible for the oxidation of fatty acids of carbon length 20 or less, while the peroxisomal pathway is physiologically more relevant for longer chain fatty acids. Mitochondrial β-oxidation is a complex process involving transport of activated acil-CoA moieties into the mitochondria, and sequential removal of two carbon acetyl-CoA units. These in turn are used as fuel for the tricarbooxylic acid cycle or the production of keton bodies. The end result is the generation of reducing equivalents, which are funneled into the electron transport chain and ultimately lead to the production of adenosine triphosphate ( ATP). Ketones derived from hepatic fat metabolism can be as an auxiliary fuel by most tissues, including the brain.(Gregersen, Andresen et al. 2008)

At least 25 enzymes and specific transport proteins are responsible for carrying out the steps of mitochondrial fatty acid metabolism and, of these, defects in at least 22 have been shown to causes disease in humans. These disorders include enzymes of β-Oxidation, defects in fatty acid and carnitine transport and acyl – CoA dehydrogenases (ACDs)(Mitchell, Gauthier et al. 2008).




Transporter proteins involved in mitochondrial β-oxidation.

## *8.1Defects in Medium chain acyl-CoA dehydrogenase* (ACADM)

Numerous diseases have been reported in relation to fatty acid, such as cardiovascular disease, cancer and diabetes. Fatty acid oxidation in mitochondria is an essential energy generation system for cells. During prolonged fasting and starvation , fatty acid are the precursors of keton bodies, which are an important alternate fuel in extrahepatic tissues at time when the supply of glucose is limited (Kozak, Hrabincova et al. 1999; Saudubray, Martin et al. 1999).

Medium-chain acyl-CoA dehydrogenase deficiency (MCADD;MIM# 201450) is the most frequent defect of fatty acid oxidation, with a significant morbidity and mortality in undiagnosed patients.

Clinical symptoms such as hypoketotic hypoglycemia, lethargy, encephalopathy, and seizures during intercurrent illness or prolonged fasting may occur at any age, but the vast majority of patients presents with symptoms during their second year of life. Approximately 20% of these patients die during their first metabolic crisis and about 40% of the survivors show irreversible neurological impairment or developmental delay. Follow-up data on symptomatic patients indicate that severe metabolic crisis and death can be prevented by appropriate counseling to avoid fasting and by simple dietary treatment, i.e., high carbohydrate intake during episodes of metabolic stress Therefore, early detection of MCADD in asymptomatic newborns by prospective newborn screening (NBS) is being propagated. Due to the advances in tandem mass spectrometry (MS/MS)-based technologies, NBS for MCADD by acylcarnitine profile analysis in dried blood spots are currently implemented in an increasing number of countries worldwide Elevated concentrations of octanoylcarnitine (C8) and other acylcarnitines in combination, with acylcarnitine ratios including hexanoylcarnitine (C6), decanoylcarnitine (C10), decenoylcarnitine (C10:1), C8/C6, C8/C10, and C8/C12 (dodecanoylcarnitine), are used to detect MCADD by various screening programs (Singh et al. 2002; Vockley and Whiteman 2002; Gregersen, Bross et al. 2004).

Although the method has been proven to be highly sensitive for detecting the biochemical hallmarks of MCADD, the confirmation of positive screening results may not always be definite, relying solely on biochemical metabolite analysis. Depending on the stringency of threshold criteria used in different NBS programs, some premature infants and heterozygous ACADM gene carriers can show elevations of the primary marker metabolite C8 that need to be differentiated from true MCADD Commonly used metabolite-based confirmation tests, such as the 3-phenylpropionic loading test and the analysis of urinary organic acids, are not always reliable. In rare instances, both have been reported to give normal results in newborns with blood acylcarnitine profiles clearly indicative for MCADD [Zschocke et al., 2001]. This reflects a gray area in the expression of MCADD, rendering both its definition as a disorder and the validation of MCADD population screening difficult. As a consequence, additional methods for confirmation or exclusion of MCADD are required, such as direct analysis of enzyme

activity, acylcarnitine profiling of cultured skin fibroblasts, or the identification of ACADM gene variations(Korman, Gutman et al. 2004; McKinney, Longo et al. 2004; Maier, Liebl et al. 2005).

Medium chain acyl CoA dehydrogenase (MCAD) catalyze the rate- determining step of fatty acid β-oxidation in mitochondria. There are two step in the MCAD- catalyzed reaction, which are called the reductive half –reaction and the oxidative half –reaction. In the first reaction substrate is dehydrogenated into 2- enoyl-CoA, and the resulting redox equivalents are transferred onto the cofactor FAD, which forms the reduced flavin. In the oxidative half-reaction , electron are transferred to electron transfer flavoprotein (ETF) and regenerate the reduced enzyme.

The crystal structure of medium chain acyl-CoA dehydrogenase shows that the binding cavity for the fatty acyl moiety is located between helices E and G and is lined with the side chains of Glu376, Tyr375, Val259, Thr168, Leu103, and Ala100. Glu99 and Thr96 form the bottom of the deep hole where the end of the alkyl chain lies. Tyr375 is conserved in all sequenced acyl-CoA dehydrogenases, and the plane of its phenolic side chain is tilted to face rather than to edge on to the alkyl chain of the substrate. Therefore, Tyr375 has relatively large contact area with the substrate and should be a very important amino acid residue for substrate binding. More than 30 mutations in this gene have been identified, and the most common mutation in this gene involves a A to G.transition at coding position 985of the gene ( K304E in the mature protein). Two copies of this common mutation are typically found in 80% of European origin individuals diagnoses with MCADD on the basis of clinical symptoms.(Maier, Liebl et al. 2005; Grosse, Khoury et al. 2006; Waddell, Wiley et al. 2006; Derks, Boer et al. 2008; Downing, Manning et al. 2008)

# 9.Homozygosity mapping

Single nucleotide polymorphisms (SNPs) are the most common sequence variation in the human genome; and are thought to be the differences in human traits and disease susceptibility within and among populations. This view was challenged recently by studies using new genome technologies that allow studying the whole genome simultaneously in many different individuals. A wealth of data has been generated and it strongly indicates that human diversity is much larger than the single nucleotide differences. Much of the human diversity might be explained by larger structural differences between individual genomes, such as copy number polymorphisms and varying numbers of deletions, insertions, inversions, and other major rearrangements. These structural polymorphisms may have a significant role in both the genetics of complex diseases and genome evolution (Iafrate, Feuk et al. 2004; Sebat, Lakshmi et al. 2004; Tuzun, Sharp et al. 2005; Conrad, Jakobsson et al. 2006; McCarroll, Hadnott et al. 2006).

Two recent reports describe the extent to which regions of extended homozygosity can be found in human subjects genotyped with large numbers of closely spaced single nucleotide polymorphism (SNP) markers (Gibson, Morton et al. 2006; Li, Ho et al. 2006; Simon-Sanchez, Scholz et al. 2007) . The first study reported 1393 tracts over 1 Mb in length among 209 subjects and the second found tracts extending over 5 Mb being present in 26 out of 272 subjects. Since the samples used were from apparently unrelated subjects in outbred populations one might not expect to find such regions by chance.

The authors of the first study (Gibson, Morton et al. 2006) thought that the tracts probably represented ancestral haplotypes but could not rule out deletions or uniparental isodisomy as alternative explanations.Likewise, the authors of the second study (Simon-Sanchez, Scholz et al. 2007) proposed that the finding represented "chance meiotic events in consanguinous parents".Through examination of the hybridization intensity of each SNP they were able to estimate copy number and hence to identify a number of structural genomic variants, which meant they could exclude the observed homozygosity as being due to deletions. However this would still not exclude the possibility that uniparental isodisomy had occurred, this being the other form of cytogenetic abnormality which can be a cause of extended homozygosity. They did note that subjects with one region of extended homozygosity were more likely to have an additional region than would be expected by chance and argued that this supported the hypothesis that such regions were due to parental consanguinity.Both studies were carried out on unrelated subjects and in such a sample it is hard to see how one could definitively rule this out such explanations.

An earlier study had used related subjects, consisting of 8 of the reference pedigrees from the Centre d'Etude du Polymorphisme Humain (CEPH). These were genotyped with 8000 short tandem repeat

markers. The authors identified numerous long homozygous segments. These were compatible with Mendelian transmission, indicating that they were indeed a result of autozygosity, that is, the subject receiving a copy of the same ancestral haplotype from each parent. Thus these stretches of homozygosity did not appear to result from cytogenetic abnormalities. However a relatively small number of markers was used yielding an average inter-marker spacing of slightly less than 0.5 cM. With this marker density it would be difficult to distinguish whether some relatively short stretches of homozygosity might or might not be due to mechanisms such as uniparental isodisomy.

A third work presents the results of an investigation of regions of extended homozygosity detected by densely spaced SNPs genotyped in a sample of CEPH trios (Curtis 2007). If such regions were the result of cytogenetic abnormalities they should be detectable as departures from Mendelian transmission. Although based on a small sample, these results conclusively demonstrate that regions of extended homozygosity are not usually due to uniparental isodisomy. If they were, we would expect that within these regions on many occasions an allele which was homozygous in a parent would fail to appear in a child. In fact, although a small number of such transmission errors do occur, they are observed more frequently outside regions of extended homozygosity than within them (Curtis 2007).

It is somewhat striking that no errors were observed within the 5 Mb regions, suggesting that even these long stretches of homozygosity reflect autozygosity rather than cytogenetic abnormality and implying that there are some very long haplotypes which are not uncommon in this outbred population. These results demonstrate that while it is fairly common for regions of homozygosity to extend over 1 Mb this is usually due to segments of such length being inherited from a common ancestor rather than being due to deletion or uniparental isodisomy. As pointed out previously, this will have implications for the ability to map diseases using association studies.

Li, Ho et al. 2006 discovered that most chromosomes in different ethnic groups contained long and contiguous segments of homozygous markers. Similar findings were reported by Gibson et al 2006. by analyzing data generated from the HapMap (www.hapmap.org) project. Such long segments of homozygosity could represent deletion polymorphisms, loss of heterozygosity (LOH), segmental uniparental disomy, or autozygosity. Their data strongly suggests these long contiguous stretches of homozygosity (LCSHs) most likely represent autozygosity, which is usually seen in related individuals or in an inbreeding population. Homozygosity tracts are surprisingly common and long in unrelated individuals from outbred populations.

The authors compared the LCSH regions identified in this study with the 20 LCSH regions published by Gibson et al. to determine if there were regions common in both studies. A total of 17 overlapping regions were identified in both studies, and most of them were very long (4.7672.98 Mb). However,

this comparison was not complete since there were only 20 longest regions reported in their work. Furthermore, they were not able to calculate the frequencies of each common region since detailed information on the number of individuals having overlapping regions were not specified in their work.
The LCSH was most likely to represent autozygosity, which is a term used to describe homozygosity for alleles at a locus that are identical by descent; that is, inherited from a common ancestor. This was supported by observation that the incidence of LCSH and the numbers of segments with LCSH increased in Taiwan aborigines, which is an isolated population, and in the offspring of consanguineous marriages. In addition, the long average length of LCSHs in Han Chinese (10.22 Mbp) also supported the idea of autozygosity rather than homozygosity by chance.On the other hand, recombination and random mating make autozygous segments a rare event in an outbred population.
The Han Chinese in Taiwan are descendant from a mixture of immigrants mainly from Northern, Southern, and Southeastern Mainland China, which were outbreeding populations. Thus, it is remarkable to see that LCSH occurs in 6.6% of the Han Chinese population.
The observation that all four Taiwan aborigines studied showed significant amounts of LCSH is of particular interest. It has been shown that within a subpopulation, the existence of a remotecommon ancestor(s) will change the heterozygosity frequency of individuals even if the subpopulation is under random mating. That is, individuals in this subpopulation are related to each other to some extent, and the existence of autozygosity in individuals is not necessary a result of inbreeding.The frequent occurrence of LCSH in outbred populations will potentially have a significant impact on all areas of genetic studies such as population genetics, disease gene studies, and cancergenetics.
In population genetics, identification of LCSH in an individual may reveal unreported parental relationships, define unclear genealogy, and refine population structure. Moreover, LCSH is a valuable resource for generating extended haplotype structure in the general population (Li, Ho et al. 2006).

### *9.1 Homozygosity mapping methodology*

Populations that are genetically isolated or have a high frequency of inbreeding tend to show an increased prevalence of recessive disorders. Historically, the study of inbred populations has been highly successful in the fine mapping of recessive traits. The effect of consanguinity reduces allelic and non-allelic heterogeneity and potentially increases linkage disequilibrium and hence the power to detect associations between genetic markers and disease. Likewise, mapping homozygous regions in affected individuals in consanguineous families is a powerful method of localising autosomal recessive genes. This technique known as homozygosity or autozygosity mapping assumes that affected offspring co-inherit two copies of a disease-related chromosomal segment from a common ancestor. Such positional cloning of recessive genes in families with rare genetic variants can provide

valuable insight into the localisation and identity of genetic susceptibility factors involved in the common types of disease with similar clinical features. The more closely related the parents are, the greater proportion of their offspring's genome is expected to be homozygous. For example, offspring of first-cousin marriages are expected to have 1/16 (6.25%) of their genome homozygous and for second cousins the fraction is 1/64 (1.56%). However, the greater the expected homozygosity in offspring, the greater the likelihood of identifying shared homozygous-by-descent (HBD) intervals in affected siblings who are unrelated to the 'disease'-causing locus (Woods, Cox et al. 2006).

**T**he recent development of microarray platforms, capable of genotyping hundreds of thousands of SNPs, has provided an opportunity to rapidly identify novel susceptibility genes for complexphenotypes. Studies employing genotyping microarrays have typically used a whole-genome association (WGA) approach, in which each SNP is examined individually for association with disease; multiple testing requires that statistical thresholds for WGA approach $10^{-7}$ or lower . Given the presumably polygenic nature of complex illness, this conservative strategy inevitably results in false negatives in the search for susceptibility genes .At the same time, structural properties of WGA datasets, including patterns of linkage disequilibrium (LD), have not yet been exploited in these analyses.

Consequently, an analytic approach, termed whole-genome homozygosity association (WGHA) was developed, which first identifies patterned clusters of SNPs demonstrating extended homozygosity and then employs both genome wide and regionally specific statistical tests for association to disease. In their study (Lencz, Lambert et al. 2007) used WGHA in a case-control dataset of patients with schizophrenia (SCZ) and healthy volunteers, genotyped at 500,000 SNPs, to detect novel susceptibility loci for SCZ.By contrast, WGHA presents an opportunity for rapidly identifying susceptibility loci broadly across the genome, yet with resolution sufficient to implicate a circumscribed set of candidate genes. WGHA is designed to be sensitive for detecting loci under selective pressure, and recent data suggest that signatures of evolutionary selection may be strongly observed in genes regulating neurodevelopment. Thus, WGHA may be particularly effective for SCZ, which is thought to have a primary pathophysiological basis in abnormal neurodevelopmental processes.

Regions of extended homozygosity across large numbers of consecutive SNPs form the basis of WGHA analysis. In general, extent of homozygosity is a function of LD within a chromosomal region, which in turn is a function of recombination rates and population history. Size and structure of LD blocks vary widely across the genome and across populations, and regions of extensive long-range LD may be indicative of partially complete selective sweeps of functional significance.. A logical consequence of such identity across unrelated chromosomes is that long stretches of homozygosity may be observed in healthy individuals from outbred populations lacking any known consanguineous parentage. However, the relative commonality of this phenomenon has not been systematically

documented in large datasets at high resolution. Using dense, whole-genome microarray SNP data, Lencz observed that ROHs ranging in size from 200 kb to 15 Mb were common even in healthy individuals from an outbred population (U.S.Caucasians residing in New York City/Long Island). These homozygous regions were both too common and too small to suggest recent consanguinity. These data are consistent with the possibility that common ROHs mark regions under selective pressure for several reasons; specifically contain genes recognized by other methods as under strong selective pressure in Caucasians, such as *SNTG1* and *ALDH2* (*iv*) are replicated in Caucasian HapMap samples; (*v*) are not replicated in non-Caucasian HapMap samples.In case-control comparison, we observed that ROHs were overrepresented in SCZ at a genomewide level.. Intriguingly, studies of population isolates and consanguineous families demonstrate elevated rates of schizophrenia. The presence of nine specific ROHs was associated with illness susceptibility both individually and cumulatively. Four of these regions implicated genes related to postsynaptic (largely glutamatergic) receptor complexes implicated in SCZ pathophysiology. It is noteworthy that most of the risk ROHs demonstrated low frequencies in the general population.

The identification of mutations in genes that cause human diseases has largely been accomplished through the use of positional cloning, which relies on linkage mapping. In studies of rare diseases, the resolution of linkage mapping is limited by the number of available meioses and informative marker density. One recent advance is the development of high-density SNP microarrays for genotyping.

The SNP arrays overcome low marker informativity by using a large number of markers to achieve greater coverage at finer resolution. Although considerable progress has been made, the majority of genes causing complex and Mendelian disorders have yet to be identified.

The discovery of most human disease-causing genes has relied on using large or multiple human pedigrees for genetic linkage mapping. The identification of additional disease-causing genes is hindered by the paucity of pedigrees that are suitable for traditional linkage studies.

Therefore, opportunities to identify novel disease genes that take advantage of smaller pedigrees and genomic resources must be sought. One recent development that can aid in the identification of disease genes are high-density SNP microarray genotyping for homozygosity mapping. This method was used in a small consanguineous Israeli Bedouin family with autosomal recessive Bardet–Biedl syndrome (BBS; obesity, pigmentary retinopathy, polydactyly, hypogonadism, renal and cardiac abnormalities, and cognitive impairment) in which previous linkage studies using short tandem repeat polymorphisms failed to identify a disease locus. SNP genotyping revealed a homozygous candidate region. Mutation analysis in the region of homozygosity identified a conserved homozygous missense mutation in the TRIM32 gene, a gene coding for an E3 ubiquitin ligase (Chiang, Beck et al. 2006). Functional analysis of this gene in zebrafish and expression correlation analyses among other BBS genes in an expression quantitative trait loci data set demonstrate that TRIM32 is a BBS gene. This

study showed the value of high-density SNP genotyping for homozygosity mapping and the use of expression correlation data for evaluation of candidate genes and identifies the proteasome degradation pathway as a pathway involved in BBS.

Furthermore Woods, Cox et al. found that in individuals with a recessive disease whose parents were first cousins, on average, 11% of their genomes were homozygous with each individual bearing 20 homozygous segments exceeding 3 cM and that the size of the homozygous segment associated with recessive disease was 26 cM These data imply that prolonged parental inbreeding has led to a background level of homozygosity increased ~5% over and above that predicted by simple models of consanguinity. This has important clinical and research implications. Individuals whose parents are related are expected to have an increased proportion of their autosomal genome that is homozygous.

The more closely the parents are related, the greater this effect is expected to be.1 Off-spring of second cousins are expected to have children with 1/64 of their genome homozygous; offspring of first cousins, 1/16; offspring of double–first cousins, 1/8; and offspring of incestuous union, 1/4.2 Furthermore, in the case of first-cousin offspring, it has been calculated that the average homozygous segment will be 20 cM.3 This degree of homozygosity is far greater than that seen in apparently outbred populations. The clinical con-sequence of this is an increased incidence of autosomal recessive diseases (and, to a lesser extent, homozygous dominant diseases) in consanguineous populations. Most populations that have practiced consanguineous marriage have done so for many generations (Woods, Cox et al. 2006).

The degree of homozygosity seen in such populations has not been estimated directly. The authors sought to address this question by examining individuals with autosomal recessive disease drawn from two populations with a long history of consanguinity (Pakistani and Arab) and a third population with a short history of consanguinity (Irish traveler, a distinct ethnic group originating from Ireland, with their own culture and language). One of the most important findings in this work was that the longest homozygous segment was the disease associated segment in 8 (17%) of 48 individuals. Since, on average, each individual bore 20 homozygous segments 13 cM, it seems reasonable to initially suggest that any 1 of these 20 segments could contain the disease-gene mutation with similar likelihood. Since the range of lengths would be greatly influenced by time (mutation and recombination) and place (chromosome and recombination), the disease-associated segment ought to have been the largest in only 5% of the time. This result may be explained by the segment being of more recent origin or by the fact that ascertainment involves greater parental consanguinity.

The clinician managing recessive disease in consanguineous families is often presented with only one or two affected children, and seeking linkage to a known locus is frequently the clinician's first step. Although homozygosity at a disease locus does not prove disease linkage in a consanguineous family, it can add support and, if no overlapping homozygous segments are found in affected individuals, can

permit exclusion. However, for a sib pair in similar circumstances, homozygosity is more significant; when there are two loci of equal frequency, homozygosity in both siblings at one of these loci, even in a highly consanguineous population, would indicate correct linkage in 95% of cases, and, even when there are eight loci of equal frequency, homozygosity in both siblings would indicate correct linkage in 81% of cases (Woods, Cox et al. 2006).

# AIMS OF THE STUDY

Congenital hyperinsulinism of infancy is a pathology characterized by a profound hypoglicaemia related to a disregulated insulin secretion.; the frequency of the disease is 1:30000-50000 live birth world –wide, but in same isolated populations with high consanguinity like the Arabic Peninsula, frequency of 1:2500 is reported. CHI has two different histological presentations, which are characterized either by a diffuse insulin secretion or by an adenomatous iperplastic focal form of the pancreatic β- cells, underling different molecular mechanism at the basis of the disease. The diffuse form shows an autosomic recessive form of inheritance and rarely an autosomic dominant one ; the focal form presents like a sporadic form (Giurgea, Bellanne-Chantelot et al. 2006) (Giurgea, Sanlaville et al. 2006) At present it is well known that disfunctions of pancreatic β- cells K-ATP channel are a common cause of CHI. The two subunits of the channel are coded by two genes, KCNJ11 and ABCC8, the last one coding for the sulphonilurea receptor, localized in the p11.5 region of chromosome 11. Mutation in these two genes inhibits the functionality of the K-ATP channel and are classified as *channelopathies*, which include mutations decreasing the channel activity till the complete lack of functionality. Mutations in ABCC8 gene are responsible of the 50-60% of focal and diffuse CHI patients, otherwise mutations in KCNJ11 are responsible for 10-15% of cases (De Leon and Stanley 2007). Focal forms of the pathology derive from a mutation in ABCC8 or KCNJ11 genes coming from the father and a contemporary loss of heterozygosity of the same allele in the mother in the pancreatic tissue bearing the focal lesion. Loss of heterozygosity of the chromosomal region 11p15.2 leads to an unbalanced expression of maternally imprinted genes like H19 and IGF2 and also p57 gene, regulating cellular growth. Diffuse form of CHI are due to the presence of two mutations in one of the two candidate genes , one paternally and one maternally inherited in a recessive way. Other less frequent molecular mechanism responsible for CHI are glutamate dehydrogenase (GDH), glucokinase (GK ) (Delonlay, Simon et al. 2007) and short- chain 3 hydroxyacyl Coenzyme A dehydrogenase (SCHAD) deficiency (Molven, Matre et al. 2004). CHI derives from a heterogeneous genetic background and many other genes involved in the pancreatic β- cells metabolism are good candidate genes for this disease.

Current genome-wide high density mapping technologies have improved the number of feasible applications of SNPs genotyping in the identification of new disease associated loci using genotype calls not only for association studies involving single polymorphisms but, thanks to the contemporary analysis of copy number, also revealing the presence of long stretches of homozygosity and their classification as structural variations (deletion, UPD, autozygosity). These studies underline how this

phenomenon is not normally due to uniparental disomy, that is a rare event associated to particular syndromes in which region of imprinted region or cancer predisposing genes are involved. This homozygosity stretches are generally located in regions of broad linkage disequilibrium and low frequency of recombination, and are considered autozygosity trait . One possible mechanism at the basis of autozygosity is the presence of a common ancestor which decreases the probability of recombination or the nature of particular chromosome trait allowing them to remain unchanged (Huqun, Izumi et al. 2007; Simon-Sanchez, Scholz et al. 2007)

This methodological approach to the analysis of SNPs genotyping data can be called "homozygosity mapping", which can be useful in the identification of disease associated locus/genes. Homozygosity mapping focuses on individuation of regions of autozygosity which are in common between affected individuals of the same family or between affected unrelated probands with the same disease which have high probability to harbor new mutations (Huqun, Izumi et al. 2007 ), and can underline regions that can lead to a new gene (or to a number of candidate loci )(Ramprasad, Soumittra et al. 2008) as responsible of the genetic basis of the disease.

In order to identify new disease susceptibility loci in a cohort of consanguineous and not consanguineous 34 CHI Italian families , we performed genome-wide homozygosity mapping using 250K NspI Gene Chip Affymetrix SNP microarrays. This approach will be useful to reveal the number and the nature of homozygosity trait in common between probands to underline new regions containing candidate genes involved in the ethiopathogenesis of CHI. Further screening for mutations in candidate genes will performed and biochemical pathways involved in the disease will be analyzed.

# MATERIAL and METHODS

## 1.Patients

This study was performed in accordance to the Italian legislation on sensible data recording. Appropriate informed consentof 110 collected sample was obtained from all CHI families members.

Inclusion criteria for CHI diagnosis are internationally accepted (Aynseley-Green et al, 2000) and patients with transient hyperinsulinism of infancy or insulinoma were excluded from the study.

All 34 CHI collected families do not have affected parents and are composed as described:

- 25 trios
- 4 families with 4 components (3 with not affected brother and one with affected sister)
- 3 families with 5 components (2 families with an unaffected brother and one family with an affected sister )
- 1 families with 7 components ( 4 not affected uncles )
- 2 single probands

We have analyzed a cohort of 35 CHI probands, 31 of Italian origin , while 4 coming respectively from Egypt ,Tunisia, Algeria and El Salvador. In five families consanguineity wasdeclared .

The vast majority of patients used diazoside to regulate glucose levels and seven patient underwent surgery , six of which subtotal and one partial pancreasectomia. Histopathological characterization was possible for six of the 7 operated patient and has evidenced the presence of 3 focal form and 3 diffuse form . Two of this patient developed diabetes after surgery.

Clinical , biochemical and therapeutic data for all 35 probands are summarized in table 1.

Table 1 . CHI probands clinical description

| Proband | Sex | Nation Of Origin | Consanguinity | Age at Diagnosiis | Blood glucose (mg/dl) | IRI mcU/ml | Glucagone response | Terapia | Chirurgy | Histology |
|---|---|---|---|---|---|---|---|---|---|---|
| HI17P | M | Italy | No | birth | | 340 | | DZ Octreotide | Pancreasectomia subtotal | diffuse |
| HI18P | F | Italy | No | 3 months | 24 | 36 | 45 mg/dl (0') 165 mg/dl | DZ | pancreasectomia subtotal | focal |
| HI15P | M | Tunisia | YES First cousin parents | birth | 91(3 years) | 40(3 years) | 39 mg/dl (0') 123 mg/dl | DZ octreotide Prednisone | pancreasectomia subtot | diffuse |
| HI20P | F | Italy | No | 2 years | 49 | 21,8 | 66 mg/dl (0') 206 mg/dl | DZ | | |
| HI21P | F | Italy | No | 4 days | 29 | 24 | | DZ Octreotide | | |
| HI16P | M | Italy | YES Second cousin parents | 1 months | 24 | 52 | | DZ | | diffuse( PET ) |
| HI22P | F | Algeria | YES First cousin parents, and grand mather and grand father too. | 9 months | 40 | 7 | 40 mg/dl (0') 58 mg/dl | DZ STOP 18 months | | |
| HI23P | F | Italy | No | birth | 5 | | 42 mg/dl (0') 124 mg/dl | DZ STOP 6 years | | |
| HI14P | M | Italy | No | 4 months | | | | Dz STOP after surgery | pancreasetomia partial | focal |
| HI01P | M | Italy | No | 13 months | | | | Dz STOP | | |
| HI02P | M | Italy | No | 3 months | | | | Dz octeotide | | |
| HI04P | F | Italy | No | 1 day | | | | Dz STOP | | |
| HI05P | F | Italy | No | 8 months | | | | Dz | | |
| HI07P | F | Italy | No | 1 year | | | | proglicem+ Valproato | | diffuse( PET ) |
| Proband | Sex | Nation Of Origin | Consanguinity | Age at Diagnosiis | Blood glucose | IRI mcU/ml | Glucagone response | Terapia | Chirurgy | Histology |

| | | | | | (mg/dl) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| HI09P | F | Italy | No | 1 week | | | | insulina+ depachin | pancreasectomia subtotal | focal |
| HI10P | M | Italy | No | | | | | | | |
| HI11P | F | El Salvador | *YES* Second cousin parents | 4 days | | | | | pancreatectomia subtotal | not determined |
| HI12P | M | Italy | No | 3 days | 11 | | | Dz+ fenobarbital | | |
| HI13P | M | Egypt | *YES* First cousin parents | 2 days | | | | Dz | | |
| HI24P | M | Italy | No | 13 months | | | | DZ | | Diffuse ( PET ) |
| HI25P | M | Italy | No | 21 days | 22 | | | Dz | | |
| HI26P | F | Italy | No | 2 days | | | | proglicem+ Sandostatin | | |
| HI27P | M | Italy | No | 8 months | 36 | | | diet | | |
| HI28P | F | Italy | No | | | | | | | |
| HI29P | M | Italy | No | 1day | 26 | | normal | Dz | | Diffuse ( PET) |
| HI30P | M | Italy | No | birth | 6 | | | Dz+ Sandostatin | pancreatectomia subtotal | focall |
| HI31P | M | Italy | No | birth | 5 | | | Dz + Luminalette | | |
| HI32P | F | Italy | No | 3 months | 48 | 55 | | Dz + Luminalette | | |
| HI33P | F | Italy | *YES* | birth | 20 | 133 | | DZ STOP | | |
| HI34P | F | Italy | No | 3 months | 39 | 10 | | DZ Octreotide | Transhepatic catheterism | diffuse |
| HI03P | M | Italy | No | 12 months | 22 | 6 | | DZ | | |
| HI35P_1 | F | Italy | No | birth | 30 | 24 | | DZOctreotide STOP | | |
| HI35P_2 | F | Italy | No | 1 month | 33 | 2 | | DZ | | |
| HI36P | M | Italy | No | | | | | | | |
| HI19P | M | Italy | No | | | | | | | |

## 2.Affymetrix 250k NspI GeneChip Snp Genotyping

### *2.1 DNA extration and purification*

DNA from peripheral blood lymphocytes was extracted using QiaAMP mini-extraction kit (Qiagen). Following DNA extraction, a salt precipitation DNA cleanup step was done as described in the Affymetrix GeneChip Mapping 500K Assay Manual.

DNA yields were quantified by spectrophotometry (NanoDrop, Nanodrop Technologies, DE, USA) and were greater than 250 ng for each sample (minimum quantity requested for Affymetrix GeneChip Human Mapping). DNA underwent electrophoresis on O.9 % agarose gel for quality control.

### *2.2 Affymetrix GeneChip Human Mapping 250k NspI Array coverage and annotation*

All 262.00 SNPs on the GeneChip Human Mapping 250K NspI Array Set were selected and tiled on arrays based on accuracy, call rate and linkage disequilibrium analysis in three populations across the genome and then went through validation process.The median physical distance between SNPs is 2.5 kb , the median physical distance between SNPs is 5.8 kb . The average heterozygosity of SNPs is 0.30 and minor allele frequency average 0.22. Eigthy-five percent of the human genome is within 10 kb of a SNP.Extensive annotation for each SNPs is provided by G-TYPE software and includes db SNP ID, nearest genes , physical map location , cytoband and alleles frequency in multiple populations.

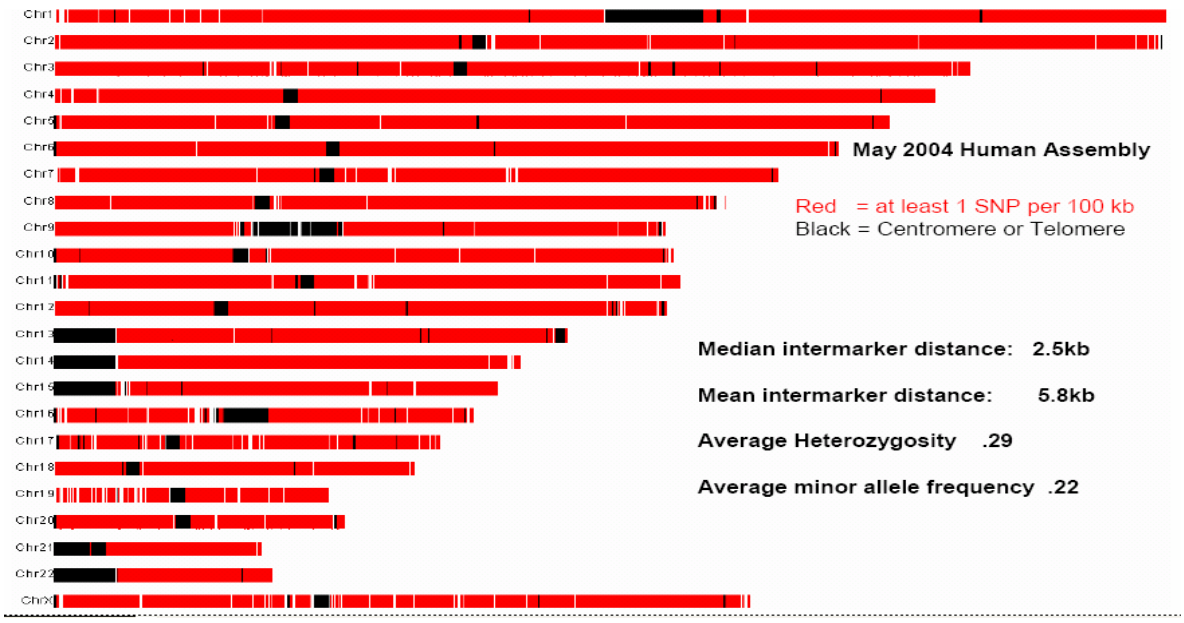



**Figure 1. Affymetrix GeneChip Human Mapping 250K NspI Array genome coverage**

***2.3 Affymetrix GeneChip Human Mapping 250k NspI Array Protocol***

DNA samples were assayed using the Mapping 250K NspI Assay kit (Affymetrix) and then analyzed on GeneChip Human Mapping 250K NspI Arrays (Affymetrix).

Briefly, 250 ng genomic DNA ( 5 µl DNA , 50 ng/ µl ) was digested with 10 units of the restriction endonuclease NspI for 2.5 h at 37°C and size of digestion products controlled by capillary electrophoresis (Agilent BIOANALYZER). Oligonucleotide adapters were ligated to the NspI digestion products using T4 DNA ligase ( 3 h at 16° C) and ligation products were added with water to 100 µl final volume. 5µl of ligation product of each sample underwent PCR amplification using Titanium Taq Polymerase in the presence of 25 µM universal primer 002 Affymetrix , 350 Mm each dNTPs and 1X Titanium Taq Polymerase buffer. Amplification protocol is described in Affymetrix GeneChip Mapping 500K Assay Manual and each sample was done in triplicate. PCR product size ( range 250-1110 bp) was controlled by capillary electrophoresis (Agilent BIOANALYZER). PCR reactions of each sample were pooled and then purified (Clontech. DNA Amplification Clean-Up Kit). Purified PCR products were quantified by spectrophotometry (NanoDrop, Nanodrop Technologies, DE, USA) an90 µg of PCR purified product was fragmented using 0.25 units of DNAase at 37° C for 35 minuts.

Fragmentations products were controlled by capillary electrophoresis (Agilent BIOANALYZER),

and their length was shorther then 180 bp. After fragmentation DNA was end-labelled with 105 units of deossinucleotide transferase at 37°C for 2 h and hybridized to an array for 18 h at 49°C.

After hybridization, the arrays were washed and stained using an Affymetrix Fluidics Station 450.

Images were obtained using a Gene Array Scanner 2500 (Affymetrix).

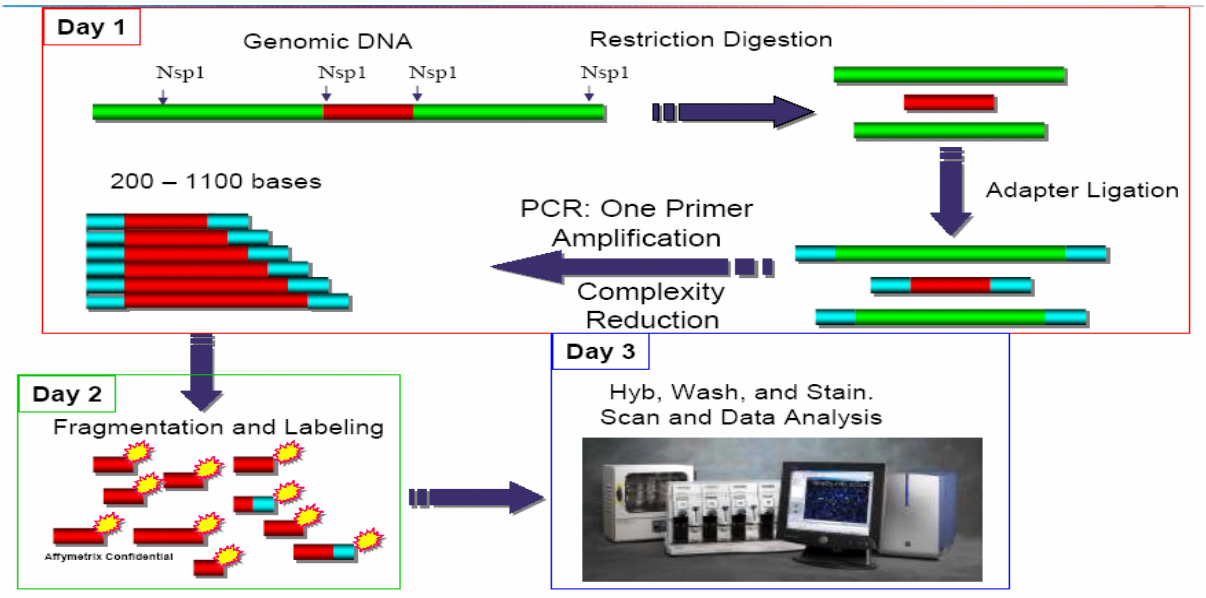



**Figure 2. Schematic rapresentation of Affymetrix GeneChip Human Mapping 250K NspI Arrays protocol.**

### *2.4 Genotypiq quality control*

Genotype calls were produced using BRLMM algorithm of GeneChip Analysis Software (version 4.0, Affymetrix) using default parameter settings:

*score threshold 0.5*

*block size 0*

*prior size 1000*

Genotyping quality control included genotype call rates (expressed as the percentage of SNPs to which BRLMM algorithm could assign an allelic call among 262.000 SNPs investigated by the array), heterozyguous call rate , mendelian error control , sex and pedigree check.

Heterozyguous call rate gives a good indications of quality and integrity of DNA samples

Samples were excluded if they showed call rates lower then 93% , mendelian errors higher then 1% (probably due to genotyping errors) and sex not corresponded to the declared.

Pedigree check in all families members was undertaken to underline transmission errors inside all 34 families due to different or not declared parental relationship.

## 3. Homozygosity mapping

The computational strategy entails several computational steps including: SNP preprocessing analysis, within-subject and across-subjects ROHs identification analysis, ROHs genes annotation. At each step, the R language was used to code the procedure (http://www.r-project.org/) and the genomic visualization of resulting ROHs was carried out using dChip (http://biosun1.harvard.edu/complab/dchip/).

**STEP1: SNP preprocessing analysis**

**a. Loss of Heterozygosity (LOH) and Copy Number (CN) analyses** were performed to determine the CN and LOH profiling for each sample by dChip2007 software. We performed unpaired LOH analysis using HMM considering haplotype correction (HC/LD-HMM) by comparing our patients dataset to CEPH 60 HapMap parents as reference samples, as suggested by the dChip manual (http://biosun1.harvard.edu/complab/dchip/manual.htm). The LOH analysis assigned to each SNP an inferred SNP LOH probability defined as P(LOH) and ranging from 0 to 1. CN analysis was performed using median smoothing and trimmed analysis, according to the software. In parallel, all samples were also analyzed with CNAG2.0 (www.genome.umin.jp) by using normal references used for dChip analysis. A matrix of SNP LOH probability data was exported from dChip2007 structured with samples in columns and SNP in rows **.**

**b. Annotation of SNP LOH probability data matrix.** Inferred SNP LOH probability data matrix were combined with additional genomic information (e.g. physical position, chromosome, cytoband, allele and genotype frequencies) included in mapping array specific annotation file. Then, annotated data matrix was sorted according the SNP physical position along the genome and used to ROH identification analysis at the level of single sample or across samples.

**STEP2: Within-subject ROHs analysis (single sample data)**We used single sample data and extracted stretches of consecutive events of LOH within-subject setting the P(LOH)>=0.5 defined as runs of homozygosity (ROHs). The procedure produced a table listing the ROHs of each patient (sample). Each ROH was described by start and end genomic SNP position (bp) and cytoband, length (bp), number of SNPs, the maximum and minimum value of P(LOH) associated to SNP belonging to ROH .

**STEP3: Across-subjects ROHs analysis (multiple sample data)a. Fingerprint Analysis.** We identified the ROH pattern across the entire dataset by setting the P(LOH)>=0.5, defined as "fingerprint analysis". The fingerprint was determined by the presence of same ROHs clusters among at least two samples in dataset.

The fingerprint ROHs data were described by start and end genomic SNP position (bp) and cytoband, length of ROH, number of SNPs in the region, the maximum and minimum number of patients (frequency) sharing the ROH region.**b. Shared ROHs Analysis.** Fixing the number of individuals carrying same ROHs namely "shared ROHs"**,** we then extracted common regions of ROHs across all the subjects under study. Data tables containing a list of shared ROHs calculated using 2, 3 or more samples can be produced. Usually, this step reduces the length of shared ROHs)and permits an enhancement of genomic regions clinically important for the disease under study .

**STEP4: Annotation of genes to ROHs**

The annotation step allowed the association of genes to selected ROHs. The list of genes associated to ROHs was obtained using the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables) and by querying the hg17 database, the following tables: KnowGenes, RefGene and RefFlat. Genes list was used to carry out enrichment and pathways analysis or to investigate for presence of new diseases candidate gene

## 4. CNAG(v3.0) copy number and LOH analisys

Homozygosity mapping of genotype calls of the CHI families members was performed using CNAG (v3.0) software for both copy number and loss of heterozygosity analysis with default parameter settings, comparing 48 Hap Map CEU samples genotype calls as reference and considering homozygosity trait including at least 10 contiguous SNPs and LOH likehood greater than 30. Homozygosity traits evidenced in the child were further analyzed for copy number changes in order to exclude the nature of deletion and compared with parents genotypes to exclude uniparental disomy.

CNAG ( v3.0 ) use allele-specific copy-number analysis using anonymous reference algorithm (AsCNAR)AsCNAR is a new algorithm developed for detection of allelic composition in the total copy number status, without dependence on the availability of constitutive DNA. By sensing subtle distortions in allele-specific signals caused by allelic imbalance with the use of anonymous controls (a pool of normal unrelated samples), it also enables sensitive detection of LOH with accurate determination of allele-specific copy numbers.

This algorithm allows simultaneously analysis of LOH and copy number changes in paired and un-paired samples comparing allele-specific signals with those in anonymous referencesAs reported Yamamoto et al., the use of multiple

references overcomes the low concordance rate with a single reference (37%), and the expected overall concordance rate for heterozygous SNPs and for all SNPs increases to 86% and 92%, respectively, with five unrelated references (Yamamoto2007). Thus, for AsCNAR, allele-specific signal ratios are calculated at all the concordant heterozygous SNP loci for individual references, and then the signal ratios for the identical SNPs are averaged across different references over the entire genome. For the analysis of total CNs, all the concordant SNPs, both homozygous and heterozygous, are included in the calculations, and the two allele-specific signal ratios for heterozygous SNP loci are summed together. His peculiarity is that it is capable to display in a single chart LOH regions, CN values and allele dosage, in particulary can determine regions of LOH with CN neutral (CNN), that is one of the two alleles is lost and the other duplicates itself, or which of the two alleles has been lost (Loss) or duplicate (Gain).

## 5. HADH gene sequencing

Genomic DNA was isolated from peripheral blood cells using a commercial kit (INVITROGEN PureLink™Genomic DNA Mini Kit).

The complete coding region of HADH isoform 1 (Ensembl Transcript ID ENST00000309522) and HADH isoform2 (Ensembl Transcript ID ENST00000351726 ) including intron-exon boundaries, was amplified by polymerase chain reaction (PCR) using appropriate primer pairs (see table below).

PCRs were performed using hotstart polymerase (Amplitaq Gold, Applied Biosystems, Foster City, CA, USA). Sequence analysis was performed on PCR products with an ABI sequencing kit (Applied Biosystem) with the original PCR primers.

Sequence reaction products were analyzed on an ABI prism 3700 Genetic Analyzer. All fragments were sequenced bidirectionally. The sequence of mutated fragment was repeated at least twice from two different DNA extractions. cDNA and protein sequences were numbered according to NCBI ORF finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi); with nucleotide numbering beginning with the first Met. Oligonucleotide primers used for PCR amplification and sequencing are listed in the table below .

| HADH gene primer name | sequence |
|---|---|
| | |
| SCHAD 1F | TCATAGAACAAGGGGCCAGT |
| SCHAD 1R | GTGAAAACTCCCTGGTGTCG |
| | |
| SCHAD 2F | TCGGTATTTTGAATGGAATGG |
| SCHAD 2R | GCCCTGCAAATGTAAAGCAG |
| | |
| SCHAD 3F | AGGCCAATACAGGTGCTCTG |
| SCHAD 3R | TAGGCATGGAACAAAACTGC |
| | |
| SCHAD 4F | AGCAGTCACCTTGTGCATTG |
| SCHAD 4R | AGGCTCAGGGACAACCTTCT |
| | |
| SCHAD 5F | TGCCTATATGGAGCCTTTTGA |
| SCHAD 5R | GGCACGGATTGCTCTTTACA |
| | |
| SCHAD 6F | TAAGAAGGGGCAATTTGGTG |
| SCHAD 6R | GCAGCTGAGCTCACAAAGC |
| | |
| SCHAD 7F | TCTAGATAGCCACCTGTGGCAG |
| SCHAD 7R | AGGAGTATTCCCCTGGCATC |
| | |
| SCHAD 8F | CCCAAGCCAGAAAGTCTCAG |
| SCHAD 8R | GGCAGGATTTTGGAAGGTCT |
| | |
| SCHAD 9.1F | TCTCTTCCCTCCGCTCAATA |
| SCHAD 9.1R | TGTGAGGGAATGTCTGACCA |
| | |
| SCHAD 9.2F | TCTCCGGCTCTGAGAAGAAC |
| SCHAD 9.2R | CCCGTTGATGAGAGAAGGAA |
| | |
| SCHAD 9.3F | TTTACACCCTTGGTGCCTTG |
| SCHAD 9.3R | GCAAACCATGCAAACAGAGA |
| | |
| SCHAD 9.4F | TTCAGCTCGTCGTGAAGATG |
| SCHAD 9.4R | AGGCAGGTTTTACACAAATGC |

## 6. ABCC8 and KCNJ11 genes sequencing

Genomic DNA was isolated from peripheral blood cells using a commercial kit (INVITROGEN PureLink™Genomic DNA Mini Kit).

Oligonucleotide primers used for PCR amplification and sequencing of the unique KCNJ11 exon and of the 37 ABCC8 exons are listed in the table below .

PCRs were performed using hotstart polymerase (Amplitaq Gold, Applied Biosystems, Foster City, CA, USA)

The amplified products were purified using EXOSAP® treatment (Amersham Biosciences, Piscataway, USA) and then directly sequenced with the forward and reverse primers previously used for the amplification. Sequenced products were purified using Montage SEQ 96 Sequencing Reaction Cleanup kit (Millipore) and fractionated by capillary electrophoresis on an automated DNA sequencer (MegaBACE 1000 DNA Analysis System; Amersham Biosciences, Piscataway, USA). The sequence of mutated fragment was repeated at least twice from two different DNA extractions.

cDNA and protein sequences were numbered according to NCBI ORF finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi; GenBank NM_000525.3) with nucleotide numbering beginning with the first Met. Oligonucleotide primers used for PCR amplification and sequencing are listed in the table below .

| KCNJ11 | Primer pairs | n° basi | PCR fragment size (bp*) |
|---|---|---|---|
| | | | |
| *esone 1A* | 5'-agaggactctgcagtgaggc-3' | 20 | 541 |
| | 5'-ttctgcacgatgaggatcag-3' | 20 | |
| | | | |
| *esone 1B* | 5'-cttcctgtgcagctggctg-3' | 19 | 769 |
| | 5'-gccaaacttggagtagtcca-3' | 20 | |
| | | | |
| *esone 1C* | 5'-tcgagatcatcgtcatcctg-3' | 20 | 471 |
| | 5'-agtgccttgtaacaccctgg-3' | 20 | |

| ABCC8 | Primer pairs | n° basi | PCR fragment size (bp*) |
|---|---|---|---|
| *esone 2* | 5'-tcctcatgtgtaggtgtcac-3' | 20 | 390 |
| | 5'-gagctgctggatgatgtaac-3' | 20 | |
| *esone 3* | 5'-ggttcagtgttcagcctttga-3' | 21 | 596 |
| | 5'-atccctatcagaggactttgc-3' | 21 | |
| *esone 4* | 5'-acacgtgcacatccacttac-3' | 20 | 591 |
| | 5'-cttatcttccctttctacgga-3' | 21 | |
| *esone 6* | 5'-agacaacaggagctagggg-3' | 19 | 529 |
| | 5'-ccgggagaggtactatggta-3' | 20 | |
| *esone 8* | 5'-gcacactgtaacaggtggca-3' | 20 | 419 |
| | 5'-tgcttgcagagtatgggaca-3' | 20 | |
| *ex** 10+intr*** 10* | 5'-ctcctgtccagtacttcgtg-3' | 20 | 625 |
| | 5'-atctcaaggcctcctgcttc-3' | 20 | |
| *intr 11+ex 12* | 5'-gtgtctccaactaaaagatcc-3' | 21 | 440 |
| | 5'-atgtccctctgaccaaccag-3' | 20 | |
| *esone 13* | 5'-gcctgagatcctgctcaga-3' | 19 | 388 |
| | 5'-gcttgcggttcacaaccct-3' | 19 | |
| *intr 15+ex 16* | 5'-ggaagaaagaaaacaggggc-3' | 20 | 402 |
| | 5'-gcgcatagtaagaggtgaata-3' | 21 | |
| *intr 18+ex 19* | 5'-caccagaggccacactaag-3' | 19 | 597 |
| | 5'-cgatgctagccctaactatta-3' | 21 | |
| *esone 24* | 5'-atgaaggggaaggactaaatc-3' | 21 | 382 |
| | 5'-tcaaaggcacactactgcac-3' | 20 | |
| *esone 28* | 5'-gaatcatttgaacccaagggg-3' | 21 | 523 |
| | 5'-gcggtggaataagatgtgga-3' | 20 | |
| *esone 29* | 5'-aggacagcagtaattgtgcc-3' | 20 | 499 |
| | 5'-ctgctggtggatatcccttg-3' | 20 | |
| *ex-intr 32+ex 33* | 5'-tccagccttaagaagagcca-3' | 20 | 569 |
| | 5'-ggtgactgcgaagccatcc-3' | 19 | |
| *ex-intr 34+ex 35* | 5'-ataccctgtgacctcccaca-3' | 20 | 896 |
| | 5'-ccaaccaacccagctgcat-3' | 19 | |
| *esone 37* | 5'-cattgttgtgtcaggtgcac-3' | 20 | 880 |
| | 5'-cccttaccgcgatggtgac-3' | 19 | |

## 7.Pax gene Blood RNA System

Pax gene Blood RNA System consists of a blood collection tube (PAX gene Blood RNA Tube) and nucleic acid purification kit (PAX gene Blood RNA Kit). The proprietary reagent composition of PAX gene Blood RNA Tube protect RNA molecules from degradation by RNAase and minimizes ex vivo change sin gene expression. Probands and controls 2.5 ml whole blood collection followed the PAX gene Blood RNA Tube instruction manual.

## 8. RNA extration protocol

Nucleic acid purification begins with a centrifugation step to pellet nucleic acid in the PAX gene Blood RNA Tube. The pellet is washed and resuspended , an incubated in optimazed buffers together with proteinase k to bring about protein digestion. An additional centrifugation through the PAX gene Shedder spin is carried out to homogenize cell lysate and remove cell debris , and the supernatant is transferred to a fresh microcentrifuge tube.

Ethanol is added to adjast binding conditions, and the lysate is applied to a PAX gene spin column. RNA is selective bound to the Pax gene silica-gel membrane and contaminants are removed in wash steps. The membrane is treated with DNAase I to remove trace amount of bound DNA. After the wash step, RNA is eluted in steril water.

## 9.cDNA Reverse Transcription Kit

Retrotranscription of 1 γ of total RNA was undertaken using High capacity cDNA Reverse Transcription Kit ( Applied Biosystems ) .

10µl of RNA sample is added to 10µl of RT master mix composed by:

- 2 µl 10X RT buffer
- 0.8 µl dNTP MIX 100 Mm
- 2 µl 10X RT random primers
- 1µl MultiScribe Riverse Trancriptase
- 1µl RNAase Inibitor
- 3,2µl H20

Thermal cycler conditions are described in the kit Information Document.

## 10. HADH gene cDNA amplification

Amplification of HADH gene cDNA was done using a couple of primers amplifying exonic sequence between exon 3 and exon ( primers sequence in table ). PCRs were performed using hot start polymerase (Amplitaq Gold, Applied Biosystems, Foster City, CA, USA) in standard condition and with an annealing temperature of 57 ° C.The amplification product has an expected length of 733 bp if Isoform 2 of the transcript is amplified; otherwise the amplification product would be of 682 bp if the first and more abundant isoform is amplified. This Isoform 1 presents skipping of exon 7.PCR fragment were controlled on 2% agarose gel.

As amplification controls commercial cDNAs were used :

liver (Clonthec), Scheletric Muscle (Clonthec), Human Control 10 (Clonthec)

| | |
|---|---|
| PTT SCHAD R | TCACTTGTATTTGTAAAATCCTTCTC |
| cDNA SCHAD ex3 F | CGGCGATGAATTTGTGGAG |

## 11. Bioinformatic Analysis of SHAD protein

In order to evidence sequence differences and homologies , we used Clustal 2.0.5 multiple sequence alignment of wild type and truncated SCHAD protein. Wild type and mutated SHAD protein ribbon diagramm was obtained with ESyPred3D (Prediction of proteins 3D structures), an automated homology modeling program. Conserved Domains were also obtained from Pfam database (a large collection of protein families, each represented by *multiple sequence alignments* and *hidden Markov models (HMMs)* of wild type and truncated SCHAD protein sequences .

## 12. MCAD gene mutation screening

Genomic DNA was isolated from peripheral blood cells using a commercial kit (INVITROGEN PureLink™Genomic DNA Mini Kit).

Oligonucleotide primers used for PCR amplification and sequencing of the unique.MCAD 13 exons are listed in the table below .

| MCAD gene primer name | sequence |
|---|---|
| ACADM 1F | GACCAGAGGAGTCCCGCGTT |
| ACADM 1R | ACCGCTGGGCTCCCTCTCA |
| ACADM 2F | TAGTCTCTTATCTGATTAATGTTTAAC |
| ACADM 2R | GCTTCATATGTATAAGTTTAAAGTC |
| ACADM 3F | TTTAACTTTTCTAAATAATTTTCCC |
| ACADM 3R | TTTCCCTCTTTAAAATGTATATACC |
| ACADM 4F | ATATATTTTAACTCAGTTCTTTTTC |
| ACADM 4R | CAGTATTAAAAATATAAAGAAAGC |
| ACADM 5F | TTTCTTGCCATTAAATGACTTTC |
| ACADM 5R | TGGGTCTGAGGAAAAGAATCA |
| ACADM 6F | ATGTGTTGAAACATTTTGATACT |
| ACADM 6R | AGTTAGGTAGTTAATTTTCTAATC |
| ACADM 7F | GATTTTTTAATGTCAATTTTCTT |
| ACADM 7R | TAAAGCGGCAGTTACACATA |
| ACADM 8F | CATTTAATTTCATTTCTCTTGTTTTTA |
| ACADM 8R | AGAAAAAATATACAAAGATGTTTTGA |
| ACADM 9F | CACCATTATCCGGTATGTGTA |
| ACADM 9R | TGGGGCCCTAATCATTAATAAA |
| ACADM 10F | GCTGGCTCAAAAAAAATTCC |
| ACADM 10R | AAGATCCTATAAAGATTATGAACCC |
| ACADM 11F | ACATGAACTTTTGCTTTATAATATCT |
| ACADM 11R | AATGAGTGTCTTTACATTTGAAC |
| ACADM 12AF | ACTTTTAAGTTTTCTCAATAAATATCCTTTAAT |
| ACADM 12AR | TATCTCCAGCAAATGCCTTT |
| ACADM 12BF | GTCGAAATACCTATTATGCTTCT |
| ACADM 12BR | ATATTCTCTCTCCTTGCAAAC |
| ACADM 13F | AAAGATATTTAACCTACACTTATATTTTTC |
| ACADM 13R | ACAGTGGCTTGTGTTCT |

## 13. Bioinformatic pathways analysis

Bioinformatic pathways analysis was undertaken uploading in David Functional Annotation Chart the list of genes in common between at least two patients.

# RESULTS

## 1.Homozygosity traits analysis of 34 CHI families with CNAG (v3.0) software

In order to realize a first description of homozygosity traits present in 34 families with CHI affected probands we used SNPs genotyping microarray platform GeneChip Human Mapping 250K NspI Array and CNAG (v3.0) software, and we obtain a list of homozygosity traits found in all families' members analyzed , their chromosomal position, starting and end physical position, start and end cytoband and length in Mb.

The analysis of homozygosity events inside every single family has evidenced presence of homozygous traits in more than one member (Table 1), included parents and healthy parents and sisters. Six of the 34 unaffected mothers presented more then one homozygosity trait, in some cases revealing 6, 7, 9, and even 14 homozygous regions; in this last highly homozygous sample the longest trait was 95.13 Mb on chromosome 3 and this was the most extended one in all our CHI dataset (data not shown). Considering the other families members, we also found that 9 of the 34 unaffected fathers showed high homozygosity degree; two of these samples bearing 8 and 12 homozygous stretches and the longest trait in a father' sample was 30.83 Mb in length.

We have also analyzed affected and unaffected sibs and we evidenced one unaffected brother with 7 homozygosity trait and two unaffected sibs of the same family with 2 and 3 stretches , the longest one of 76.61 Mb in length on chromosome 8 (Table1).

Homozygosity trait analysis in our 35 CHI probands evidenced that homozygosity traits of more than 1 Mb were present in only 10 probands and they spanned between 3.4 and 79 Mb in length (Table1).The evidenced homozygosity regions were located on 16 different chromosomes, and they were not preferentially distributed on one or few chromosomes. When we considered the chromosomal position of every single homozygosity stretch we elucidated that their location along the chromosomes was peculiar to every single CHI patient and none of the affected children presented perfect match between start and end position of the loss of etherozygosity regions. Nevertheless, we were able with this first analysis to underline that in 6 of the 16 chromosomes that presented homozygosity stretches there were same regions in common between CHI patients (Table1). On Chromosome 4 two probands shared a common region between 102152133 and 106370477, on chromosome 6 between 19936192 and 33975418, on chromosome 10 between 126834837 and 127762981, and on chromosome 13 between 81164254 and 84504847 (Table1). On chromosome 11 we evidenced three CHI probands with a common LOH trait that covers the 11p15.2 region and on chromosome 18 one CHI proband shared a common region with two other probands, but the two overlapping traits

had a different chromosomal position (Table 1).

Further investigation of the parental origin of homozygous stretches has confirmed that they were not inherited and the contemporary analysis of the copy number with CNAG (v3.0) software did not reveled any copy number loss, concluding that they were not deletions. Mendelian errors check (all below 0.1%, threshold due to genotyping errors and not transmission errors) and parental haplotypes analysis has excluded uniparental disomy (UPD) and they must be considered as autozygosity traits.

| Sample_name | Chromosome | StartPos | EndPos | Mb | START CYTOBAND | END CYTOBAND |
|---|---|---|---|---|---|---|
| HI22P | 1 | 54594479 | 80226727 | 25,63 | 1p32.3 | 1p31.1 |
| HI22P | 1 | 155917158 | 160086719 | 4,17 | 1q23.1 | 1q23.3 |
| HI13P | 1 | 185811639 | 189786806 | 3,98 | 1p36.21 | 1q44 |
| HI16P | 2 | 230112156 | 242642787 | 12,53 | 2q37.1 | 2q37.3 |
| HI06P | 4 | 102152133 | 106370477 | 4,22 | 4q21-q24 | 4q24 |
| HI16P | 4 | 68321842 | 147336231 | 79,01 | 4q13.2 | 4q31.21 |
| HI11P | 5 | 116226943 | 120983392 | 4,76 | 5q23.1 | 5q23.1 |
| HI11P | 5 | 160387464 | 167444580 | 7,06 | 5q34 | 5q34-q35.1 |
| HI13P | 5 | 146708683 | 159601881 | 12,89 | 5q33.1 | 5q33.3 |
| HI13P | 5 | 125396244 | 141943941 | 16,55 | 5q31.1 | 5q32-q33 |
| HI28P | 5 | 38435479 | 50536897 | 13,61 | 5p13.2 | 5q11.2 |
| HI16P | 6 | 6076323 | 10601757 | 4,53 | 6p25.3-p24 | 6p24.3 |
| HI22P | 6 | 11311011 | 33975418 | 22,66 | 6p24.1 | 6pter-p21.31 |
| HI16P | 6 | 19936192 | 41091259 | 21,16 | 6p22.3-p22.2 | 6p21 |
| HI16P | 6 | 88778372 | 96025216 | 7,25 | 6p25.3 | 6q27 |
| HI18P | 6 | 102233451 | 109256119 | 7,02 | 6q16.3-q21 | 6q21 |
| HI11P | 7 | 13550657 | 16986312 | 3,44 | 7p21.2 | 7p21.1 |
| HI16P | 7 | 132026854 | 146183963 | 14,16 | 7q32.3 | 7q35 |
| HI13P | 8 | 17627431 | 38405382 | 20,78 | 8p22 | 8p11.2 |
| HI11P | 8 | 40509288 | 65657887 | 25,15 | 8p11 | 8q12.3 |
| HI13P | 8 | 97623577 | 107482718 | 9,86 | 8q22.1 | 8q23.1 |
| HI22P | 8 | 123099200 | 134674046 | 11,57 | 8q24.13 | 8q24.22 |
| HI13P | 10 | 121898842 | 127762981 | 5,86 | 10q26.12 | 10q26.2 |
| HI22P | 10 | 126834837 | 133718894 | 6,88 | 10q26.13 | 10q26.3 |
| HI15P | 11 | 8555549 | 24098510 | 15,54 | 11p15 | 11p14.2 |
| HI11P | 11 | 11995777 | 18704815 | 6,71 | 11p15.3 | 11p15.1 |
| HI09P | 11 | 12394102 | 17467332 | 5,07 | 11p15.3 | 11p15.1 |
| HI13P | 11 | 114173286 | 11998603 | 5,81 | 11q23.2 | 11q23.3 |
| HI13P | 13 | 69627005 | 93563813 | 23,94 | 13q22.1 | 13q32 |
| HI28P | 13 | 81164254 | 84504847 | 13,55 | 13q31.1 | 13q31.1 |
| HI13P | 14 | 29766259 | 84248261 | 54,48 | 14q12 | 14q24.3-q31 |
| HI22P | 15 | 83921005 | 91352551 | 7,43 | 15q24-q25 | 15q26.1 |
| HI13P | 16 | 60780281 | 69551276 | 8,77 | 16q22.1 | 16q22.1 |
| HI13P | 17 | 7804069 | 16391547 | 8,59 | 17p13.1 | 17p11.2 |
| HI15P | 18 | 13331062 | 22131173 | 8,80 | 18p11.2 | 18q11.2 |
| HI13P | 18 | 14842335 | 40392980 | 25,55 | 18p11.21 | 18q12.3 |
| HI16P | 18 | 22609157 | 44240170 | 21,63 | 18q11.2 | 18q21.1 |
| HI11P_ | 22 | 41674130 | 47904529 | 6,23 | 22q13.2-13.33 | 22q13.31 |

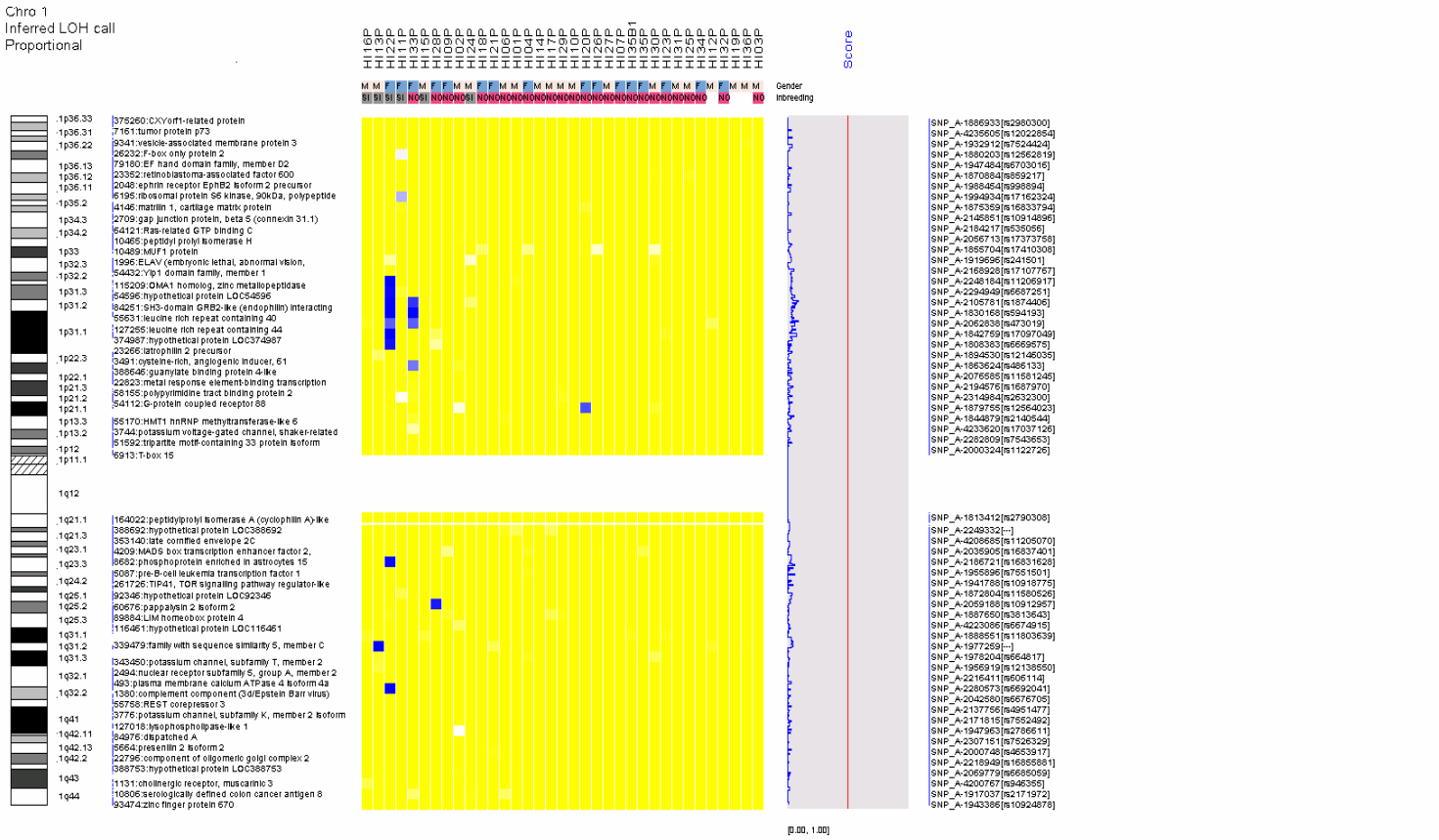
**Table1.Homozygosity trait analysis with CNAG(v3.0) software** . List of homozygosity trait in 35 CHI probands with indication of chromosomal position , start and end physical position , start and end cytoband and length in Mb.

# 2.Homozygosity traits analysis of 35 CHI probands with dChip software

To further characterize homozygosity traits in 35 CHI probands, we analyzed genotyping microarray GeneChip Human Mapping 250K NspI Array data using dCHIP software . In this analysis we decided to consider as a run of homozygosity (ROH) an homozygous trait with at least 50 SNPs, with the minimal length of 1 Mb and a loss of heterozygosity (LOH) probability greater or equal to 0.5. In order to compare homzygosity profile of all 35 probands, we also used dChip software to visualize LOH analysis chromosome by chromosome. Using parameter setting previously decribed, we have produced a list of ROHs present in the 35 CHI analyzed probands , each ROH is described by start and end physical position , start and end cytoband, number of SNPs included in that region, maximum and minimal value of LOH probability in that trait (data not shown), and ROH length in base pair. We underlined that 22 patients presented one or more ROH and the maximum number of ROHs owing to a single proband was 31 distributed on 15 chromosomes, while the second proband presented 30 ROHs on 12 different chromosomes.
When we also considered the chromosomal distribution of the previously evidenced ROHs, we could not underline a physical preferential positioning, even if chromosome 1, 2, 4, 5, 6 and 10 showed at least 10 homozygous regions. Chromosome 1 LOH analysis evidenced the presence of homozygosity traits in 5 CHI probands (Table2); the long contiguous trait was of 25 Mb on 1p31.1, with a small overlapping homozygous segment in a second CHI patient; chromosome 2 analysis (Table3) listed 5 homozygous patient too, one patient showing a telomeric region of LOH of approximately 1Mb. On chromosome 4 (Table 4) one probands showed the longest homozygosity stretch of our CHI patient dataset ( from 4q13.2 to 4q31.21) which ranged 79,01 Mb and included 312 known genes among which the known CHI causative HADH gene. This gene is genotyped on the microarray by 7 SNPs covering 45 kb and transmission analysis of these SNPs parents' haplotypes excluded the presence of uniparental disomy. These two long and contiguous stretches were previously identified with CNAG3 software too.
Considering other homozygous chromosomal region, an high degree of homozygosity was also revealed in one patient on chromosome 5 (Table5), which includes two LOH traits, the first one spanning between 5q23.2 and 5q33.3 and the second one between 5q35.2 and 5q35.3. On chromosome 6 (Table6) we evidenced two probands with two extended and partially overlapping homozygous region,with extension of respectively 12063077-33775934 and 20121370-41617879; one proband bore an LOH trait of 13.5 Mb on chromosome 7.
Chromosome 8 (Table7) revealed 4 long and contiguous LOH regions in 3 patients , one presenting two LOH traits on two different chromosomal arms and another one bearing a long stretch starting from p arm (8p11.21) and ending in 8q11.23 cytoband. All these findings were confirmed by CNAG software LOH analysis. When we further consider the most extend LOH traits of chromosome 10 (Table8) and 11 (Table9), we confirmed that two probands presented a partially common region of homozygosity spanning 10q26.12-10q26.2 and 10q26.13-10q26.3 while we underlined 6

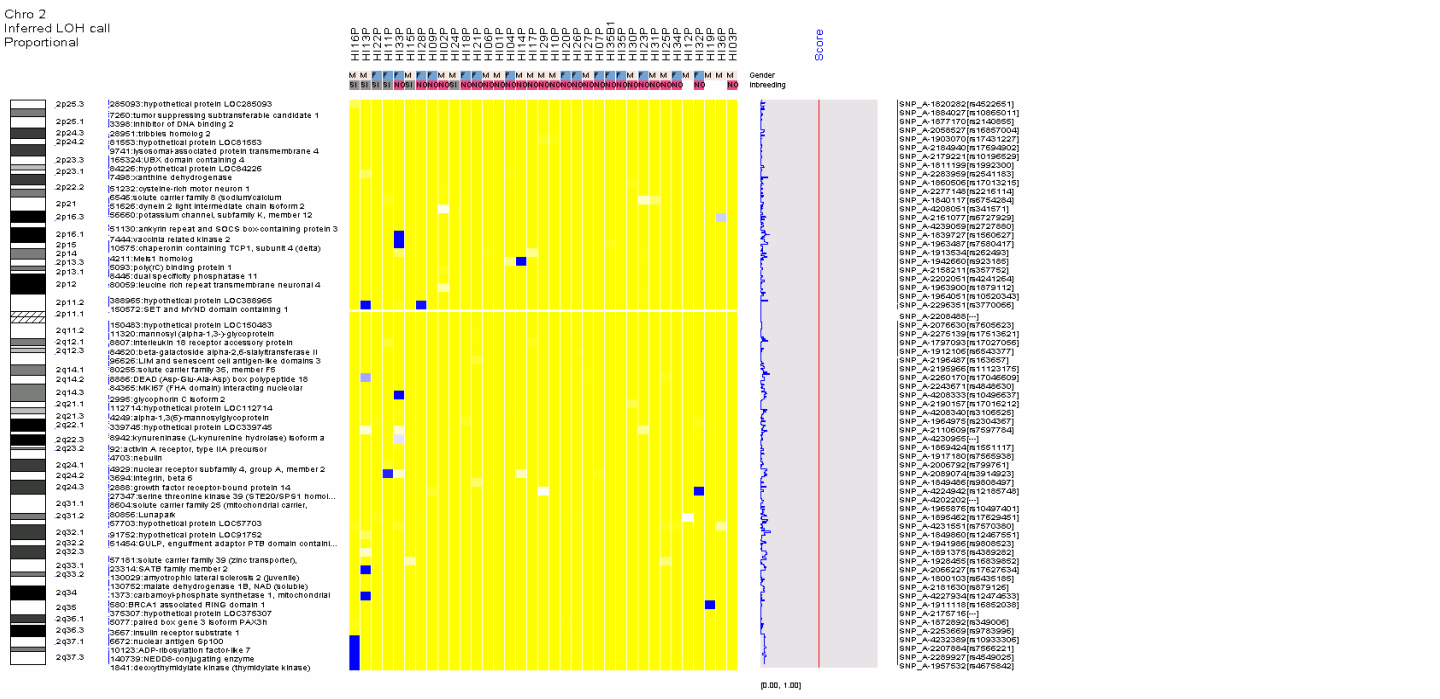
patients with homozygous traits on chromosome 11, evidencing a forth proband owning LOH trait that covers the 11p15.2 region not previously evidenced with CNAG software LOH analysis.

Otherh very long and contiguous regions of homozigosity were found in one patient on chromosomes 13, 14 (data not shown ) and in two patients on chromosomes 18 (Table10) of respectively 24, 54.5, 25.5 and 21.6 Mbase. Chromosomal positioning of ROHs of this entity had a nearly perfect correlation with homozygosity analysis previously done with CNAG software, even if dChip inferred LOH was seen to cluster them in more then one ROH (three different ROHs in the first two chromosomes cited ), with few SNPs dividing each other, particularly when their LOH probability dropped below 0.6 ( data not shown )..



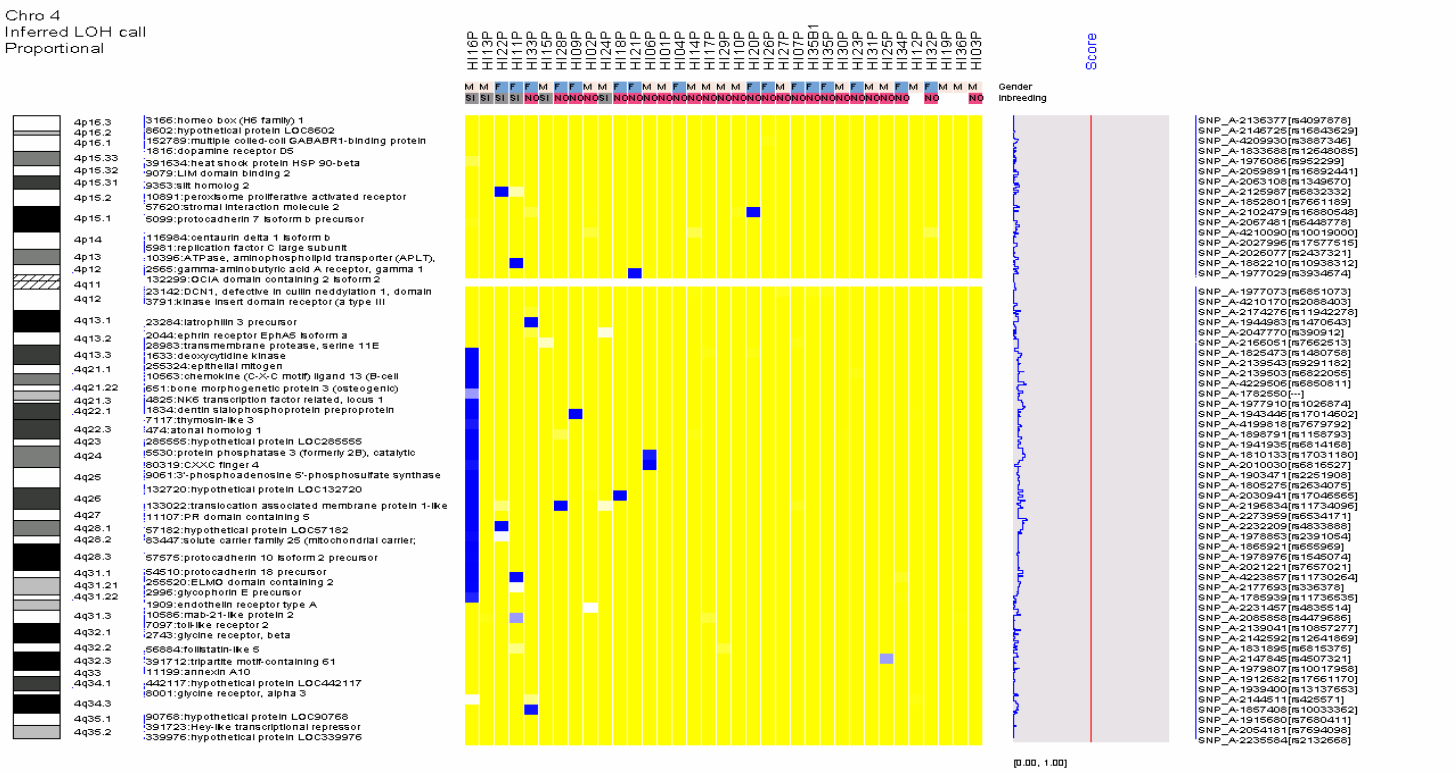

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 1 | 170139711 | 172275804 | q25.1 | q25.1 | 150 | 2136094 | ArrRaP_HI28P |
| 1 | 54671315 | 71125345 | p32.3 | p31.1 | 1950 | 16454031 | Ben.probando |
| 1 | 71714715 | 72767289 | p31.1 | p31.1 | 100 | 1052575 | Ben.probando |
| 1 | 73701875 | 77106323 | p31.1 | p31.1 | 400 | 3404449 | Ben.probando |
| 1 | 77554882 | 80127921 | p31.1 | p31.1 | 250 | 2573040 | Ben.probando |
| 1 | 156244808 | 159015715 | q23.2 | q23.3 | 250 | 2770908 | Ben.probando |
| 1 | 186337133 | 189476447 | q31.1 | q31.2 | 400 | 3139315 | LalIP_HI13P |
| 1 | 10266368 | 12273885 | p36.22 | p36.22 | 100 | 2007518 | SaKaP_HI11P |
| 1 | 25748046 | 26996958 | p36.11 | p36.11 | 50 | 1248913 | SaKaP_HI11P |
| 1 | 71125403 | 72767289 | p31.1 | p31.1 | 150 | 1641887 | SpiAnP_HI33P |
| 1 | 73701875 | 74847962 | p31.1 | p31.1 | 100 | 1146088 | SpiAnP_HI33P |
| 1 | 92543590 | 93987929 | p22.1 | p22.1 | 100 | 1444340 | SpiAnP_HI33P |
| 1 | 106760930 | 108031431 | p21.1 | p13.3 | 150 | 1270502 | SpiAnP_HI33P |

**Table2.CHROMOSOME 1 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).



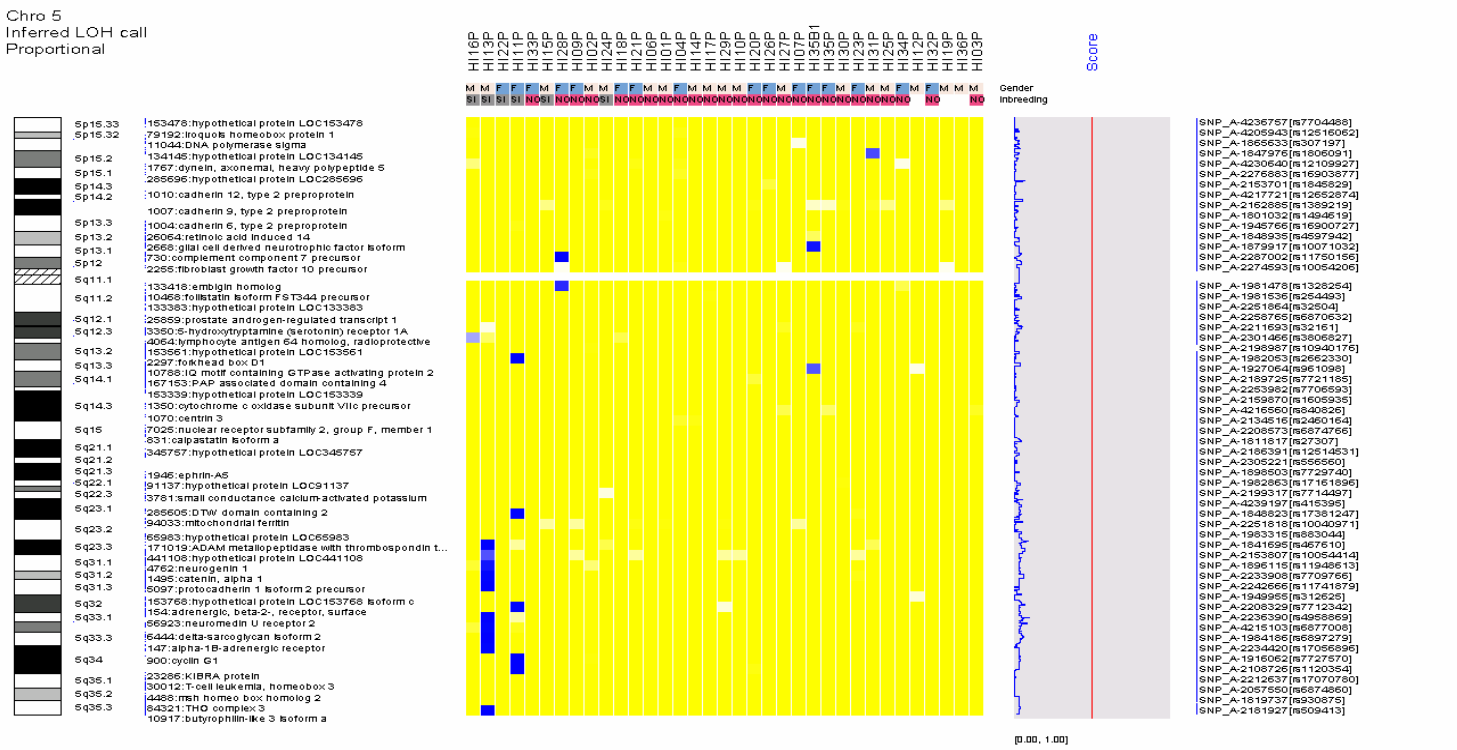

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 2 | 185273737 | 187123635 | q32.1 | q32.1 | 150 | 1849899 | AgnAleP_HI35P |
| 2 | 199618880 | 202225833 | q33.1 | q33.1 | 250 | 2606954 | LalIP_HI13P |
| 2 | 208868359 | 211687603 | q33.3 | q34 | 250 | 2819245 | LalIP_HI13P |
| 2 | 12779571 | 13860317 | p24.3 | p24.3 | 100 | 1080747 | SaKaP_HI11P |
| 2 | 125580472 | 128775793 | q14.3 | q14.3 | 250 | 3195322 | SaKaP_HI11P |
| 2 | 158821507 | 160664848 | q24.1 | q24.2 | 200 | 1843342 | SaKaP_HI11P |
| 2 | 230112156 | 238041387 | q36.3 | q37.3 | 650 | 7929232 | Sco.proband |
| 2 | 238790067 | 242235086 | q37.3 | q37.3 | 200 | 3445020 | Sco.proband |
| 2 | 59911787 | 61898960 | p16.1 | p15 | 150 | 1987174 | SpiAnP_HI33P |
| 2 | 114321833 | 115414003 | q14.1 | q14.1 | 100 | 1092171 | SpiAnP_HI33P |
| 2 | 117873546 | 120548651 | q14.1 | q14.2 | 250 | 2675106 | SpiAnP_HI33P |
| 2 | 133065490 | 134647518 | q21.2 | q21.2 | 200 | 1582029 | SpiAnP_HI33P |
| 2 | 160666646 | 162102808 | q24.2 | q24.2 | 150 | 1436163 | SpiAnP_HI33P |

**Table3.CHROMOSOME 2 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).



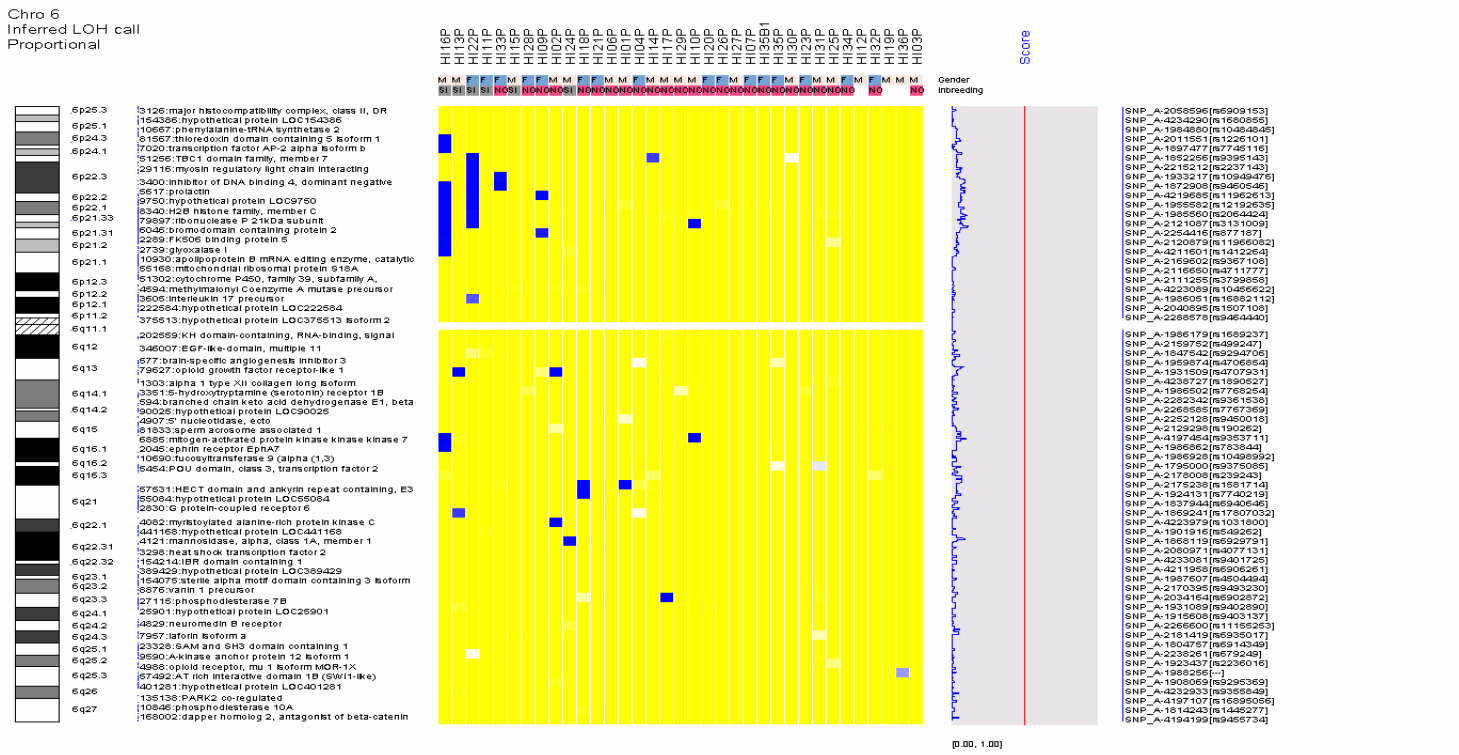

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 4 | 123882940 | 125063533 | q27 | q28.1 | 100 | 1180594 | Ben.probando |
| 4 | 125636494 | 127825459 | q28.1 | q28.1 | 200 | 2188966 | Ben.probando |
| 4 | 102215684 | 106049347 | q24 | q24 | 350 | 3833664 | Bi.proband_HI06P |
| 4 | 87270913 | 90351364 | q21.23 | q22.1 | 250 | 3080452 | RoNoP_HI09P |
| 4 | 151678846 | 153470722 | q31.23 | q31.3 | 150 | 1791877 | SaKaP_HI11P |
| 4 | 68630116 | 71433001 | q13.2 | q13.3 | 250 | 2802886 | Sco.proband |
| 4 | 72481637 | 107059283 | q13.3 | q24 | 3400 | 34577647 | Sco.proband |
| 4 | 107653134 | 147108571 | q24 | q31.21 | 3750 | 39455438 | Sco.proband |
| 4 | 61590871 | 63144614 | q13.1 | q13.1 | 150 | 1553744 | SpiAnP_HI33P |
| 4 | 180430069 | 182671024 | q34.3 | q34.3 | 350 | 2240956 | SpiAnP_HI33P |

**Table4. CHROMOSOME 4 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. TROHs regions were visualized by dChip (upper figure)..



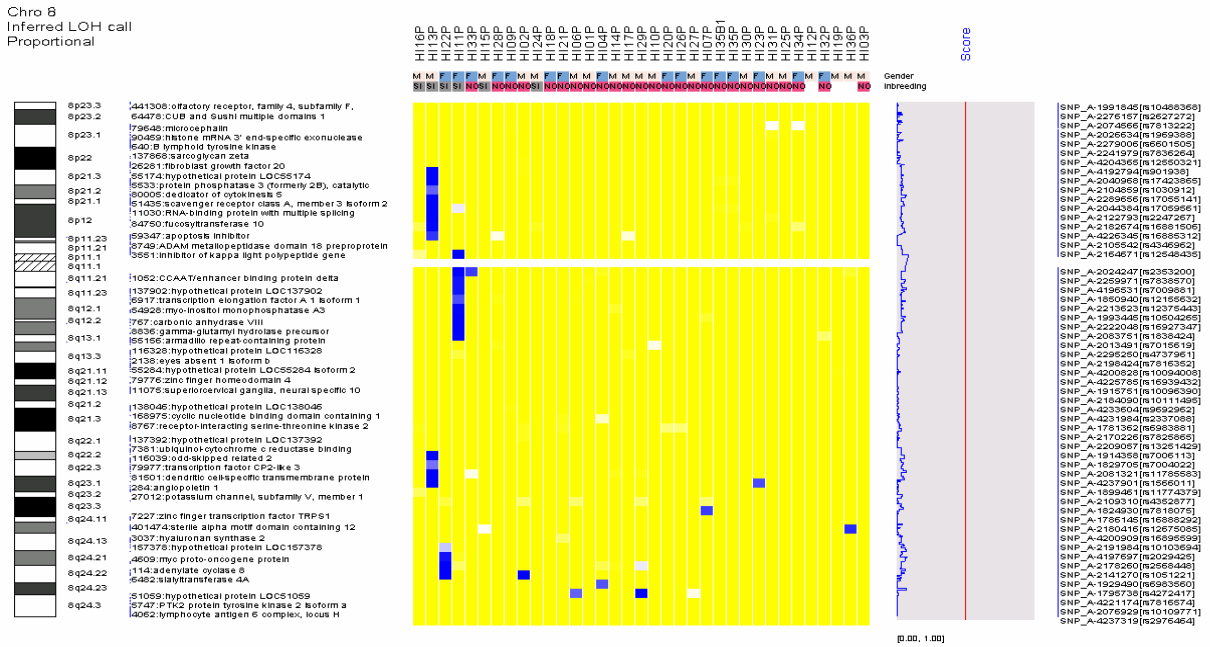

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length_cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 5 | 38873836 | 41540570 | p13.1 | p13.1 | 300 | 2666735 | ArrRaP_HI28P |
| 5 | 44371577 | 50139375 | p12 | q11.1 | 100 | 5767799 | ArrRaP_HI28P |
| 5 | 6941510 | 8042257 | p15.31 | p15.31 | 150 | 1100748 | Ben.probando |
| 5 | 97013319 | 98038320 | q15 | q21.1 | 100 | 1025002 | LaIlP_HI13P |
| 5 | 125337442 | 141836006 | q23.2 | q31.3 | 1300 | 16498565 | LaIlP_HI13P |
| 5 | 146785743 | 159094448 | q32 | q33.3 | 1300 | 12308706 | LaIlP_HI13P |
| 5 | 175136628 | 180425968 | q35.2 | q35.3 | 300 | 5289341 | LaIlP_HI13P |
| 5 | 62438367 | 63849152 | q12.1 | q12.3 | 100 | 1410786 | Pe.proband |
| 5 | 156051076 | 157212338 | q33.3 | q33.3 | 100 | 1161263 | RoDaP_HI10P |
| 5 | 116188761 | 120597319 | q23.1 | q23.1 | 487 | 4408559 | SaKaP_HI11P |
| 5 | 138241993 | 139323573 | q31.2 | q31.2 | 50 | 1081581 | SaKaP_HI11P |
| 5 | 146271461 | 147554192 | q32 | q32 | 150 | 1282732 | SaKaP_HI11P |
| 5 | 160578993 | 167157158 | q34 | q34 | 700 | 6578166 | SaKaP_HI11P |

**Table5.CHROMOSOME 5 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label.ROHs regions were visualized by dChip (upper figure).



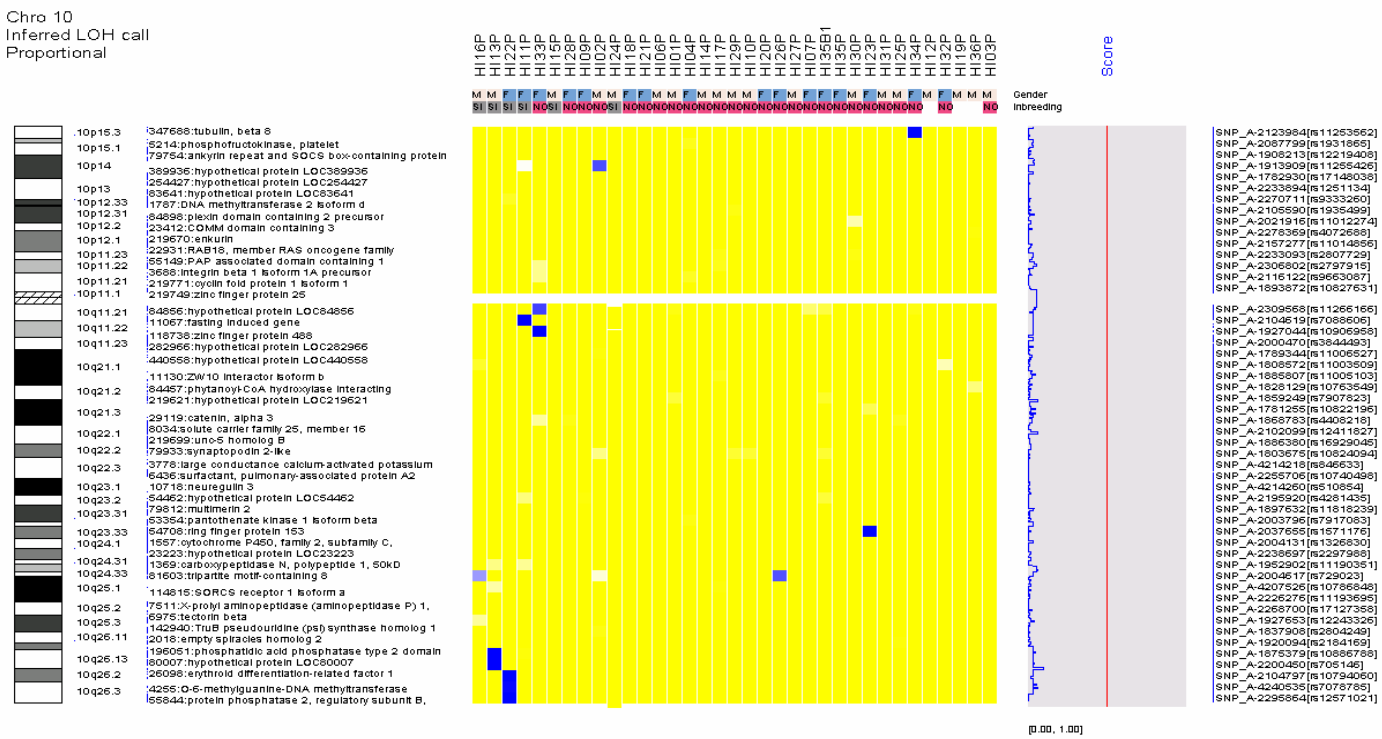

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 6 | 12063077 | 27912139 | p24.1 | p22.1 | 1700 | 15849063 | Ben.probando |
| 6 | 28410257 | 31350628 | p22.1 | p21.33 | 350 | 2940372 | Ben.probando |
| 6 | 31495946 | 33775943 | p21.33 | p21.31 | 200 | 2279998 | Ben.probando |
| 6 | 102233451 | 108970234 | q16.3 | q21 | 650 | 6736784 | Dol.proband |
| 6 | 72247314 | 74054537 | q13 | q13 | 200 | 1807224 | LaIlP_HI13P |
| 6 | 33776492 | 34873096 | p21.31 | p21.31 | 50 | 1096605 | RoNoP_HI09P |
| 6 | 119157216 | 120308807 | q22.31 | q22.31 | 100 | 1151592 | SaAnP_HI24P |
| 6 | 126032563 | 127111003 | q22.31 | q22.32 | 50 | 1078441 | SaKaP_HI11P |
| 6 | 5538964 | 10343900 | p25.1 | p24.3 | 600 | 4804937 | Sco.proband |
| 6 | 20121370 | 39072227 | p22.3 | p21.2 | 1850 | 18950858 | Sco.proband |
| 6 | 39734763 | 41617879 | p21.2 | p21.1 | 200 | 1883117 | Sco.proband |
| 6 | 88879335 | 95613212 | q15 | q16.1 | 735 | 6733878 | Sco.proband |
| 6 | 20121370 | 21504377 | p22.3 | p22.3 | 150 | 1383008 | SpiAnP_HI33P |
| 6 | 137954424 | 139551903 | q23.3 | q24.1 | 150 | 1597480 | SpiAnP_HI33P |

**Table6. CHROMOSOME 6 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).



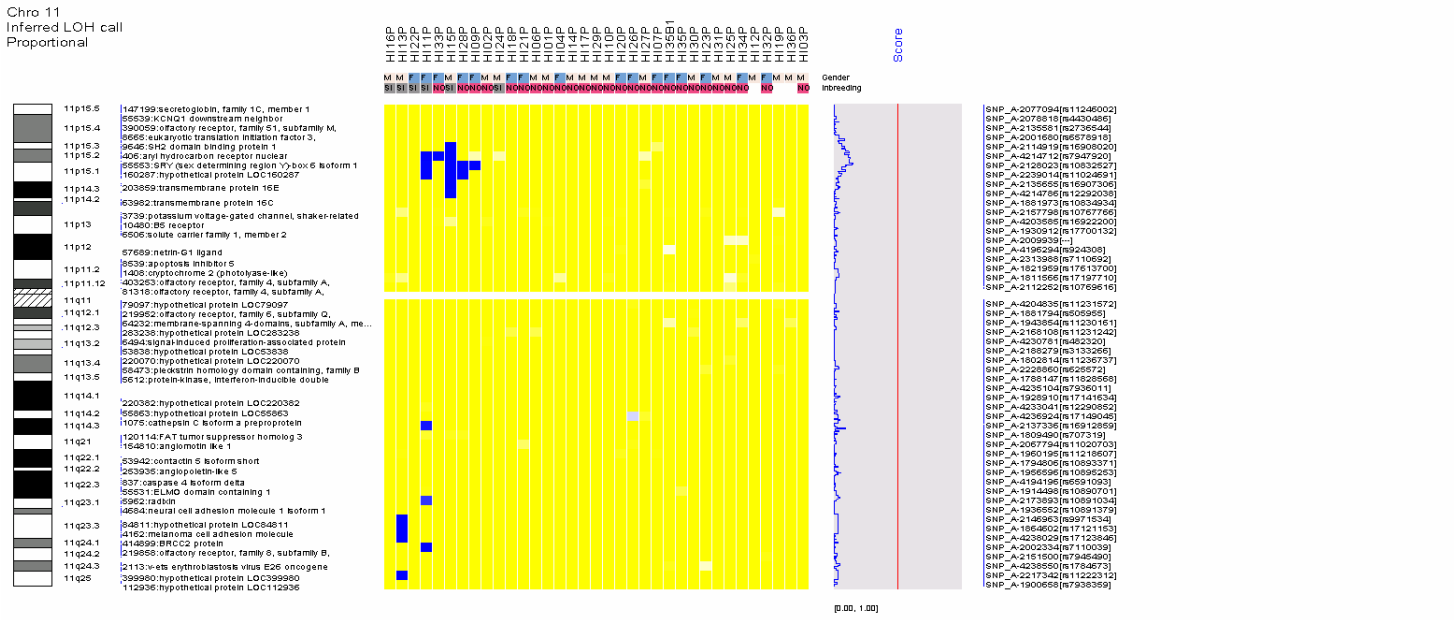

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 8 | 123294832 | 134427936 | q24.13 | q24.22 | 1187 | 11133105 | Ben.probando |
| 8 | 17627431 | 32757594 | p22 | p12 | 1750 | 15130164 | LaIlP_HI13P |
| 8 | 33235086 | 37709604 | p12 | p12 | 400 | 4474519 | LaIlP_HI13P |
| 8 | 98188237 | 107222273 | q22.1 | q23.1 | 750 | 9034037 | LaIlP_HI13P |
| 8 | 28418711 | 29552681 | p21.1 | p12 | 89 | 1133971 | SaKaP_HI11P |
| 8 | 40648225 | 65632682 | p11.21 | q12.3 | 1950 | 24984458 | SaKaP_HI11P |
| 8 | 43275955 | 48059141 | p11.1 | q11.1 | 50 | 4783187 | SpiAnP_HI33P |

**Table 7.CHROMOSOME 8 Single sample ROHs data.** ROHs events are extracted for each CHI proband. ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).



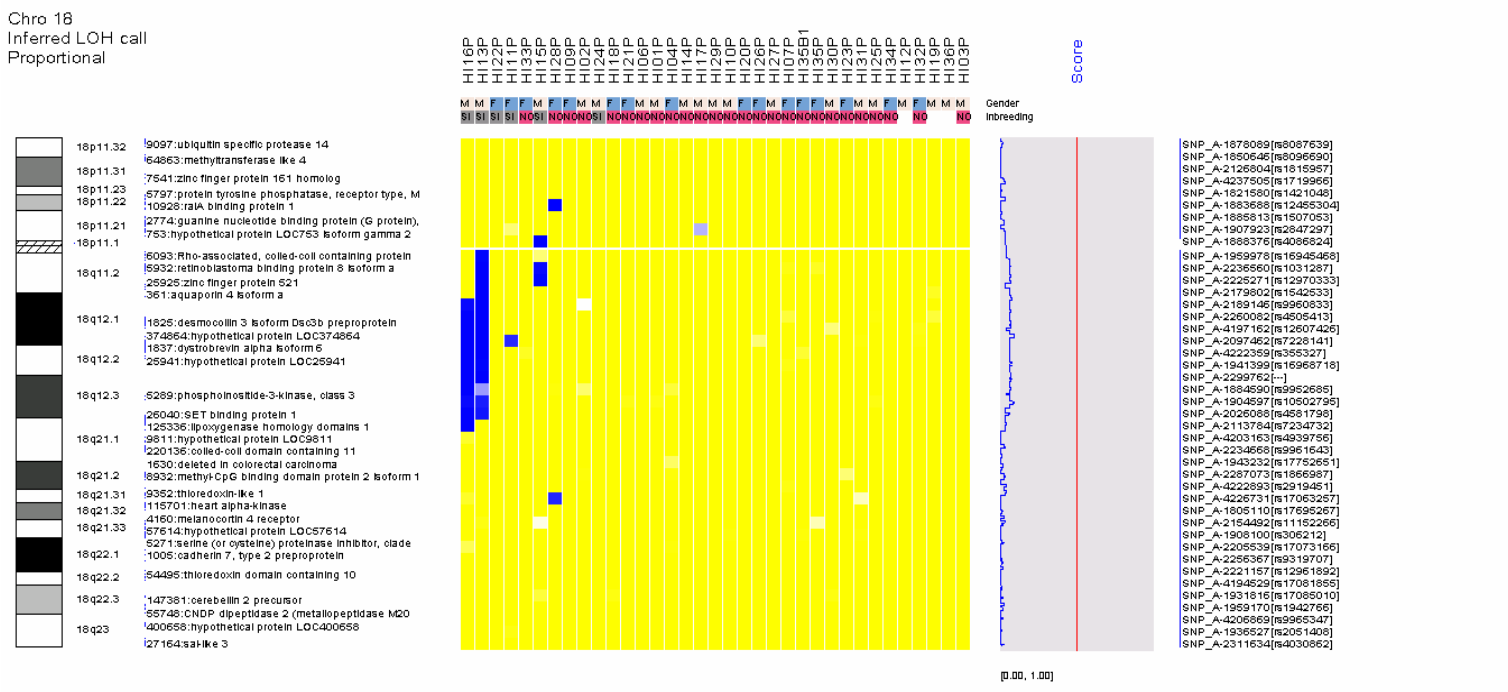

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 10 | 71304645 | 72327324 | q22.1 | q22.1 | 100 | 1022680 | Ben.probando |
| 10 | 127241907 | 129247017 | q26.13 | q26.2 | 250 | 2005111 | Ben.probando |
| 10 | 129750224 | 133317914 | q26.2 | q26.3 | 450 | 3567691 | Ben.probando |
| 10 | 103274164 | 105635715 | q24.32 | q24.33 | 100 | 2361552 | DaIlP_HI26P |
| 10 | 3166363 | 4906614 | p15.2 | p15.1 | 293 | 1740252 | LaIlP_HI13P |
| 10 | 7990564 | 9071510 | p14 | p14 | 200 | 1080947 | LaIlP_HI13P |
| 10 | 122116262 | 127727920 | q26.12 | q26.2 | 650 | 5611659 | LaIlP_HI13P |
| 10 | 32366043 | 33424122 | p11.22 | p11.22 | 100 | 1058080 | Risi.Probando |
| 10 | 38499125 | 42533657 | p11.21 | q11.21 | 50 | 4034533 | SaAnP_HI24P |
| 10 | 103274164 | 104650246 | q24.32 | q24.32 | 50 | 1376083 | Sco.proband |
| 10 | 29738242 | 31241111 | p11.23 | p11.23 | 200 | 1502870 | SpiAnP_HI33P |
| 10 | 31785387 | 32978388 | p11.22 | p11.22 | 100 | 1193002 | SpiAnP_HI33P |
| 10 | 38499125 | 43632112 | p11.21 | q11.21 | 150 | 5132988 | SpiAnP_HI33P |
| 10 | 44328733 | 49062113 | q11.21 | q11.22 | 150 | 4733381 | SpiAnP_HI33P |

**Table 8.CHROMOSOME 10 Single sample ROHs data.** ROHs events are extracted for each CHI proband. ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).




| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 11 | 14860212 | 17570989 | p15.2 | p15.1 | 250 | 2710778 | ArrRaP_HI28P |
| 11 | 114354151 | 119744172 | q23.2 | q23.3 | 450 | 5390022 | LaIlP_HI13P |
| 11 | 130116269 | 132698363 | q24.3 | q25 | 300 | 2582095 | LaIlP_HI13P |
| 11 | 8687209 | 23705016 | p15.4 | p14.3 | 1800 | 15017808 | My.proband |
| 11 | 13578990 | 16926330 | p15.2 | p15.1 | 300 | 3347341 | RoNoP_HI09P |
| 11 | 11806675 | 18725932 | p15.3 | p15.1 | 700 | 6919258 | SaKaP_HI11P |
| 11 | 10991925 | 12222118 | p15.3 | p15.3 | 150 | 1230194 | SpiAnP_HI33P |

**Table 9.CHROMOSOME 11 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).




| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster | Sample Name |
|---|---|---|---|---|---|---|---|
| 18 | 51895370 | 53511090 | q21.2 | q21.31 | 200 | 1615721 | ArrRaP_HI28P |
| 18 | 16857220 | 28414791 | q11.1 | q12.1 | 1100 | 11557572 | LaIlP_HI13P |
| 18 | 28782880 | 37725939 | q12.1 | q12.3 | 850 | 8943060 | LaIlP_HI13P |
| 18 | 38183758 | 40350391 | q12.3 | q12.3 | 250 | 2166634 | LaIlP_HI13P |
| 18 | 13544310 | 15096727 | p11.21 | p11.21 | 100 | 1552418 | My.proband |
| 18 | 18042301 | 21828290 | q11.2 | q11.2 | 350 | 3785990 | My.proband |
| 18 | 22487439 | 38183491 | q11.2 | q12.3 | 1600 | 15696053 | Sco.proband |
| 18 | 38493571 | 43812253 | q12.3 | q21.1 | 550 | 5318683 | Sco.proband |

**Tabele 10. CHROMOSOME 18 Single sample ROHs data.** ROHs events are extracted for each CHI proband . ROHs are described by chromosome, physical position (start-end), cytoband (start-end), number of SNPs markers, , length (bp) and sample label. ROHs regions were visualized by dChip (upper figure).

# 3.Homozygosity mapping

In order to realize an homozygosity mapping of our CHI probands dataset, we produced a list of homozygous regions in common between at least two probands ( chromosome clusters ) including more then 50 SNPs, with start and end physical position, cytoband, length of the cluster , minimal and maximal number of patients who are enclosed in that cluster. To further visualize the presence of these common ROHs on every proband genome-wide LOH profile, we have also obtained the clinical dataset " fingerprint ", which gave us the opportunity to visualize at same time the physical positioning of the common ROHs and the probands bearing them.

We have evidenced clusters of homozygosity traits in common between at least two probands (Table 11) on nearly all the chromosomes, except chromosome 21, with a total amount of 133 clusters including more then 50 SNPs; chromosome 11, with 16 ROHs, counted the largest amount of cluster followed by chromosomes 1 and 4 (13 clusters). The longest chromosome cluster was on chromosome 18, of 8943060 base pair in length, including 850 homozygous SNPs. A full list of 965 annotated genes enclosed in these regions was also produced (Table 16 ).

In this list we evidenced genes of the fatty acid metabolism as class III alcohol dehydrogenase 5 $\chi$ subunit , class II alcohol dehydrogenase 4 $\rho$ subunit , class V alcohol dehydrogenase 6 isoform 1, class I alcohol dehydrogenase,$\alpha$ subunit, alcohol dehydrogenase 1B, $\gamma$ subunit and class IV alcohol dehydrogenase,7 $\mu$ or $\sigma$ subunit on chromosome 4, acyl-CoA synthetase long-chain family member 6 (ACSL6) on chromosome 6 and cytochrome P450, family 27, subfamily C on chromosome 2, cytochrome P450, family 2, subfamily D on chromosome 22 and 3-hydroxybutyrate dehydrogenase, and type 2 (BDH2) on chromosome 4. One of the numerous homozygous traits in common between two patient on chromosome 1 was in 1p31; one of the two patients bore a long contiguous trait of 25 Mb. This common region contains medium-chain acyl-CoA dehydrogenase (MCAD) gene, involved in organic acid metabolism deficiency, that we considered a good candidate for mutational screening .

CHI probands'" fingerprint dataset showed us that patients with high homozygosity content, i.e number and extension of ROHs, had a common history of consanguinity and, as we can visualize in the fingerprint dataset wiev (Fig1), they clustered together on the base of homozygosity content. Consanguineous probands bearing a high number of autozygous regions had a good probability to bear the disease causing mutation on one of those chromosomal loci, but to better characterize them we focused our analysis to LOH traits shared between a broad number of patients. When we considered ROHs shared between at least two patients, we underlined that some of them were indeed partially in common between 3, 4, and also 5 probands; these regionswere the best candidate to bear disease associated mutations.

To further characterize our dataset, we analyzed only ROHs strictly in common between 3 and more than 3 probands (Fig.2) This analysis resulted in 35 cluster of common ROHs between 3 patients (Table12), 3 ROHs between 4 patients and one ROH between 5 patients and Table13). Five patients shared a common homozygous trait on cytoband 11p15, where the two major CHI candidate genes are located. Other relevant homozigosity regions in

common between 4 patients are present on chromosome 6 (p21.33-p21.31), on chromosome 10 (q26.13- q26.2) and on chromosome 11 (p15.2). A full list of 101 annotated genes in common between 3 probands and 42 genes in common between more than 3 probands were produced ( Table14 and Table15)

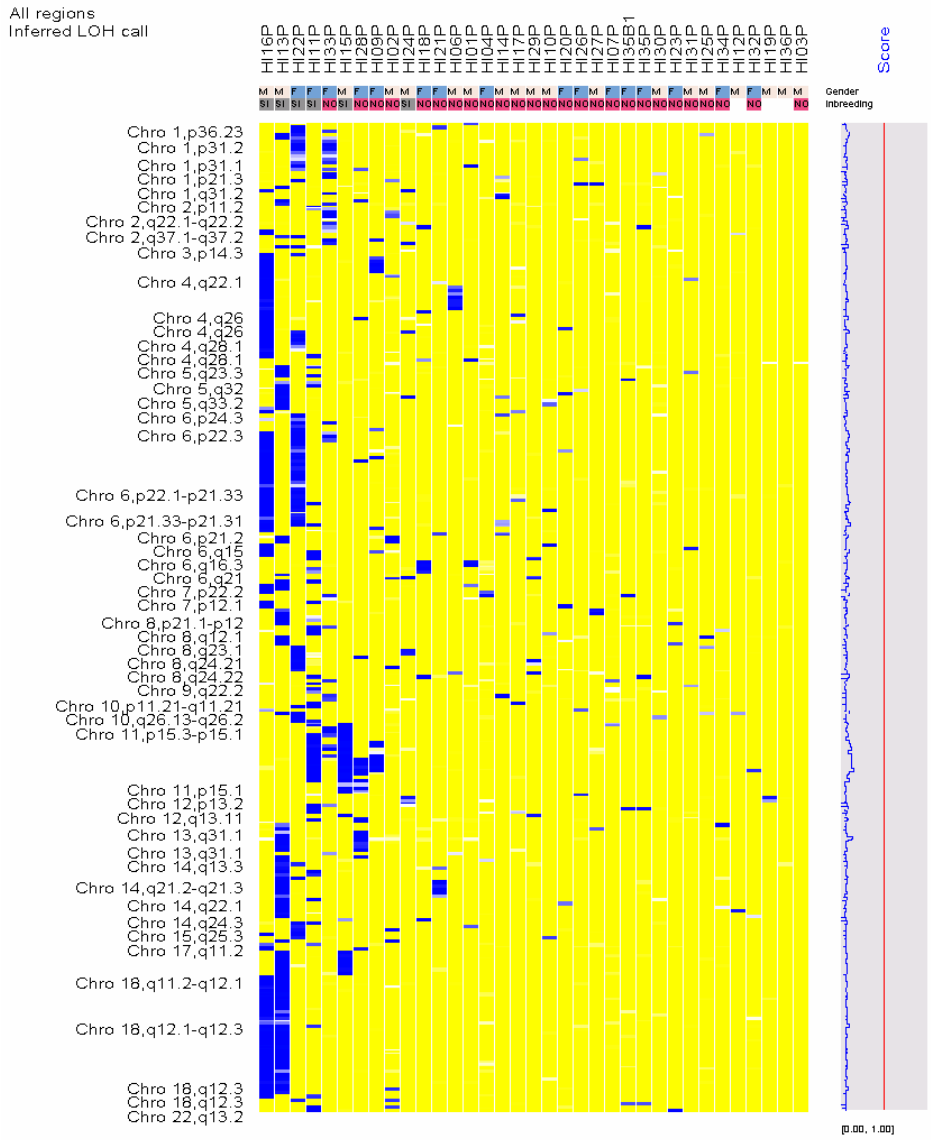



**Figure 1.Fingerprint dataset view.** The plot shows the fingerprint of 35 CHI patients composed by 164 ROHs clusters (blue boxes). Rows represent chromosome ROH regions and columns represent samples. We noticed that the number and the length of ROHs events were mainly associated to patients derived from consanguineous families (inbreeding).

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | freq.max | freq.min | Length_cluster |
|---|---|---|---|---|---|---|---|---|
| 1 | 54671315 | 55176764 | p32.3 | p32.3 | 50 | 2 | 2 | 505450 |
| 1 | 64101570 | 64493815 | p31.3 | p31.3 | 50 | 2 | 2 | 392246 |
| 1 | 65856897 | 66497951 | p31.2 | p31.2 | 100 | 3 | 3 | 641055 |
| 1 | 66795692 | 67566623 | p31.2 | p31.2 | 100 | 2 | 2 | 770932 |
| 1 | 68028825 | 68528821 | p31.2 | p31.2 | 50 | 2 | 2 | 499997 |
| 1 | 71714715 | 72767289 | p31.1 | p31.1 | 100 | 2 | 2 | 1052575 |
| 1 | 73701875 | 74847962 | p31.1 | p31.1 | 100 | 3 | 2 | 1146088 |
| 1 | 75905253 | 76192135 | p31.1 | p31.1 | 50 | 2 | 2 | 286883 |
| 1 | 77554882 | 78408842 | p31.1 | p31.1 | 50 | 3 | 3 | 853961 |
| 1 | 96195413 | 96568739 | p21.3 | p21.3 | 50 | 2 | 2 | 373327 |
| 1 | 106760930 | 107285607 | p21.1 | p13.3 | 50 | 2 | 2 | 524678 |
| 1 | 159693218 | 160026992 | q23.3 | q23.3 | 50 | 2 | 2 | 333775 |
| 1 | 163220916 | 163524154 | q24.1 | q24.1 | 50 | 2 | 2 | 303239 |
| 2 | 33013559 | 33375717 | p22.3 | p22.3 | 50 | 2 | 2 | 362159 |
| 2 | 67213820 | 67633078 | p14 | p14 | 50 | 2 | 2 | 419259 |
| 2 | 86069655 | 86605928 | p11.2 | p11.2 | 50 | 2 | 2 | 536274 |
| 2 | 117873546 | 118185629 | q14.1 | q14.1 | 50 | 2 | 2 | 312084 |
| 2 | 127365955 | 127971114 | q14.3 | q14.3 | 50 | 2 | 2 | 605160 |
| 2 | 133065490 | 133479538 | q21.2 | q21.2 | 50 | 2 | 2 | 414049 |
| 2 | 142243085 | 142502458 | q22.1 | q22.2 | 50 | 2 | 2 | 259374 |
| 2 | 176001354 | 176439366 | q31.1 | q31.1 | 50 | 2 | 2 | 438013 |
| 2 | 180348266 | 180702892 | q31.2 | q31.3 | 50 | 2 | 2 | 354627 |
| 2 | 185273737 | 186042074 | q32.1 | q32.1 | 50 | 3 | 3 | 768338 |
| 2 | 235212409 | 235654219 | q37.1 | q37.2 | 50 | 2 | 2 | 441811 |
| 3 | 23026938 | 24144365 | p24.3 | p24.2 | 100 | 2 | 2 | 1117428 |
| 3 | 54904586 | 55116477 | p14.3 | p14.3 | 50 | 2 | 2 | 211892 |
| 3 | 67774336 | 68236875 | p14.1 | p14.1 | 50 | 2 | 2 | 462540 |
| 4 | 84718250 | 85154258 | q21.23 | q21.23 | 50 | 2 | 2 | 436009 |
| 4 | 87270913 | 90351364 | q21.23 | q22.1 | 250 | 2 | 2 | 3080452 |
| 4 | 95037989 | 95636392 | q22.3 | q22.3 | 50 | 2 | 2 | 598404 |
| 4 | 98610487 | 99479372 | q22.3 | q23 | 50 | 2 | 2 | 868886 |
| 4 | 100363823 | 100765388 | q23 | q23 | 50 | 2 | 2 | 401566 |
| 4 | 102215684 | 106049347 | q24 | q24 | 350 | 2 | 2 | 3833664 |
| 4 | 115119102 | 115491760 | q26 | q26 | 50 | 2 | 2 | 372659 |

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | freq.max | freq.min | Length_cluster |
|---|---|---|---|---|---|---|---|---|
| 4 | 116551743 | 117096679 | q26 | q26 | 50 | 2 | 2 | 544937 |
| 4 | 118150143 | 118444914 | q26 | q26 | 50 | 2 | 2 | 294772 |
| 4 | 119964813 | 120606395 | q26 | q26 | 96 | 2 | 2 | 641583 |
| 4 | 123882940 | 125063533 | q27 | q28.1 | 100 | 3 | 2 | 1180594 |
| 4 | 125636494 | 127825459 | q28.1 | q28.1 | 200 | 2 | 2 | 2188966 |
| 4 | 139853752 | 140601427 | q31.1 | q31.1 | 50 | 2 | 2 | 747676 |
| 5 | 19121196 | 19533456 | p14.3 | p14.3 | 50 | 3 | 2 | 412261 |
| 5 | 115510693 | 115760721 | q23.1 | q23.1 | 50 | 2 | 2 | 250029 |
| 5 | 128013505 | 128664075 | q23.3 | q23.3 | 78 | 3 | 2 | 650571 |
| 5 | 131348629 | 131968240 | q23.3 | q23.3 | 50 | 2 | 2 | 619612 |
| 5 | 138241993 | 139323573 | q31.2 | q31.2 | 50 | 2 | 2 | 1081581 |
| 5 | 146785743 | 147554192 | q32 | q32 | 100 | 2 | 2 | 768450 |
| 5 | 150654041 | 151630108 | q33.1 | q33.1 | 100 | 2 | 2 | 976068 |
| 5 | 152939962 | 153345621 | q33.2 | q33.2 | 50 | 3 | 3 | 405660 |
| 5 | 156051076 | 157727618 | q33.3 | q33.3 | 150 | 2 | 2 | 1676543 |
| 6 | 8987431 | 9629747 | p24.3 | p24.3 | 50 | 2 | 2 | 642317 |
| 6 | 12871825 | 13275190 | p24.1 | p24.1 | 50 | 2 | 2 | 403366 |
| 6 | 14698976 | 15229194 | p23 | p23 | 50 | 2 | 2 | 530219 |
| 6 | 17567309 | 18574729 | p22.3 | p22.3 | 100 | 2 | 2 | 1007421 |
| 6 | 19702679 | 27912139 | p22.3 | p22.1 | 850 | 3 | 2 | 8209461 |
| 6 | 28410257 | 31350628 | p22.1 | p21.33 | 350 | 3 | 2 | 2940372 |
| 6 | 31495946 | 34873096 | p21.33 | p21.31 | 250 | 4 | 2 | 3377151 |
| 6 | 68925053 | 69328692 | q12 | q12 | 50 | 2 | 2 | 403640 |
| 6 | 72247314 | 72977664 | q13 | q13 | 100 | 2 | 2 | 730351 |
| 6 | 90256502 | 90834670 | q15 | q15 | 50 | 2 | 2 | 578169 |
| 6 | 91255941 | 91600801 | q15 | q15 | 50 | 2 | 2 | 344861 |
| 6 | 92077331 | 92870077 | q15 | q16.1 | 100 | 3 | 2 | 792747 |
| 6 | 102730295 | 103357098 | q16.3 | q16.3 | 50 | 2 | 2 | 626804 |
| 6 | 103759637 | 104648730 | q16.3 | q16.3 | 94 | 2 | 2 | 889094 |
| 6 | 107648780 | 108185814 | q21 | q21 | 50 | 2 | 2 | 537035 |
| 6 | 119789644 | 120308807 | q22.31 | q22.31 | 50 | 3 | 2 | 519164 |
| 7 | 1855676 | 2779462 | p22.3 | p22.2 | 50 | 2 | 2 | 923787 |
| 7 | 3129724 | 3844994 | p22.2 | p22.2 | 100 | 3 | 2 | 715271 |
| 7 | 49162995 | 49611620 | p12.3 | p12.3 | 50 | 3 | 3 | 448626 |
| 7 | 52616401 | 52933626 | p12.1 | p12.1 | 50 | 2 | 2 | 317226 |

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | freq.max | freq.min | Length_cluster |
|---|---|---|---|---|---|---|---|---|
| 7 | 143167422 | 144403185 | q35 | q35 | 100 | 2 | 2 | 1235764 |
| 8 | 26395573 | 27057092 | p21.2 | p21.2 | 100 | 2 | 2 | 661520 |
| 8 | 28418711 | 29552681 | p21.1 | p12 | 89 | 2 | 2 | 1133971 |
| 8 | 36836117 | 37207258 | p12 | p12 | 50 | 2 | 2 | 371142 |
| 8 | 43275955 | 48059141 | p11.1 | q11.1 | 50 | 2 | 2 | 4783187 |
| 8 | 56337971 | 56606117 | q12.1 | q12.1 | 50 | 2 | 2 | 268147 |
| 8 | 61765166 | 62165454 | q12.2 | q12.2 | 50 | 2 | 2 | 400289 |
| 8 | 104598943 | 105273361 | q22.3 | q22.3 | 50 | 2 | 2 | 674419 |
| 8 | 106727853 | 107222273 | q23.1 | q23.1 | 50 | 2 | 2 | 494421 |
| 8 | 126718727 | 127587550 | q24.13 | q24.21 | 100 | 2 | 2 | 868824 |
| 8 | 129218353 | 130444986 | q24.21 | q24.21 | 144 | 2 | 2 | 1226634 |
| 8 | 132749604 | 133137755 | q24.22 | q24.22 | 50 | 2 | 2 | 388152 |
| 8 | 138224786 | 138524186 | q24.23 | q24.23 | 50 | 2 | 2 | 299401 |
| 9 | 18109147 | 18464821 | p22.2 | p22.2 | 50 | 3 | 3 | 355675 |
| 9 | 74404364 | 74972741 | q21.13 | q21.13 | 50 | 2 | 2 | 568378 |
| 9 | 124127004 | 125801382 | q33.3 | q33.3 | 100 | 3 | 2 | 1674379 |
| 10 | 32366043 | 32978388 | p11.22 | p11.22 | 50 | 2 | 2 | 612346 |
| 10 | 38499125 | 42533657 | p11.21 | q11.21 | 50 | 2 | 2 | 4034533 |
| 10 | 64229033 | 64732014 | q21.3 | q21.3 | 50 | 2 | 2 | 502982 |
| 10 | 71877866 | 72327324 | q22.1 | q22.1 | 50 | 2 | 2 | 449459 |
| 10 | 103274164 | 104650246 | q24.32 | q24.32 | 50 | 2 | 2 | 1376083 |
| 10 | 127241907 | 127727920 | q26.13 | q26.2 | 50 | 4 | 4 | 486014 |
| 10 | 130964966 | 131345318 | q26.3 | q26.3 | 50 | 2 | 2 | 380353 |
| 10 | 131839011 | 132215151 | q26.3 | q26.3 | 50 | 2 | 2 | 376141 |
| 11 | 9571659 | 10053592 | p15.4 | p15.4 | 50 | 3 | 2 | 481934 |
| 11 | 10991925 | 18725932 | p15.3 | p15.1 | 800 | 5 | 2 | 7734008 |
| 11 | 19035090 | 20031569 | p15.1 | p15.1 | 146 | 2 | 2 | 996480 |
| 12 | 10821531 | 11860931 | p13.2 | p13.2 | 100 | 2 | 2 | 1039401 |
| 12 | 13630617 | 13952836 | p13.1 | p13.1 | 50 | 2 | 2 | 322220 |
| 12 | 17419385 | 17762650 | p12.3 | p12.3 | 50 | 3 | 3 | 343266 |
| 12 | 46517877 | 47210041 | q13.11 | q13.11 | 50 | 2 | 2 | 692165 |
| 12 | 51187562 | 51421288 | q13.13 | q13.13 | 50 | 2 | 2 | 233727 |
| 12 | 84742783 | 85757807 | q21.31 | q21.32 | 75 | 2 | 2 | 1015025 |
| 13 | 31605691 | 31934759 | q13.1 | q13.1 | 50 | 2 | 2 | 329069 |
| 13 | 77917728 | 78317321 | q31.1 | q31.1 | 50 | 2 | 2 | 399594 |

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | freq.max | freq.min | Length_cluster |
|---|---|---|---|---|---|---|---|---|
| 13 | 81539340 | 83150644 | q31.1 | q31.1 | 200 | 3 | 2 | 1611305 |
| 13 | 83651454 | 84443330 | q31.1 | q31.1 | 100 | 2 | 2 | 791877 |
| 14 | 32098567 | 32515204 | q12 | q13.1 | 50 | 2 | 2 | 416638 |
| 14 | 36130878 | 36341562 | q13.3 | q13.3 | 50 | 2 | 2 | 210685 |
| 14 | 38480432 | 39102940 | q21.1 | q21.1 | 50 | 2 | 2 | 622509 |
| 14 | 39694219 | 40050563 | q21.1 | q21.1 | 50 | 2 | 2 | 356345 |
| 14 | 41725025 | 43486802 | q21.1 | q21.2 | 150 | 2 | 2 | 1761778 |
| 14 | 45906862 | 47937414 | q21.2 | q21.3 | 250 | 2 | 2 | 2030553 |
| 14 | 51565125 | 51811396 | q22.1 | q22.1 | 50 | 2 | 2 | 246272 |
| 14 | 57373756 | 58482525 | q23.1 | q23.1 | 100 | 2 | 2 | 1108770 |
| 14 | 67049466 | 67914548 | q24.1 | q24.1 | 83 | 2 | 2 | 865083 |
| 15 | 36248992 | 36751867 | q14 | q14 | 50 | 2 | 2 | 502876 |
| 15 | 79627239 | 80154346 | q25.2 | q25.2 | 50 | 2 | 2 | 527108 |
| 15 | 83824542 | 84072454 | q25.3 | q25.3 | 50 | 2 | 2 | 247913 |
| 15 | 84572965 | 84946846 | q25.3 | q25.3 | 50 | 2 | 2 | 373882 |
| 15 | 88824604 | 89518149 | q26.1 | q26.1 | 50 | 2 | 2 | 693546 |
| 16 | 26150765 | 26543513 | p12.1 | p12.1 | 50 | 2 | 2 | 392749 |
| 16 | 69792490 | 70285072 | q22.2 | q22.2 | 50 | 2 | 2 | 492583 |
| 17 | 47531938 | 47943987 | q22 | q22 | 50 | 2 | 2 | 412050 |
| 17 | 65866643 | 66285786 | q24.3 | q24.3 | 50 | 2 | 2 | 419144 |
| 18 | 18042301 | 21828290 | q11.2 | q11.2 | 350 | 2 | 2 | 3785990 |
| 18 | 22487439 | 28414791 | q11.2 | q12.1 | 650 | 2 | 2 | 5927353 |
| 18 | 28782880 | 37725939 | q12.1 | q12.3 | 850 | 3 | 2 | 8943060 |
| 18 | 38493571 | 40350391 | q12.3 | q12.3 | 200 | 3 | 2 | 1856821 |
| 18 | 41210575 | 41498356 | q12.3 | q12.3 | 50 | 2 | 2 | 287782 |
| 19 | 48230781 | 49136256 | q13.31 | q13.31 | 50 | 2 | 2 | 905476 |
| 20 | 12307879 | 12604055 | p12.1 | p12.1 | 50 | 2 | 2 | 296177 |
| 22 | 40122324 | 41385992 | q13.2 | q13.2 | 50 | 2 | 2 | 1263669 |
| 22 | 42110517 | 42576227 | q13.2 | q13.2 | 50 | 2 | 2 | 465711 |

**Table 11. Fingerprint data.** 35 CHI patients homozygosity mapping analysis. In table are listed a total of 133 ROHs including more than 50 SNPs. The ROHs regions were visualized by dChip giving the fingerprint of CHI patients.

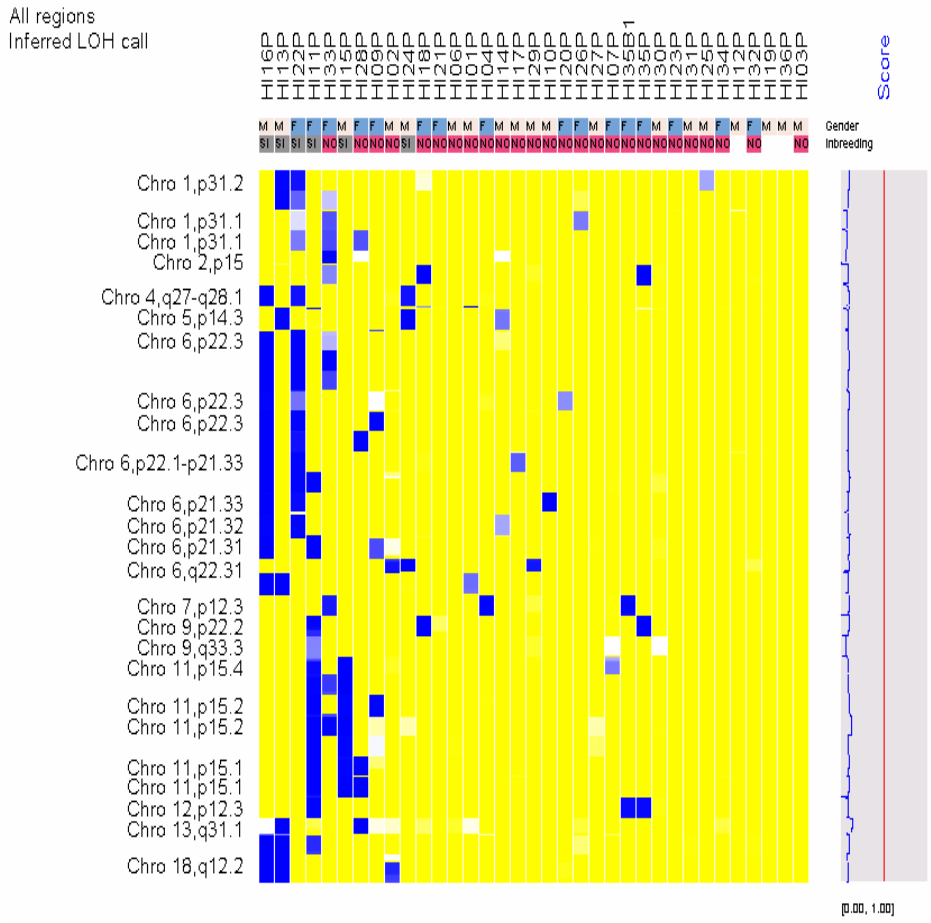



**Figure 2. Three patients shared ROHs view.** The plot shows 35 shared ROHs pattern (blue boxes) obtained setting the number of CHI patients equal to 3.

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Length _cluster |
|---|---|---|---|---|---|---|
| 1 | 65856897 | 66497951 | p31.2 | p31.2 | 100 | 641055 |
| 1 | 73701875 | 74190336 | p31.1 | p31.1 | 50 | 488462 |
| 1 | 77554882 | 78408842 | p31.1 | p31.1 | 50 | 853961 |
| 2 | 61288114 | 61898960 | p15 | p15 | 33 | 610847 |
| 2 | 185273737 | 186042074 | q32.1 | q32.1 | 50 | 768338 |
| 4 | 123882940 | 124448184 | q27 | q28.1 | 50 | 565245 |
| 5 | 19269375 | 19269562 | p14.3 | p14.3 | 2 | 188 |
| 5 | 128226162 | 128228401 | q23.3 | q23.3 | 3 | 2240 |
| 5 | 152939962 | 153345621 | q33.2 | q33.2 | 50 | 405660 |

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | Lenght _cluster |
|---|---|---|---|---|---|---|
| 6 | 20121370 | 21504377 | p22.3 | p22.3 | 150 | 1383008 |
| 6 | 22516941 | 22722709 | p22.3 | p22.3 | 50 | 205769 |
| 6 | 23279793 | 23976573 | p22.3 | p22.3 | 100 | 696781 |
| 6 | 29916011 | 30455218 | p22.1 | p21.33 | 100 | 539208 |
| 6 | 31094483 | 31350628 | p21.33 | p21.33 | 50 | 256146 |
| 6 | 32717405 | 33087748 | p21.32 | p21.32 | 50 | 370344 |
| 6 | 33645780 | 33775943 | p21.31 | p21.31 | 8 | 130164 |
| 6 | 92077331 | 92531884 | q15 | q16.1 | 50 | 454554 |
| 6 | 120034143 | 120308807 | q22.31 | q22.31 | 33 | 274665 |
| 7 | 3129724 | 3476429 | p22.2 | p22.2 | 50 | 346706 |
| 7 | 49162995 | 49611620 | p12.3 | p12.3 | 50 | 448626 |
| 9 | 18109147 | 18464821 | p22.2 | p22.2 | 50 | 355675 |
| 9 | 125236073 | 125801382 | q33.3 | q33.3 | 50 | 565310 |
| 11 | 9681193 | 10053592 | p15.4 | p15.4 | 45 | 372400 |
| 11 | 11806675 | 12222118 | p15.3 | p15.3 | 50 | 415444 |
| 11 | 12621117 | 12976818 | p15.2 | p15.2 | 50 | 355702 |
| 11 | 13187209 | 13982361 | p15.2 | p15.2 | 100 | 795153 |
| 11 | 16938063 | 17570989 | p15.1 | p15.1 | 50 | 632927 |
| 11 | 18154646 | 18725932 | p15.1 | p15.1 | 50 | 571287 |
| 12 | 17419385 | 17762650 | p12.3 | p12.3 | 50 | 343266 |
| 13 | 82383339 | 82639797 | q31.1 | q31.1 | 34 | 256459 |
| 18 | 29391054 | 29875256 | q12.1 | q12.1 | 50 | 484203 |
| 18 | 33555925 | 33699639 | q12.2 | q12.2 | 17 | 143715 |
| 18 | 34170274 | 34170274 | q12.2 | q12.2 | 1 | 1 |
| 18 | 36034385 | 36035495 | q12.3 | q12.3 | 2 | 1111 |
| 18 | 39445931 | 39907894 | q12.3 | q12.3 | 50 | 461964 |

**Table 12.Three patients shared ROHs data**. In table identified ROHs based on fixed frequency across the dataset are listed . Using three patients as fixed frequency, we found a total of 34clusters. ROHs regions were visualized by dChip giving the view of common ROH pattern of CHI patients.

| Gene Symbol | Chromosome | Start | End | refLink.product | hg17.refLink.locus | hg17.refLink.omimId |
|---|---|---|---|---|---|---|
| PDE4B | chr1 | 65970213 | 66552282 | phosphodiesterase 4B, cAMP-specific isoform 1 | 5142 | 600127 |
| AK5 | chr1 | 77459762 | 77737673 | adenylate kinase 5 isoform 1 | 26289 | 608009 |
| ZZZ3 | chr1 | 77742210 | 77860364 | zinc finger, ZZ-type containing 3 | 26009 | 0 |
| USP33 | chr1 | 77873696 | 77937558 | ubiquitin specific protease 33 isoform 1 | 23032 | 0 |
| FAM73A | chr1 | 77957329 | 78056102 | hypothetical protein LOC374986 | 374986 | 0 |
| NEXN | chr1 | 78066220 | 78121599 | nexilin (F actin binding protein) | 91624 | 0 |
| FUBP1 | chr1 | 78125613 | 78156798 | far upstream element-binding protein | 8880 | 603444 |
| DNAJB4 | chr1 | 78182656 | 78195014 | DnaJ (Hsp40) homolog, subfamily B, member 4 | 11080 | 611327 |
| GIPC2 | chr1 | 78223609 | 78315132 | PDZ domain protein GIPC2 | 54810 | 0 |
| AHSA2 | chr2 | 61316471 | 61325709 | AHA1, activator of heat shock 90kDa protein | 130872 | 0 |
| USP34 | chr2 | 61326241 | 61609500 | ubiquitin specific protease 34 | 9736 | 0 |
| XPO1 | chr2 | 61616720 | 61677069 | exportin 1 | 7514 | 602559 |
| ZNF804A | chr2 | 185288598 | 185629718 | zinc finger protein 804A | 91752 | 0 |
| IL21 | chr4 | 123891387 | 123899816 | interleukin 21 | 59067 | 605384 |
| BBS12 | chr4 | 124011461 | 124023702 | Bardet-Biedl syndrome 12 | 166379 | 610683 |
| FGF2 | chr4 | 124105467 | 124176995 | fibroblast growth factor 2 | 2247 | 134920 |
| NUDT6 | chr4 | 124171410 | 124201728 | nudix-type motif 6 isoform b | 11162 | 606261 |
| SPATA5 | chr4 | 124201829 | 124598209 | spermatogenesis associated 5 | 166378 | 0 |
| GRIA1 | chr5 | 152850276 | 153173622 | glutamate receptor, ionotropic, AMPA 1 isoform | 2890 | 138248 |
| MBOAT1 | chr6 | 20208913 | 20320649 | membrane bound O-acyltransferase domain | 154141 | 611732 |
| E2F3 | chr6 | 20510115 | 20601924 | E2F transcription factor 3 | 1871 | 600427 |
| CDKAL1 | chr6 | 20642666 | 21339743 | CDK5 regulatory subunit associated protein | 54901 | 611259 |
| HDGFL1 | chr6 | 22677656 | 22678728 | hepatoma derived growth factor-like 1 | 154150 | 0 |
| HLA-A29.1 | chr6 | 30018304 | 30085130 | major histocompatibility complex class I | 649853 | 0 |
| HLA-A | chr6 | 30018309 | 30021632 | major histocompatibility complex, class I, A | 3105 | 142800 |
| HCG9 | chr6 | 30050870 | 30054148 | HLA complex group 9 | 10255 | 0 |
| ZNRD1 | chr6 | 30137014 | 30140664 | zinc ribbon domain containing 1 | 30834 | 607525 |
| PPP1R11 | chr6 | 30142910 | 30146085 | protein phosphatase 1, regulatory (inhibitor) | 6992 | 606670 |
| RNF39 | chr6 | 30146022 | 30151607 | ring finger protein 39 isoform 1 | 80352 | 607524 |
| TRIM31 | chr6 | 30178654 | 30188846 | tripartite motif protein 31 | 11074 | 609316 |
| TRIM40 | chr6 | 30212488 | 30224491 | tripartite motif-containing 40 | 135644 | 0 |
| TRIM10 | chr6 | 30227702 | 30236690 | tripartite motif-containing 10 isoform 1 | 10107 | 605701 |
| TRIM15 | chr6 | 30238961 | 30248452 | tripartite motif protein 15 | 89870 | 0 |
| TRIM26 | chr6 | 30260212 | 30289132 | tripartite motif-containing 26 | 7726 | 600830 |

| | | | | | | |
|---|---|---|---|---|---|---|
| FLJ45422 | chr6 | 30335352 | 30342707 | hypothetical protein LOC441140 | 441140 | 0 |
| TRIM39 | chr6 | 30403018 | 30419484 | tripartite motif-containing 39 isoform 1 | 56658 | 605700 |
| RPP21 | chr6 | 30420915 | 30422611 | ribonuclease P/MRP 21kDa subunit | 79897 | 0 |
| C6orf15 | chr6 | 31186980 | 31188311 | STG protein | 29113 | 611401 |
| PSORS1C1 | chr6 | 31190601 | 31215816 | SEEK1 protein | 170679 | 0 |
| CDSN | chr6 | 31190848 | 31196202 | corneodesmosin precursor | 1041 | 602593 |
| PSORS1C2 | chr6 | 31213289 | 31215106 | SPR1 protein | 170680 | 0 |
| CCHCR1 | chr6 | 31218194 | 31233545 | coiled-coil alpha-helical rod protein 1 isoform | 54535 | 605310 |
| TCF19 | chr6 | 31234281 | 31239970 | transcription factor 19 | 6941 | 600912 |
| POU5F1 | chr6 | 31240093 | 31242006 | POU domain, class 5, transcription factor 1 | 5460 | 164177 |
| HCG27 | chr6 | 31273577 | 31279723 | HLA complex group 27 | 253018 | 0 |
| HLA-C | chr6 | 31344508 | 31347834 | major histocompatibility complex, class I, C | 3107 | 142840 |
| HLA-DQA1 | chr6 | 32713160 | 32719406 | major histocompatibility complex, class II, DQ | 3117 | 146880 |
| HLA-DQB1 | chr6 | 32735634 | 32742444 | major histocompatibility complex, class II, DQ | 3119 | 604305 |
| HLA-DQA2 | chr6 | 32817140 | 32823197 | major histocompatibility complex, class II, DQ | 3118 | 0 |
| HLA-DOB | chr6 | 32888518 | 32892803 | major histocompatibility complex, class II, DO | 3112 | 600629 |
| TAP2 | chr6 | 32897587 | 32914525 | transporter 2, ATP-binding cassette, sub-family | 6891 | 170261 |
| PSMB8 | chr6 | 32916476 | 32919794 | proteasome beta 8 subunit isoform E2 proprotein | 5696 | 177046 |
| TAP1 | chr6 | 32920964 | 32929726 | transporter 1, ATP-binding cassette, sub-family | 6890 | 170260 |
| PSMB9 | chr6 | 32929915 | 32935604 | proteasome beta 9 subunit isoform 1 proprotein | 5698 | 177045 |
| HLA-DMB | chr6 | 33010393 | 33016795 | major histocompatibility complex, class II, DM | 3109 | 142856 |
| HLA-DMA | chr6 | 33024373 | 33028831 | major histocompatibility complex, class II, DM | 3108 | 142855 |
| BRD2 | chr6 | 33044414 | 33057260 | bromodomain containing 2 | 6046 | 601540 |
| HLA-DOA | chr6 | 33079939 | 33085367 | major histocompatibility complex, class II, DO | 3111 | 142930 |
| BAK1 | chr6 | 33648301 | 33656048 | BCL2-antagonist/killer 1 | 578 | 600516 |
| FLJ43752 | chr6 | 33661860 | 33669093 | hypothetical protein LOC401253 | 401253 | 0 |
| ITPR3 | chr6 | 33697138 | 33772326 | inositol 1,4,5-triphosphate receptor, type 3 | 3710 | 147267 |
| C6orf125 | chr6 | 33773323 | 33787482 | hypothetical protein LOC84300 | 84300 | 0 |
| SDK1 | chr7 | 3114320 | 4081870 | sidekick 1 isoform 1 | 221935 | 607216 |
| VWC2 | chr7 | 49590517 | 49729399 | von Willebrand factor C domain containing 2 | 375567 | 611108 |
| ADAMTSL1 | chr9 | 18464103 | 18674952 | ADAMTS-like 1 isoform 2 precursor | 92949 | 609198 |
| MAPKAP1 | chr9 | 125279227 | 125549067 | mitogen-activated protein kinase associated | 79109 | 610558 |
| PBX3 | chr9 | 125589219 | 125809207 | pre-B-cell leukemia homeobox 3 | 5090 | 176312 |
| SWAP70 | chr11 | 9642203 | 9731081 | SWAP-70 protein | 23075 | 604762 |
| SBF2 | chr11 | 9756789 | 10272330 | SET binding factor 2 | 81846 | 607697 |
| USP47 | chr11 | 11819545 | 11937446 | ubiquitin specific protease 47 | 55031 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| DKK3 | chr11 | 11941120 | 11987205 | dickkopf homolog 3 precursor | 27122 | 605416 |
| MICAL2 | chr11 | 12088713 | 12241905 | microtubule associated monoxygenase, calponin | 9645 | 608881 |
| TEAD1 | chr11 | 12652544 | 12922875 | TEA domain family member 1 | 7003 | 189967 |
| ARNTL | chr11 | 13255900 | 13365386 | aryl hydrocarbon receptor nuclear | 406 | 602550 |
| BTBD10 | chr11 | 13366133 | 13441414 | K+ channel tetramerization protein | 84280 | 0 |
| PTH | chr11 | 13470176 | 13474143 | parathyroid hormone preproprotein | 5741 | 168450 |
| FAR1 | chr11 | 13646846 | 13710467 | fatty acyl CoA reductase 1 | 84188 | 0 |
| SPON1 | chr11 | 13940489 | 14246222 | spondin 1, extracellular matrix protein | 10418 | 604989 |
| PLEKHA7 | chr11 | 16765787 | 16992535 | pleckstrin homology domain containing, family A | 144100 | 0 |
| RPS13 | chr11 | 17052515 | 17055796 | ribosomal protein S13 | 6207 | 180476 |
| PIK3C2A | chr11 | 17064699 | 17147930 | phosphoinositide-3-kinase, class 2, alpha | 5286 | 603601 |
| NUCB2 | chr11 | 17254861 | 17309645 | nucleobindin 2 | 4925 | 608020 |
| KCNJ11 | chr11 | 17363373 | 17366782 | potassium inwardly-rectifying channel J11 | 3767 | 600937 |
| ABCC8 | chr11 | 17371007 | 17455025 | ATP-binding cassette, sub-family C, member 8 | 6833 | 600509 |
| USH1C | chr11 | 17472035 | 17522539 | harmonin isoform a | 10083 | 605242 |
| SAA4 | chr11 | 18209479 | 18214910 | serum amyloid A4, constitutive | 6291 | 104752 |
| SAA2 | chr11 | 18217163 | 18226758 | serum amyloid A2 isoform b | 6289 | 104751 |
| SAA1 | chr11 | 18244347 | 18248102 | serum amyloid A1 preproprotein | 6288 | 104750 |
| HPS5 | chr11 | 18256792 | 18300297 | Hermansky-Pudlak syndrome 5 isoform b | 11234 | 607521 |
| GTF2H1 | chr11 | 18300718 | 18345152 | general transcription factor IIH, polypeptide 1, | 2965 | 189972 |
| LDHA | chr11 | 18372557 | 18385974 | lactate dehydrogenase A | 3939 | 150000 |
| LDHC | chr11 | 18390433 | 18429369 | lactate dehydrogenase C | 3948 | 150150 |
| LDHAL6A | chr11 | 18434006 | 18457723 | lactate dehydrogenase A-like 6A | 160287 | 0 |
| tumor susceptibility gene 101 | chr11 | 18458434 | 18505065 | | 7251 | 601387 |
| ubiquitin-conjugating enzyme E2-like isoform a | chr11 | 18509820 | 18566857 | | 55293 | 610985 |
| SPTY2D1 | chr11 | 18584524 | 18612596 | SPT2, Suppressor of Ty, domain containing 1 | 144108 | 0 |
| TMEM86A | chr11 | 18676926 | 18682908 | transmembrane protein 86A | 144110 | 0 |
| IGSF22 | chr11 | 18682427 | 18704353 | immunoglobulin superfamily, member 22 | 283284 | 0 |
| PTPN5 | chr11 | 18706053 | 18769965 | protein tyrosine phosphatase, non-receptor type | 84867 | 176879 |
| ASXL3 | chr18 | 29412538 | 29581397 | additional sex combs like 3 | 80816 | 0 |
| NOL4 | chr18 | 29685061 | 30057444 | nucleolar protein 4 | 8715 | 603577 |

**Table 14 .Three patients shared ROHs gene annotation.** The list of genes associated to ROHs was obtained using the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables) and by querying the hg17 database.

| Chr_cluster | Start_cluster | End_cluster | Start_cytoband | End_cytoband | N_SNPs | freq.max | freq.min | Length_cluster |
|---|---|---|---|---|---|---|---|---|
| 6 | 33105555 | 33599035 | p21.32 | p21.32 | 42 | 4 | 4 | 493481 |
| 10 | 127241907 | 127727920 | q26.13 | q26.2 | 50 | 4 | 4 | 486014 |
| 11 | 12977034 | 13184740 | p15.2 | p15.2 | 50 | 4 | 4 | 207707 |
| 11 | 13994054 | 16926330 | p15.2 | p15.1 | 250 | 5 | 4 | 2932277 |

**Table 13 .Four and five patients shared ROHs data.** In table identified ROHs based on fixed frequency across the dataset are listed . Using more than three patients as fixed frequency, we found a total of 4 clusters. The ROHs regions were visualized by dChip giving the view of common ROH pattern of CHI patients

| Gene Symbol | Chromosome | Start | End | hg17.refLink.product | hg17.refLink.locus LinkId | hg17.refLink.omi mId |
|---|---|---|---|---|---|---|
| HLA-DPA1 | chr6 | 33140771 | 33149356 | major histocompatibility complex, class II, DP | 3113 | 142880 |
| HLA-DPB1 | chr6 | 33151737 | 33162954 | major histocompatibility complex, class II, DP | 3115 | 142858 |
| COL11A2 | chr6 | 33238446 | 33268223 | collagen, type XI, alpha 2 isoform 2 | 1302 | 120290 |
| RXRB | chr6 | 33269343 | 33276410 | retinoid X receptor, beta | 6257 | 180246 |
| SLC39A7 | chr6 | 33276580 | 33280191 | solute carrier family 39, member 7 | 7922 | 601416 |
| HSD17B8 | chr6 | 33280396 | 33282585 | estradiol 17 beta-dehydrogenase 8 | 7923 | 601417 |
| RING1 | chr6 | 33284263 | 33288476 | ring finger protein 1 | 6015 | 602045 |
| VPS52 | chr6 | 33326027 | 33347640 | vacuolar protein sorting 52 | 6293 | 603443 |
| RPS18 | chr6 | 33347829 | 33352259 | ribosomal protein S18 | 6222 | 180473 |
| B3GALT4 | chr6 | 33352894 | 33354580 | UDP-Gal:betaGlcNAc beta | 8705 | 603095 |
| WDR46 | chr6 | 33354863 | 33364969 | WD repeat domain 46 | 9277 | 0 |
| PFDN6 | chr6 | 33365355 | 33366689 | HLA class II region expressed gene KE2 | 10471 | 605660 |
| RGL2 | chr6 | 33367415 | 33374716 | ral guanine nucleotide dissociation | 5863 | 602306 |
| TAPBP | chr6 | 33375451 | 33389967 | tapasin isoform 1 precursor | 6892 | 601962 |
| ZBTB22 | chr6 | 33390173 | 33393472 | zinc finger and BTB domain containing 22 | 9278 | 0 |
| DAXX | chr6 | 33394378 | 33398682 | death-associated protein 6 | 1616 | 603186 |
| KIFC1 | chr6 | 33467290 | 33485677 | kinesin family member C1 | 3833 | 603763 |
| PHF1 | chr6 | 33486771 | 33492191 | PHD finger protein 1 isoform b | 5252 | 602881 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CUTA | chr6 | 33492296 | 33494043 | cutA divalent cation tolerance homolog isoform | 51596 | 0 |
| SYNGAP1 | chr6 | 33495824 | 33529444 | synaptic Ras GTPase activating protein 1 isoform | 8831 | 603384 |
| ZBTB9 | chr6 | 33530333 | 33533296 | zinc finger and BTB domain containing 9 | 221504 | 0 |
| C10orf122 | chr10 | 127334252 | 127361703 | hypothetical protein LOC387718 | 387718 | 0 |
| C10orf137 | chr10 | 127398073 | 127442702 | erythroid differentiation-related factor 1 | 26098 | 0 |
| MMP21 | chr10 | 127445015 | 127454380 | matrix metalloproteinase 21 preproprotein | 118856 | 608416 |
| UROS | chr10 | 127467136 | 127501827 | uroporphyrinogen III synthase | 7390 | 606938 |
| BCCIP | chr10 | 127502093 | 127532254 | BRCA2 and CDKN1A-interacting protein isoform | 56647 | 611883 |
| DHX32 | chr10 | 127514898 | 127559874 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 32 | 55760 | 607960 |
| FANK1 | chr10 | 127575097 | 127688151 | fibronectin type III and ankyrin repeat domains | 92565 | 611640 |
| ADAM12 | chr10 | 127693414 | 128067055 | ADAM metallopeptidase domain 12 isoform 1 | 8038 | 602714 |
| RASSF10 | chr11 | 12987271 | 12989223 | Ras association (RalGDS/AF-6) domain family | 644943 | 0 |
| SPON1 | chr11 | 13940489 | 14246222 | spondin 1, extracellular matrix protein | 10418 | 604989 |
| RRAS2 | chr11 | 14256041 | 14336607 | related RAS viral (r-ras) oncogene homolog 2 | 22800 | 600098 |
| COPB1 | chr11 | 14435628 | 14477974 | coatomer protein complex, subunit beta | 1315 | 600959 |
| PSMA1 | chr11 | 14482998 | 14498524 | proteasome alpha 1 subunit isoform 2 | 5682 | 602854 |
| PDE3B | chr11 | 14621906 | 14848926 | phosphodiesterase 3B, cGMP-inhibited | 5140 | 602047 |
| CYP2R1 | chr11 | 14856132 | 14870327 | cytochrome P450, family 2, subfamily R, | 120227 | 608713 |
| CALCA | chr11 | 14944791 | 14950408 | calcitonin isoform CGRP preproprotein | 796 | 114130 |
| CALCB | chr11 | 15051721 | 15056753 | calcitonin-related polypeptide, beta | 797 | 114160 |
| INSC | chr11 | 15090545 | 15225328 | inscuteable isoform a | 387755 | 610668 |
| SOX6 | chr11 | 15948370 | 16380968 | SRY (sex determining region Y)-box 6 isoform 1 | 55553 | 607257 |
| C11orf58 | chr11 | 16716769 | 16734149 | small acidic protein | 10944 | 0 |
| PLEKHA7 | chr11 | 16765787 | 16992535 | pleckstrin homology domain containing, family A | 144100 | 0 |

**Table 15 .Three patients shared ROHs gene annotation.** The list of genes associated to ROHs was obtained using the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables) and by querying the hg17 database.

## 4.ABCC8 and KCNJ11 genes mutational screening

Direct sequencing analysis of complete coding region including intron-exon boundaries of ABCC8 and KCNJ11 genes was performed in 3 patients who showed an homozygous chromosomal region on 11p15.2 were the two genes are located. In 2 of these 3 probands complete sequencing was undertaken and the causative mutation was revealed; in the third patient sequencing is actually on work. In the first probands the homozygous mutation 521 C>T of KCNJ11 gene was identified, resulting in the missense mutation A174V( Fig.3), and in the second one the homozygous mutation 4787 G>A of ABCC8 gene was identified, resulting in the missense mutation G1554D (Fig.4). Sequence analysis of parent's DNA and siblings showed in both cases that they were heterozygous for the same mutation.

## 5.MCAD gene mutational screening

The complete coding region of MCAD (Ensembl Transcript ID ENST00000370841), including intron-exon boundaries, was sequenced in the proband bearing an homozygous trait in 1p32.3 - 1p31.1 and in her parents and brother too . The homozygous synonymous mutation P195P in exon 8 of MCAD gene was identified (587 T >G) while her parents and brother evidenced the same etherozygous nucleotidic variation (Fig.5).As synonimous variation was les tha 20 nucleotides from intron-exon junction, we investigated if this polimorphism would produce any change in the gene splicing process . Using NNSPLICE 0.9 software, we predicted that this single base change could exclud the canonical splicing acceptor site and could create a new splicing acceptor site ( Fig.6 ).No difference were underlined in the donor site prediction .We need patient' RNA to confirm the presence of altered splicing process.

A. aagGCTATTGTGTAACAGAACCTGGAGCAGGCTCTGATGTAGCTGGTATAAAGACCAAAGCAGA

B aagGCTATTGTGTAACAGAACCGGGAGCAGGCTCTGATGTAGCTGGTATAAAGACCAAAGCAGA

**Fig 6.P195P MCAD ex 8 splice site prediction (NNSPLICE 0.9 software)** A. Wt MCAD ex 8 acceptor splice site is underlined in yellow and the wild type 587 position in red ( intron 7 in black and exon 8 in blu). B. P195P MCAD ex 8 acceptor splice site is underlined in green and the 587 T>G variation in red ( intron 7 in black and exon 8 in blu).

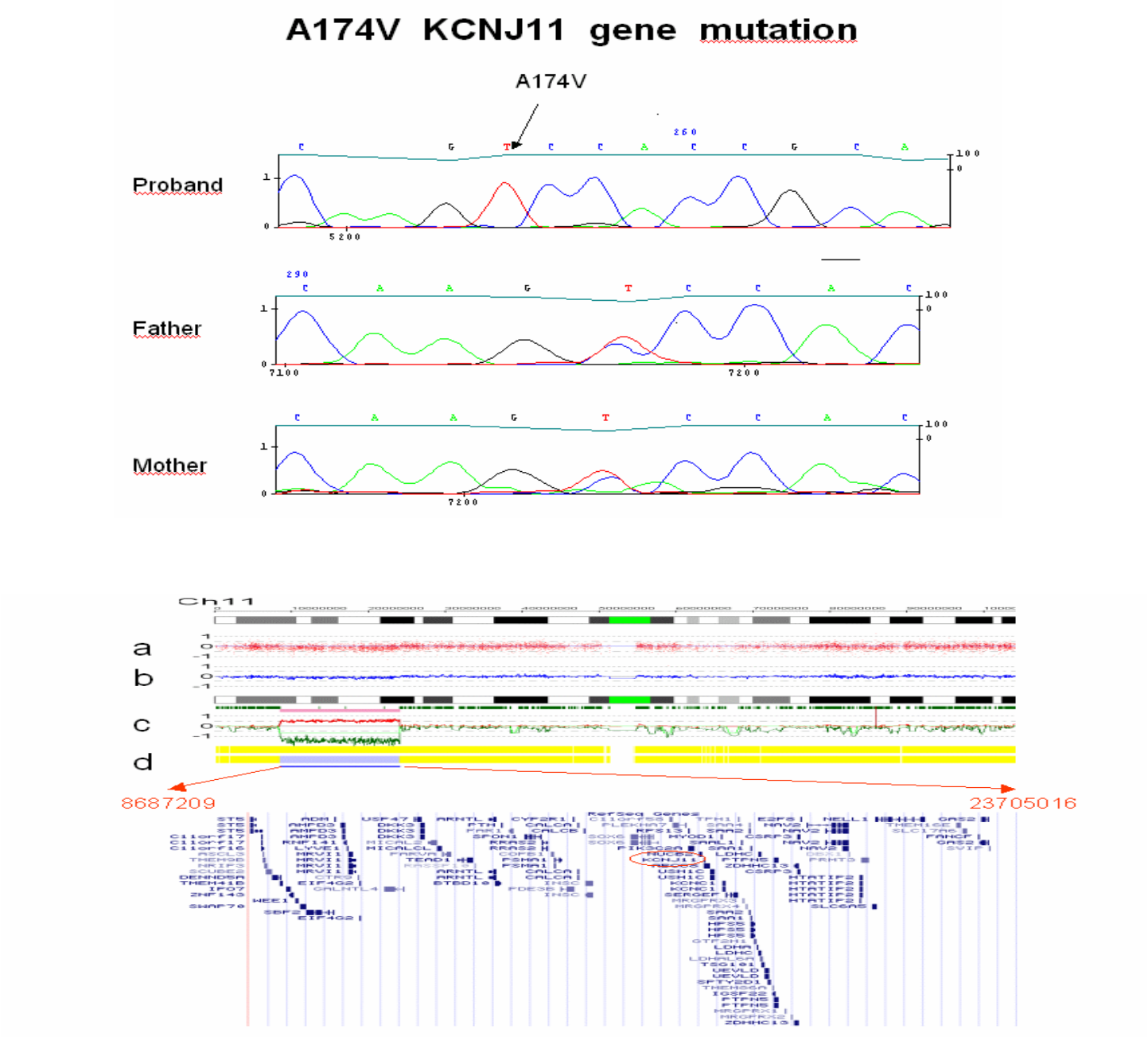



**Fig3. *KNJ11*gene mutation in CHI family**. Genomic DNA from the affected subject shows a homozygous A174V mutation of the synthesized protein. Unaffected parents are heterozygous for the same mutation. Below, map of DNA analysis of patient's chromosome 1 is shown. DNA was analysed on Affymetrix GeneChip 250K Nsp SNP Mapping Arrays. Whole-genome DNA profile was assembled using Copy Number Analyzer for Affymetrix GeneChip (CNAG, v3) software and comparing each sample to a pool of normal controls (48 HapMap samples available on Affymetrix website). Chromosome 11 map is shown from p to q end (from left to right). **A.** Red dots represent single SNP copy number signals on log scale, **B.** Total copy number values averaged on adjacent 10 SNPs (blue lines), **C.** Copy number values for each allele (red and green lines); bars in the middle represent heterozygous genotype calls (green) and homozygous calls (pink) between the sample and normal controls, **D**. The two bars at the bottom represent the colour-coded visualisation of total copy number status (yellow, diploidy) and LOH (loss of heterozygosity);(blue, significant LOH; yellow, no LOH). Copy Number Neutral LOH stretch from 8687208 to 23705016 bp according to UCS Genome Browser, Human Assembly May 2004) observed in CHI affected patient, containing the already known KCNJ11 CHI causative gene.

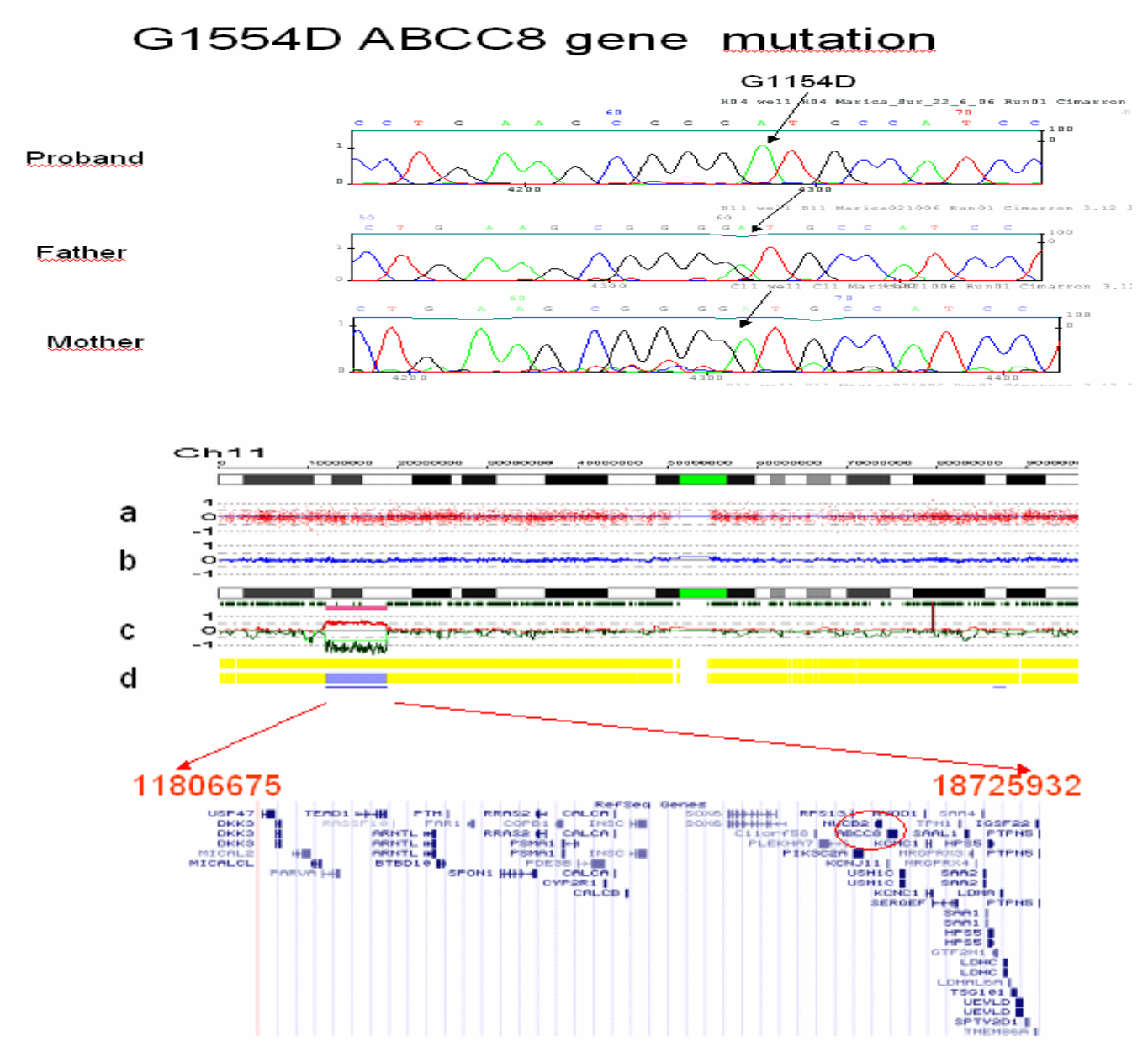



**Fig.4. *ABCC8* gene mutation in CHI family**. Genomic DNA from the affected subject shows a homozygous G1554D mutation of the synthesized protein. Unaffected parents are heterozygous for the same mutation. Below, map of DNA analysis of patient's chromosome 1 is shown. DNA was analysed on Affymetrix GeneChip 250K Nsp SNP Mapping Arrays. Whole-genome DNA profile was assembled using Copy Number Analyzer for Affymetrix GeneChip (CNAG, v3) software and comparing each sample to a pool of normal controls (48 HapMap samples available on Affymetrix website). Chromosome 11 map is shown from p to q end (from left to right). **A.** Red dots represent single SNP copy number signals on log scale, **B.** Total copy number values averaged on adjacent 10 SNPs (blue lines), **C.** Copy number values for each allele (red and green lines); bars in the middle represent heterozygous genotype calls (green) and homozygous calls (pink) between the sample and normal controls, **D**. The two bars at the bottom represent the colour-coded visualisation of total copy number status (yellow, diploidy) and LOH (loss of heterozygosity);(blue, significant LOH; yellow, no LOH). Copy Number Neutral LOH stretch from 11806675 to 18725932 bp according to UCS Genome Browser, Human Assembly May 2004) observed in CHI affected patient, containing the already known ABCC8 CHI causative gene (entrez gene ID) is shown.

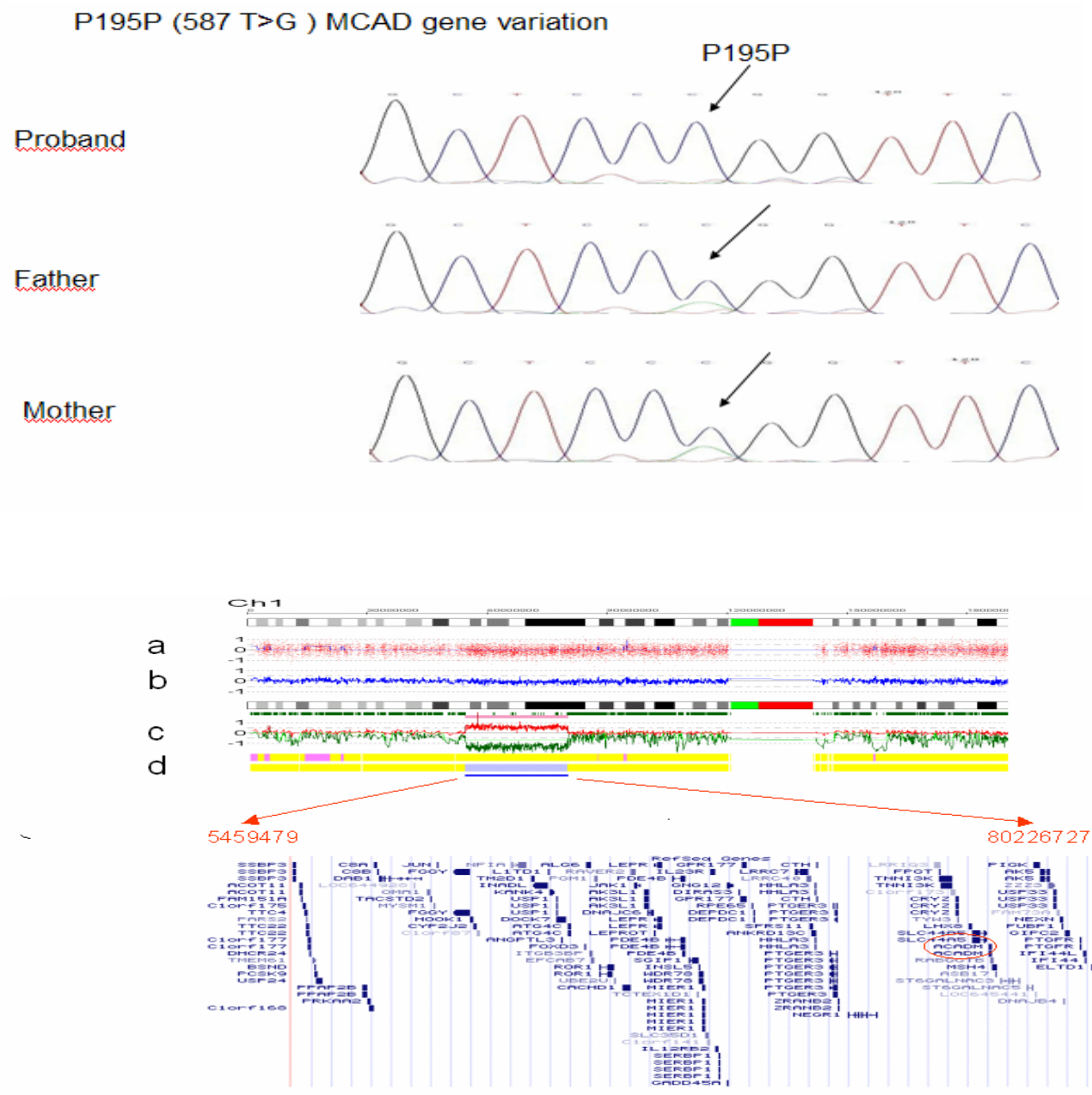
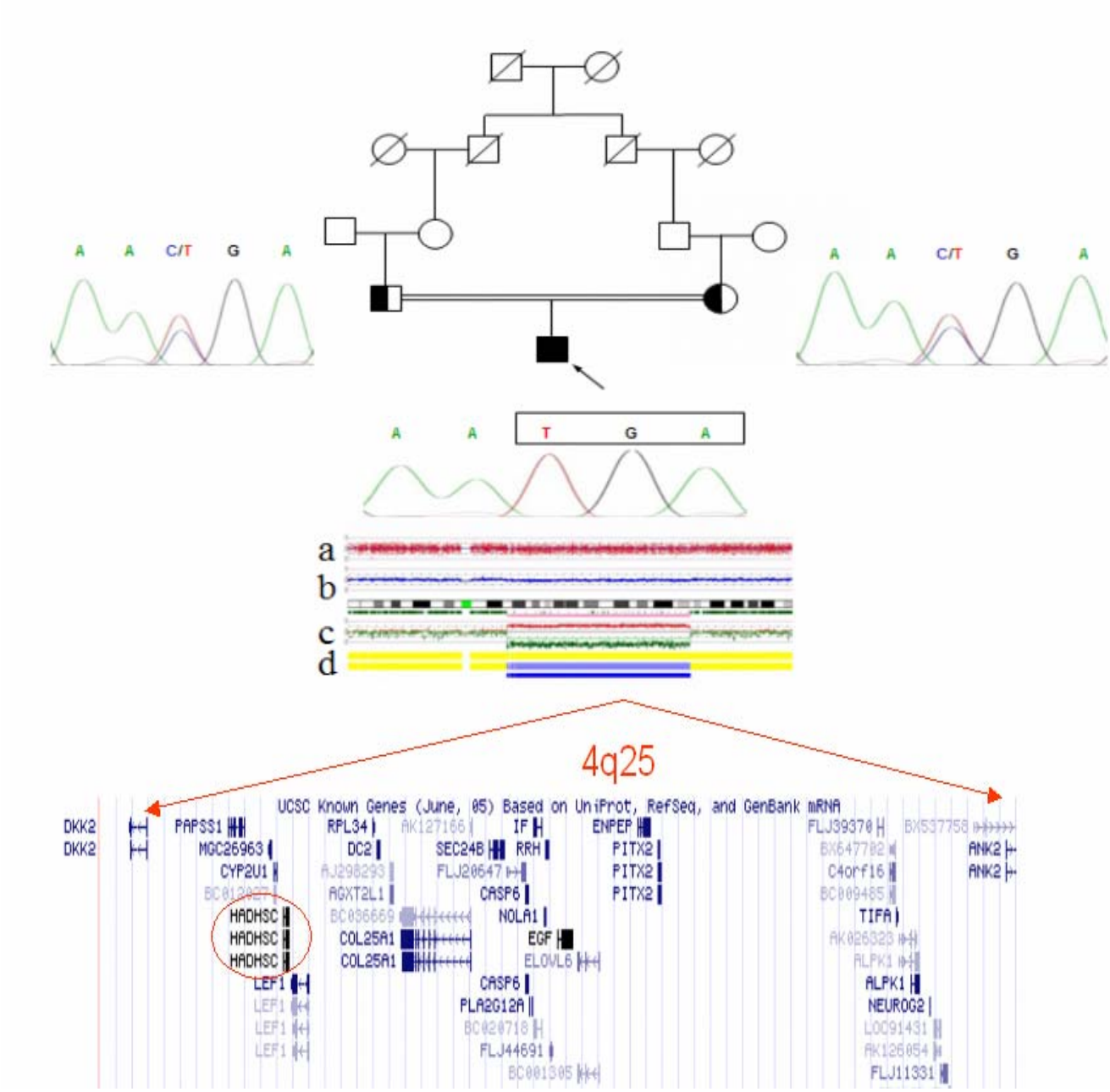



**Fig.5. *MCAD* gene variation in CHI family**. Genomic DNA from the affected subject shows a homozygous T to G transition at position 587 (P195P of the synthesized protein). Unaffected parents are heterozygous for the same mutation. Below, map of DNA analysis of patient's chromosome 1 is shown. DNA was analysed on Affymetrix GeneChip 250K Nsp SNP Mapping Arrays. Whole-genome DNA profile was assembled using Copy Number Analyzer for Affymetrix GeneChip (CNAG, v3) software and comparing each sample to a pool of normal controls (48 HapMap samples available on Affymetrix website). Chromosome 1 map is shown from p to q end (from left to right). **A.** Red dots represent single SNP copy number signals on log scale, **B.** Total copy number values averaged on adjacent 10 SNPs (blue lines), **C.** Copy number values for each allele (red and green lines); bars in the middle represent heterozygous genotype calls (green) and homozygous calls (pink) between the sample and normal controls, **D**. The two bars at the bottom represent the colour-coded visualisation of total copy number status (yellow, diploidy) and LOH (loss of heterozygosity);(blue, significant LOH; yellow, no LOH). Copy Number Neutral LOH stretch from 5459479 to80226727 bp according to UCS Genome Browser, Human Assembly May 2004) observed in CHI affected patient, containing the already known MCAD CHI causative gene .




**F ig.7.Pedigree and HADH gene mutation in CHI family**. Genomic DNA from the affected subject shows a homozygous C to T transition changing Arg to STOP codon at position 236 of the synthesized protein (R236X). Unaffected parents are heterozygous for the same mutation. Below, map of DNA analysis of patient's chromosome 4 is shown. DNA was analysed on Affymetrix GeneChip 250K Nsp SNP Mapping Arrays. Whole-genome DNA profile was assembled using Copy Number Analyzer for Affymetrix GeneChip (CNAG, v3) software and comparing each sample to a pool of normal controls (48 HapMap samples available on Affymetrix website). Chromosome 4 map is shown from p to q end (from left to right). **A.** Red dots represent single SNP copy number signals on log scale, **B.** Total copy number values averaged on adjacent 10 SNPs (blue lines), **C.** Copy number values for each allele (red and green lines); bars in the middle represent heterozygous genotype calls (green) and homozygous calls (pink) between the sample and normal controls, **D**. The two bars at the bottom represent the colour-coded visualisation of total copy number status (yellow, diploidy) and LOH (loss of heterozygosity);(blue, significant LOH; yellow, no LOH). Copy Number Neutral LOH stretch at 4q13.2-q31.21 (from 107900001 to 114100000 bp according to UCS Genome Browser, Human Assembly May 2004) observed in CHI affected patient, containing the already known CHI causative gene HADH (entrez gene ID 3033) is shown..

## 6. HADH gene mutational screening

The complete coding region of HADH gene (Ensembl Transcript ID ENST00000309522), including intron-exon boundaries, was sequenced in the proband bearing homozygosity trait in 4q13.2 to 4q31.21 and in his parents.In this patient a novel homozygous mutation was identified at codon 236 (706 C>T) in exon 6 of HADH gene, resulting in the nonsense mutation R236X. The truncated protein lacks of 78 aminiacids. Parents DNA sequencing showed they were heterozygous for this mutation (Fig.7). Clustal 2.0.5 multiple alignment visualized the 235 amminoacid which are in common between R236X truncated protein and wild type SCHAD protein (Fig.8).

We further investigated HADH gene cDNA expression in various tissues and in patient's and his mother blood linphocytes and healthy controls in order to evidence if SCHAD mRNA bearing the novel mutation would be suppressed by RNA nonsense-mediated decay. Amplification of HADH gene cDNA of the etherozygous mother, controls and different tissues revealed the presence of the amplification of SCHAD protein isoform 1 , which is of 682 bp from exon 3 to exon 8 due to exon 7 skipping. In the proband we underlined the presence of two different cDNA amplification product, the first one corresponding to the 682 bp band and the other one corresponding to a 622 bp band (Fig 9). Direct sequencing of the purified 622 bp gel band evidenced exon 6 skipped SCHAD mRNA. The additional 622 bp fragment is probably determined by the increased instability of mutated RNA.(data not shown ).

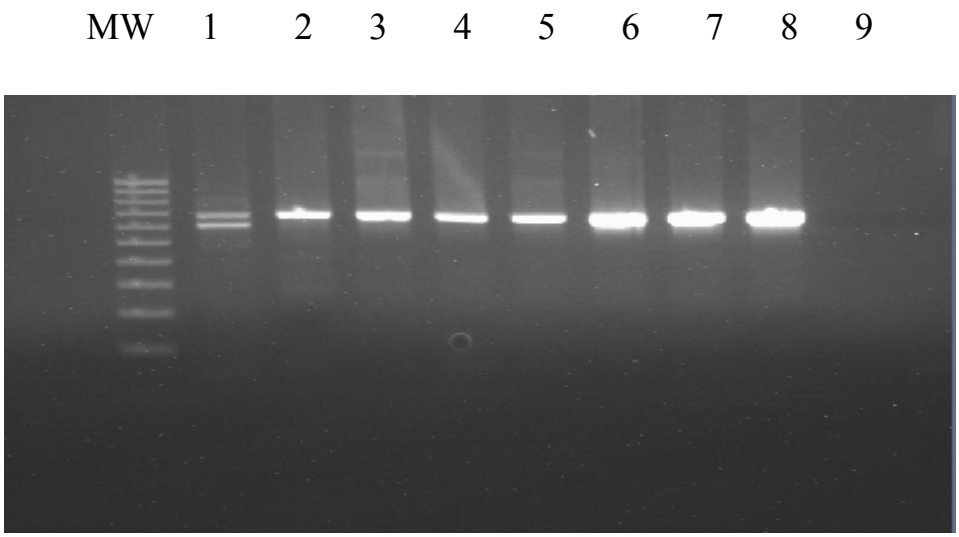



**Fig.9 SCHAD cDNA expression.** Lanes description 1. Proban 2. Mother 3.Healthy Contro1 4. Healthy Control 5. Healthy Control 6. Liver (Clonthec) 7. Skeletric Muscle (Clonthec) 8. Human Control 10. (Clonthec) 9. Blank

```
CLUSTAL 2.0.5 multiple sequence alignment

Q16836|HCDH_HUMAN       MAFVTRQFMRSVSSSSTASASAKKIIVKHVTVIGGGLMGAGIAQVAAATGHTVVLVDQTE 60
SCHAD_T                 MAFVTRQFMRSVSSSSTASASAKKIIVKHVTVIGGGLMGAGIAQVAAATGHTVVLVDQTE 60
                        ************************************************************

Q16836|HCDH_HUMAN       DILAKSKKGIEESLRKVAKKKFAENPKAGDEFVEKTLSTIATSTDAASVVHSTDLVVEAI 120
SCHAD_T                 DILAKSKKGIEESLRKVAKKKFAENPKAGDEFVEKTLSTIATSTDAASVVHSTDLVVEAI 120
                        ************************************************************

Q16836|HCDH_HUMAN       VENLKVKNELFKRLDKFAAEHTIFASNTSSLQITSIANATTRQDRFAGLHFFNPVPVMKL 180
SCHAD_T                 VENLKVKNELFKRLDKFAAEHTIFASNTSSLQITSIANATTRQDRFAGLHFFNPVPVMKL 180
                        ************************************************************

Q16836|HCDH_HUMAN       VEVIKTPMTSQKTFESLVDFSKALGKHPVSCKDTPGFIVNRLLVPYLMEAIRLYERGDAS 240
SCHAD_T                 VEVIKTPMTSQKTFESLVDFSKALGKHPVSCKDTPGFIVNRLLVPYLMEAIRLYE----- 235
                        ************************************************

Q16836|HCDH_HUMAN       KEDIDTAMKLGAGYPMGPFELLDYVGLDTTKFIVDGWHEMDAENPLHQPSPSLNKLVAEN 300
SCHAD_T                 ------------------------------------------------------------

Q16836|HCDH_HUMAN       KFGKKTGEGFYKYK 314
SCHAD_T                 -------------
```

**Fig 8.Clustal 2.0.5 multiple alignment of wild type and truncated SCHAD protein**

## 7. SCHAD protein bioinformatic analysis

ESyPred3D homology modeling ribbon diagramm and Conserved Domains obtained from Pfam database exhibited a two-domain topology of wild type SCHAD protein ( Fig.10), with the N- terminal domain of the monomer (residues 29-214) similar to a NAD(P)+ -binding enzyme and the C- terminal domain ( residues 216-313 ). Nonsense mutation R236X SCHAD protein was modeled by Pfam as a truncated protein lacking the most part of the C- terminal domain, which in this cases consisted of 20 residues instead of 105 residues .ESyPred3D homology modeling ribbon diagram of R236X protein also showed lack of subunits dimerization and no formation of the active cleft site (Fig11 )

## 8. Bioinformatic pathways analysis

In order to identify biochimical and metabolic pathways involved in CHI disease, we used DAVID Bioinformatic Database Funtional Annotation Tool to investigte a list of genes which were previously found to be present in homozygosity traits in comon between at least two patients. Our gene list contained 898 genes functionally annotated in DAVID database, 236 of which (28%) were found to be enclosed in different pathways (Fig.12) listed in the funtional annotation chart. Molecolar patways like type I diabetes mellitus, antigen processing and presentation , taste tranduction cell adhesion and communication were the more enriched cathegories, conteining from 15 to 22 genes (Fig 13), but other important cathegories related with CHI disease were also found, like glycolysis / gluconeogenesis (9 genes) and fatty acid metabolism (7 genes). Evaluating fatty acid metabolism pathway, we identified acyl-CoA dehydrogenase C4- C12 , previously described as included in the long and contigous trait on chromosome 1 , alcohol dehydrogenase class I-V, cytochrome P450 family 2 and acyl CoA synthetase long chain family member 6. Glycolysis/ gluconeogenesis pathways analysis added to this list lactate dehydrogenase A, C, and phosphofructokinase, the last one present in muscle cells.

**A L-3-Hydroxyacyl-CoA dehydrogenase (1F12)** **B. SCHAD Truncated Protein**

**Fig.10. The ribbon diagramm obtained with ESyPred3D (Prediction of proteins 3D structures) an automated homology modeling program.**

**Conserved Domain of L-3-Hydroxyacyl-CoA dehydrogenase** **Conserved Domain of Truncated Protein**



| Pfam-A | Description | Entry type | Sequence | | HMM | |
|---|---|---|---|---|---|---|
| | | | Start | End | From | To |
| 3HCDH_N | 3-hydroxyacyl-CoA dehydrogenase, NAD binding domain | Domain | 29 | 214 | 1 | 188 |
| 3HCDH | 3-hydroxyacyl-CoA dehydrogenase, C-terminal domain | Domain | 216 | 313 | 1 | 105 |

| Pfam-A | Description | Entry type | Sequence | | HMM | |
|---|---|---|---|---|---|---|
| | | | Start | End | From | To |
| 3HCDH_N | 3-hydroxyacyl-CoA dehydrogenase, NAD binding domain | Domain | 29 | 214 | 1 | 188 |
| 3HCDH | 3-hydroxyacyl-CoA dehydrogenase, C-terminal domain | Domain | 216 | 235 | 1 | 20 |

**Fig.11. The Conserved Domains obtained from Pfam database (a large collection of protein families, each represented by*multiple sequence alignments* and *hidden Markov models (HMMs)* of wild type and truncated SCHAD protein sequences.**

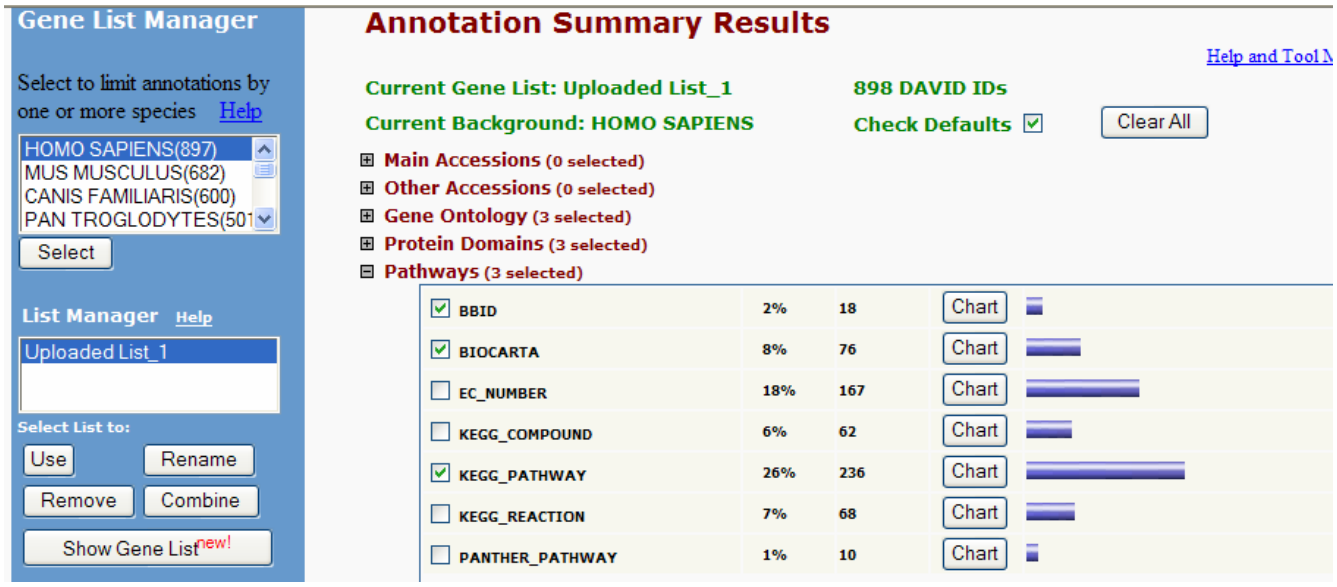
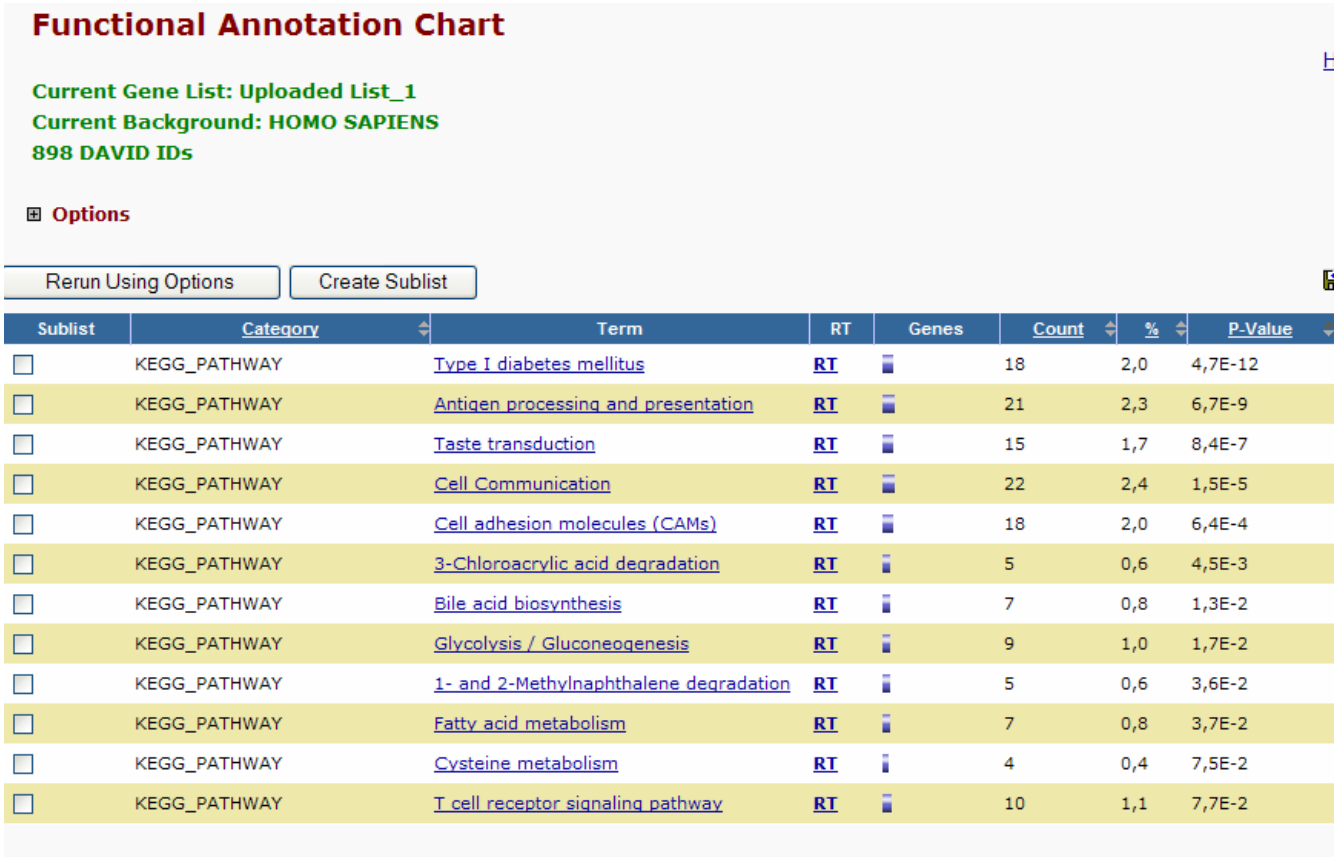



**Fig.12. DAVID Bioinformatic pathways analysis of the list of genes included in the Fingerprint ROHs:** David pathways analisys was enriched by 263 genes , that are the 26% of our gene list.



## Functional Annotation Chart

**Current Gene List: Uploaded List_1**
**Current Background: HOMO SAPIENS**
**898 DAVID IDs**

⊞ **Options**

Rerun Using Options | Create Sublist

| Sublist | Category | Term | RT | Genes | Count | % | P-Value |
|---|---|---|---|---|---|---|---|
| ☐ | KEGG_PATHWAY | Type I diabetes mellitus | RT | | 18 | 2,0 | 4,7E-12 |
| ☐ | KEGG_PATHWAY | Antigen processing and presentation | RT | | 21 | 2,3 | 6,7E-9 |
| ☐ | KEGG_PATHWAY | Taste transduction | RT | | 15 | 1,7 | 8,4E-7 |
| ☐ | KEGG_PATHWAY | Cell Communication | RT | | 22 | 2,4 | 1,5E-5 |
| ☐ | KEGG_PATHWAY | Cell adhesion molecules (CAMs) | RT | | 18 | 2,0 | 6,4E-4 |
| ☐ | KEGG_PATHWAY | 3-Chloroacrylic acid degradation | RT | | 5 | 0,6 | 4,5E-3 |
| ☐ | KEGG_PATHWAY | Bile acid biosynthesis | RT | | 7 | 0,8 | 1,3E-2 |
| ☐ | KEGG_PATHWAY | Glycolysis / Gluconeogenesis | RT | | 9 | 1,0 | 1,7E-2 |
| ☐ | KEGG_PATHWAY | 1- and 2-Methylnaphthalene degradation | RT | | 5 | 0,6 | 3,6E-2 |
| ☐ | KEGG_PATHWAY | Fatty acid metabolism | RT | | 7 | 0,8 | 3,7E-2 |
| ☐ | KEGG_PATHWAY | Cysteine metabolism | RT | | 4 | 0,4 | 7,5E-2 |
| ☐ | KEGG_PATHWAY | T cell receptor signaling pathway | RT | | 10 | 1,1 | 7,7E-2 |

**Fig.13.DAVID Bioinformatic functional annotation chart of the list of genes included in the Fingerprint ROHs:** pathways description, number of genes involved and their % in the uploaded list is indicated.

| Gene Symbol | Chromosome | Start | End | hg17.refLink.product | hg17.refLink.loc | hg17.refLink.omimId |
|---|---|---|---|---|---|---|
| VAMP3 | chr1 | 7765594 | 7775757 | vesicle-associated membrane protein 3 | 9341 | 603657 |
| PER3 | chr1 | 7779028 | 7839502 | period 3 | 8863 | 603427 |
| UTS2 | chr1 | 7841940 | 7847361 | urotensin 2 preproprotein, isoform b | 10911 | 604097 |
| TNFRSF9 | chr1 | 7914173 | 7935153 | tumor necrosis factor receptor superfamily, | 3604 | 602250 |
| PARK7 | chr1 | 7955979 | 7979608 | Parkinson disease protein 7 | 11315 | 602533 |
| ERRFI1 | chr1 | 8006060 | 8020622 | mitogen-inducible gene 6 protein | 54206 | 608069 |
| ACOT11 | chr1 | 54725921 | 54788021 | thioesterase, adipose associated isoform BFIT2 | 26027 | 606803 |
| FAM151A | chr1 | 54786870 | 54801221 | hypothetical protein LOC338094 | 338094 | 0 |
| C1orf175 | chr1 | 54819464 | 54887960 | hypothetical protein LOC374977 | 374977 | 0 |
| TTC4 | chr1 | 54893549 | 54920002 | tetratricopeptide repeat domain 4 | 7268 | 606753 |
| PARS2 | chr1 | 54934593 | 54942208 | prolyl-tRNA synthetase | 25973 | 612036 |
| TTC22 | chr1 | 54958772 | 54978962 | tetratricopeptide repeat domain 22 isoform 1 | 55001 | 0 |
| C1orf177 | chr1 | 54983756 | 55019958 | hypothetical protein LOC163747 isoform 2 | 163747 | 0 |
| DHCR24 | chr1 | 55027320 | 55064942 | 24-dehydrocholesterol reductase precursor | 1718 | 606418 |
| TMEM61 | chr1 | 55158485 | 55169986 | transmembrane protein 61 | 199964 | 0 |
| BSND | chr1 | 55176637 | 55186485 | barttin | 7809 | 606412 |
| ROR1 | chr1 | 63951710 | 64321073 | receptor tyrosine kinase-like orphan receptor 1 | 4919 | 602336 |
| UBE2U | chr1 | 64381510 | 64422048 | ubiquitin-conjugating enzyme E2U (putative) | 148581 | 0 |
| PDE4B | chr1 | 65970213 | 66552282 | phosphodiesterase 4B, cAMP-specific isoform 1 | 5142 | 600127 |
| SGIP1 | chr1 | 66711845 | 66922788 | SH3-domain GRB2-like (endophilin) interacting | 84251 | 611540 |
| TCTEX1D1 | chr1 | 66930162 | 66956491 | Tctex1 domain containing 1 | 200132 | 0 |
| INSL5 | chr1 | 66975445 | 66978960 | insulin-like 5 precursor | 10022 | 606413 |
| WDR78 | chr1 | 66990594 | 67102591 | WD repeat domain 78 isoform 1 | 79819 | 0 |
| MIER1 | chr1 | 67102662 | 67164374 | mesoderm induction early response 1 isoform | 57708 | 0 |
| SLC35D1 | chr1 | 67181873 | 67231749 | solute carrier family 35 (UDP-glucuronic | 23169 | 610804 |
| C1orf141 | chr1 | 67269879 | 67306241 | hypothetical protein LOC400757 | 400757 | 0 |
| IL23R | chr1 | 67344189 | 67437669 | interleukin 23 receptor precursor | 149233 | 607562 |
| IL12RB2 | chr1 | 67485067 | 67574603 | interleukin 12 receptor, beta 2 precursor | 3595 | 601642 |
| DIRAS3 | chr1 | 68223667 | 68228481 | DIRAS family, GTP-binding RAS-like 3 | 9077 | 605193 |
| GPR177 | chr1 | 68276172 | 68410274 | G protein-coupled receptor 177 isoform 2 | 79971 | 611514 |
| NEGR1 | chr1 | 71580646 | 72460298 | neuronal growth regulator 1 | 257194 | 0 |
| LRRIQ3 | chr1 | 74203722 | 74375892 | leucine-rich repeats and IQ motif containing 3 | 127255 | 0 |
| FPGT | chr1 | 74375946 | 74385141 | fucose-1-phosphate guanyltransferase | 8790 | 603609 |
| TNNI3K | chr1 | 74375967 | 74722129 | TNNI3 interacting kinase isoform a | 51086 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| C1orf173 | chr1 | 74745816 | 74851443 | hypothetical protein LOC127254 | 127254 | 0 |
| ACADM | chr1 | 75902063 | 75941376 | acyl-Coenzyme A dehydrogenase, C-4 to C-12 | 34 | 607008 |
| RABGGTB | chr1 | 75963906 | 75972785 | Rab geranylgeranyltransferase, beta subunit | 5876 | 179080 |
| MSH4 | chr1 | 75974650 | 76090944 | mutS homolog 4 | 4438 | 602105 |
| ASB17 | chr1 | 76096580 | 76110137 | ankyrin repeat and SOCS box-containing 17 | 127247 | 0 |
| AK5 | chr1 | 77459762 | 77737673 | adenylate kinase 5 isoform 1 | 26289 | 608009 |
| ZZZ3 | chr1 | 77742210 | 77860364 | zinc finger, ZZ-type containing 3 | 26009 | 0 |
| USP33 | chr1 | 77873696 | 77937558 | ubiquitin specific protease 33 isoform 1 | 23032 | 0 |
| FAM73A | chr1 | 77957329 | 78056102 | hypothetical protein LOC374986 | 374986 | 0 |
| NEXN | chr1 | 78066220 | 78121599 | nexilin (F actin binding protein) | 91624 | 0 |
| FUBP1 | chr1 | 78125613 | 78156798 | far upstream element-binding protein | 8880 | 603444 |
| DNAJB4 | chr1 | 78182656 | 78195014 | DnaJ (Hsp40) homolog, subfamily B, member 4 | 11080 | 611327 |
| GIPC2 | chr1 | 78223609 | 78315132 | PDZ domain protein GIPC2 | 54810 | 0 |
| RGS4 | chr1 | 159770053 | 159778250 | regulator of G-protein signaling 4 isoform 1 | 5999 | 602516 |
| RGS5 | chr1 | 159843755 | 159904530 | regulator of G-protein signalling 5 | 8490 | 603276 |
| NUF2 | chr1 | 160023380 | 160057211 | NUF2, NDC80 kinetochore complex component | 83540 | 611772 |
| RGS18 | chr1 | 188859248 | 188886600 | regulator of G-protein signalling 18 | 64407 | 607192 |
| RGS21 | chr1 | 189017778 | 189068071 | regulator of G-protein signaling 21 | 431704 | 0 |
| LTBP1 | chr2 | 33084042 | 33536224 | latent transforming growth factor beta binding | 4052 | 150390 |
| NRXN1 | chr2 | 50059138 | 50486545 | neurexin 1 isoform beta precursor | 9378 | 600565 |
| AHSA2 | chr2 | 61316471 | 61325709 | AHA1, activator of heat shock 90kDa protein | 130872 | 0 |
| USP34 | chr2 | 61326241 | 61609500 | ubiquitin specific protease 34 | 9736 | 0 |
| XPO1 | chr2 | 61616720 | 61677069 | exportin 1 | 7514 | 602559 |
| ETAA1 | chr2 | 67536092 | 67549185 | ETAA16 protein | 54465 | 0 |
| POLR1A | chr2 | 86165110 | 86244936 | polymerase (RNA) I polypeptide A, 194kDa | 25885 | 0 |
| PTCD3 | chr2 | 86244995 | 86280938 | Pentatricopeptide repeat domain 3 | 55037 | 0 |
| IMMT | chr2 | 86282712 | 86334551 | inner membrane protein, mitochondrial isoform 2 | 10989 | 600378 |
| MRPL35 | chr2 | 86338213 | 86352135 | mitochondrial ribosomal protein L35 isoform a | 51318 | 611841 |
| REEP1 | chr2 | 86352777 | 86476435 | receptor accessory protein 1 | 65055 | 609139 |
| JMJD1A | chr2 | 86580011 | 86631497 | jumonji domain containing 1A | 55818 | 611512 |
| BIN1 | chr2 | 127521836 | 127581094 | bridging integrator 1 isoform 6 | 274 | 601248 |
| CYP27C1 | chr2 | 127657641 | 127679573 | cytochrome P450, family 27, subfamily C, | 339761 | 0 |
| ERCC3 | chr2 | 127731095 | 127767982 | excision repair cross-complementing rodent | 2071 | 133510 |
| MAP3K2 | chr2 | 127778368 | 127817000 | mitogen-activated protein kinase kinase kinase | 10746 | 609487 |
| PROC | chr2 | 127892246 | 127903046 | protein C (inactivator of coagulation factors Va | 5624 | 176860 |
| IWS1 | chr2 | 127954619 | 128000274 | IWS1 homolog | 55677 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GPR39 | chr2 | 133007878 | 133237901 | G protein-coupled receptor 39 | 2863 | 602886 |
| LYPD1 | chr2 | 133236068 | 133262213 | LY6/PLAUR domain containing 1 isoform a | 116372 | 610450 |
| NAP5 | chr2 | 133263103 | 134159763 | Nck-associated protein 5 isoform 1 | 344148 | 608789 |
| LRP1B | chr2 | 140822727 | 142723002 | low density lipoprotein-related protein 1B | 53353 | 608766 |
| ZNF385B | chr2 | 180132216 | 180436275 | zinc finger protein 385B isoform 2 | 151126 | 0 |
| KIAA1604 | chr2 | 180635109 | 180697286 | hypothetical protein LOC57703 | 57703 | 0 |
| ZNF804A | chr2 | 185288598 | 185629718 | zinc finger protein 804A | 91752 | 0 |
| SH3BP4 | chr2 | 235642627 | 235746356 | SH3-domain binding protein 4 | 23677 | 605611 |
| ESPNL | chr2 | 238790950 | 238823928 | espin-like | 339768 | 0 |
| KLHL30 | chr2 | 238831416 | 238843547 | kelch-like 30 | 377007 | 0 |
| ILKAP | chr2 | 238861043 | 238894324 | integrin-linked kinase-associated protein | 80895 | 0 |
| HES6 | chr2 | 238928909 | 238930670 | hairy and enhancer of split 6 | 55502 | 610331 |
| PER2 | chr2 | 238934678 | 238979207 | period 2 | 8864 | 603426 |
| TRAF3IP1 | chr2 | 239011081 | 239089796 | TNF receptor-associated factor 3 interacting | 26146 | 607380 |
| ASB1 | chr2 | 239117625 | 239142890 | ankyrin repeat and SOCS box-containing protein | 51665 | 605758 |
| UBE2E2 | chr3 | 23219811 | 23607074 | ubiquitin-conjugating enzyme E2E 2 (UBC4/5 | 7325 | 602163 |
| UBE2E1 | chr3 | 23822442 | 23907806 | ubiquitin-conjugating enzyme E2E 1 isoform 1 | 7324 | 602916 |
| NKIRAS1 | chr3 | 23908575 | 23933541 | kappa B-ras 1 | 28512 | 604496 |
| RPL15 | chr3 | 23933642 | 23937334 | ribosomal protein L15 | 6138 | 604174 |
| NR1D2 | chr3 | 23961809 | 23996241 | nuclear receptor subfamily 1, group D, member 2 | 9975 | 602304 |
| THRB | chr3 | 24133648 | 24511317 | thyroid hormone receptor, beta | 7068 | 190160 |
| CACNA2D3 | chr3 | 54131732 | 55083622 | calcium channel, voltage-dependent, alpha | 55799 | 606399 |
| LRTM1 | chr3 | 54927424 | 54937112 | leucine-rich repeats and transmembrane domains | 57408 | 0 |
| SUCLG2 | chr3 | 67507832 | 67787728 | succinate-CoA ligase, GDP-forming, beta subunit | 8801 | 603922 |
| FAM19A1 | chr3 | 68136143 | 68677459 | family with sequence similarity 19 (chemokine | 407738 | 0 |
| HEL308 | chr4 | 84685679 | 84734204 | DNA helicase HEL308 | 113510 | 606769 |
| MRPS18C | chr4 | 84734296 | 84740108 | mitochondrial ribosomal protein S18C | 51023 | 611983 |
| FAM175A | chr4 | 84739273 | 84763469 | coiled-coil domain containing 98 | 84142 | 611143 |
| AGPAT9 | chr4 | 84814831 | 84884204 | lysophosphatidic acid acyltransferase theta | 84803 | 610958 |
| ARHGAP24 | chr4 | 86753462 | 87281001 | Rho GTPase activating protein 24 isoform 1 | 83478 | 610586 |
| MAPK10 | chr4 | 87294811 | 87731462 | mitogen-activated protein kinase 10 isoform 1 | 5602 | 602897 |
| PTPN13 | chr4 | 87873063 | 88093481 | protein tyrosine phosphatase, non-receptor type | 5783 | 600267 |
| SLC10A6 | chr4 | 88102019 | 88127447 | sodium-dependent organic anion transporter | 345274 | 0 |
| C4orf36 | chr4 | 88154536 | 88170754 | hypothetical protein LOC132989 | 132989 | 0 |
| AFF1 | chr4 | 88285331 | 88419370 | myeloid/lymphoid or mixed-lineage leukemia | 4299 | 159557 |
| KLHL8 | chr4 | 88439392 | 88498853 | kelch-like 8 | 57563 | 611967 |

| | | | | | | |
|---|---|---|---|---|---|---|
| HSD17B13 | chr4 | 88582119 | 88601217 | hydroxysteroid (17-beta) dehydrogenase 13 | 345275 | 612127 |
| HSD17B11 | chr4 | 88614869 | 88669634 | dehydrogenase/reductase (SDR family) member 8 | 51170 | 0 |
| NUDT9 | chr4 | 88700913 | 88736676 | nudix -type motif 9 isoform a | 53343 | 606022 |
| SPARCL1 | chr4 | 88751666 | 88807834 | SPARC-like 1 | 8404 | 606041 |
| DSPP | chr4 | 88886859 | 88895202 | dentin sialophosphoprotein preproprotein | 1834 | 125485 |
| DMP1 | chr4 | 88928637 | 88942689 | dentin matrix acidic phosphoprotein isoform 2 | 1758 | 600980 |
| IBSP | chr4 | 89077880 | 89090780 | integrin-binding sialoprotein precursor | 3381 | 147563 |
| MEPE | chr4 | 89111318 | 89125121 | matrix, extracellular phosphoglycoprotein with | 56955 | 605912 |
| SPP1 | chr4 | 89253980 | 89261741 | secreted phosphoprotein 1 isoform b | 6696 | 166490 |
| PKD2 | chr4 | 89285998 | 89356107 | polycystin 2 | 5311 | 173910 |
| ABCG2 | chr4 | 89368595 | 89437190 | ATP-binding cassette, sub-family G, member 2 | 9429 | 603756 |
| PPM1K | chr4 | 89538710 | 89563067 | protein phosphatase 1K (PP2C domain containing) | 152926 | 611065 |
| HERC6 | chr4 | 89657069 | 89721426 | hect domain and RLD 6 | 55008 | 609249 |
| HERC5 | chr4 | 89735445 | 89784483 | hect domain and RLD 5 | 51191 | 608242 |
| PIGY | chr4 | 89799312 | 89802133 | phosphatidylinositol glycan anchor biosynthesis, | 84992 | 610662 |
| HERC3 | chr4 | 89870824 | 89986864 | hect domain and RLD 3 | 8916 | 605200 |
| NAP1L5 | chr4 | 89974244 | 89976158 | nucleosome assembly protein 1-like 5 | 266812 | 0 |
| FAM13A1 | chr4 | 90004283 | 90101580 | family with sequence similarity 13, member A1 | 10144 | 0 |
| GRID2 | chr4 | 93582727 | 95050827 | glutamate receptor, ionotropic, delta 2 | 2895 | 602368 |
| ATOH1 | chr4 | 95107255 | 95108320 | atonal homolog 1 | 474 | 601461 |
| SMARCAD1 | chr4 | 95485936 | 95569621 | SWI/SNF-related, matrix-associated | 56916 | 0 |
| PGDS | chr4 | 95576884 | 95621205 | prostaglandin-D synthase | 27306 | 602598 |
| C4orf37 | chr4 | 98837211 | 99421569 | hypothetical protein LOC285555 | 285555 | 0 |
| ADH5 | chr4 | 100349307 | 100367109 | class III alcohol dehydrogenase 5 chi subunit | 128 | 103710 |
| ADH4 | chr4 | 100402012 | 100422627 | class II alcohol dehydrogenase 4 pi subunit | 127 | 103740 |
| ADH6 | chr4 | 100480972 | 100497581 | class V alcohol dehydrogenase 6 isoform 1 | 130 | 103735 |
| ADH1A | chr4 | 100554701 | 100569320 | class I alcohol dehydrogenase, alpha subunit | 124 | 103700 |
| ADH1B | chr4 | 100584704 | 100599750 | alcohol dehydrogenase 1B (class I), beta | 125 | 103720 |
| ADH1C | chr4 | 100614828 | 100631095 | class I alcohol dehydrogenase, gamma subunit | 126 | 103730 |
| ADH7 | chr4 | 100690595 | 100713703 | class IV alcohol dehydrogenase 7 mu or sigma | 131 | 600086 |
| PPP3CA | chr4 | 102301764 | 102625531 | protein phosphatase 3, catalytic subunit, alpha | 5530 | 114105 |
| BANK1 | chr4 | 103068941 | 103353147 | B-cell scaffold protein with ankyrin repeats 1 | 55024 | 610292 |
| SLC39A8 | chr4 | 103539998 | 103623526 | solute carrier family 39 (zinc transporter), | 64116 | 608732 |
| NFKB1 | chr4 | 103779672 | 103895661 | nuclear factor kappa-B, subunit 1 | 4790 | 164011 |
| MANBA | chr4 | 103909846 | 104039351 | mannosidase, beta A, lysosomal | 4126 | 609489 |
| UBE2D3 | chr4 | 104074372 | 104147599 | ubiquitin-conjugating enzyme E2D 3 isoform 3 | 7323 | 602963 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CISD2 | chr4 | 104147730 | 104167317 | CDGSH iron sulfur domain 2 | 493856 | 611507 |
| NHEDC1 | chr4 | 104163798 | 104298480 | Na+/H+ exchanger domain containing 1 isoform 2 | 150159 | 611527 |
| NHEDC2 | chr4 | 104304954 | 104355534 | Na+/H+ exchanger domain containing 2 | 133308 | 611789 |
| BDH2 | chr4 | 104356386 | 104378628 | 3-hydroxybutyrate dehydrogenase, type 2 | 56898 | 0 |
| CENPE | chr4 | 104384566 | 104477170 | centromere protein E | 1062 | 117143 |
| TACR3 | chr4 | 104868228 | 104998577 | tachykinin receptor 3 | 6870 | 162332 |
| CXXC4 | chr4 | 105750948 | 105770071 | CXXC finger 4 | 80319 | 611645 |
| ARSJ | chr4 | 115179043 | 115258482 | arylsulfatase J | 79642 | 610010 |
| TRAM1L1 | chr4 | 118362312 | 118364339 | translocation associated membrane protein 1-like | 133022 | 0 |
| KIAA1627 | chr4 | 119964176 | 119989678 | hypothetical protein LOC57721 | 57721 | 0 |
| SEC24D | chr4 | 120001580 | 120114929 | Sec24-related protein D | 9871 | 607186 |
| SYNPO2 | chr4 | 120167598 | 120317272 | synaptopodin 2 isoform b | 171024 | 0 |
| MYOZ2 | chr4 | 120414541 | 120466540 | myozenin 2 | 51778 | 605602 |
| USP53 | chr4 | 120491384 | 120574276 | ubiquitin specific protease 53 | 54532 | 0 |
| LOC401152 | chr4 | 120575176 | 120579571 | hypothetical protein LOC401152 | 401152 | 0 |
| FABP2 | chr4 | 120596007 | 120600919 | intestinal fatty acid binding protein 2 | 2169 | 134640 |
| IL21 | chr4 | 123891387 | 123899816 | interleukin 21 | 59067 | 605384 |
| BBS12 | chr4 | 124011461 | 124023702 | Bardet-Biedl syndrome 12 | 166379 | 610683 |
| FGF2 | chr4 | 124105467 | 124176995 | fibroblast growth factor 2 | 2247 | 134920 |
| NUDT6 | chr4 | 124171410 | 124201728 | nudix-type motif 6 isoform b | 11162 | 606261 |
| SPATA5 | chr4 | 124201829 | 124598209 | spermatogenesis associated 5 | 166378 | 0 |
| SPRY1 | chr4 | 124675560 | 124682512 | sprouty homolog 1, antagonist of FGF signaling | 10252 | 602465 |
| ANKRD50 | chr4 | 125943072 | 125989537 | ankyrin repeat domain 50 | 57182 | 0 |
| FAT4 | chr4 | 126595171 | 126771692 | FAT tumor suppressor homolog 4 | 79633 | 0 |
| CCRN4L | chr4 | 140294547 | 140324696 | CCR4 carbon catabolite repression 4-like | 25819 | 608468 |
| ELF2 | chr4 | 140336475 | 140363252 | E74-like factor 2 (ets domain transcription | 1998 | 0 |
| OSAP | chr4 | 140544921 | 140559097 | ovary-specific acidic protein | 84709 | 0 |
| NDUFC1 | chr4 | 140568695 | 140574562 | NADH dehydrogenase (ubiquinone) 1, subcomplex | 4717 | 603844 |
| NARG1 | chr4 | 140580280 | 140669540 | NMDA receptor regulated 1 | 80155 | 608000 |
| CDH18 | chr5 | 19508913 | 20017046 | cadherin 18, type 2 preproprotein | 1016 | 603019 |
| COMMD10 | chr5 | 115448625 | 115656877 | COMM domain containing 10 | 51397 | 0 |
| SLC27A6 | chr5 | 128329108 | 128397233 | solute carrier family 27 (fatty acid | 28965 | 604196 |
| ISOC1 | chr5 | 128458340 | 128477616 | isochorismatase domain containing 1 | 51015 | 0 |
| ACSL6 | chr5 | 131317050 | 131375248 | acyl-CoA synthetase long-chain family member 6 | 23305 | 604443 |
| IL3 | chr5 | 131424245 | 131426795 | interleukin 3 precursor | 3562 | 147740 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CSF2 | chr5 | 131437383 | 131439757 | colony stimulating factor 2 precursor | 1437 | 138960 |
| P4HA2 | chr5 | 131556204 | 131590834 | prolyl 4-hydroxylase, alpha II subunit isoform 1 | 8974 | 600608 |
| PDLIM4 | chr5 | 131621285 | 131637044 | PDZ and LIM domain 4 | 8572 | 603422 |
| SLC22A4 | chr5 | 131658043 | 131707796 | solute carrier family 22 member 4 | 6583 | 604190 |
| SLC22A5 | chr5 | 131733299 | 131759205 | solute carrier family 22 member 5 | 6584 | 603377 |
| LOC441108 | chr5 | 131774571 | 131825958 | hypothetical protein LOC441108 | 441108 | 0 |
| IRF1 | chr5 | 131846683 | 131854326 | interferon regulatory factor 1 | 3659 | 147575 |
| IL5 | chr5 | 131905034 | 131907113 | interleukin 5 precursor | 3567 | 147850 |
| RAD50 | chr5 | 131920528 | 132007492 | RAD50 homolog isoform 2 | 10111 | 604040 |
| CTNNA1 | chr5 | 138117005 | 138298621 | catenin, alpha 1 | 1495 | 116805 |
| SIL1 | chr5 | 138310310 | 138561964 | SIL1 protein precursor | 64374 | 608005 |
| MATR3 | chr5 | 138637690 | 138694027 | matrin 3 | 9782 | 164015 |
| PAIP2 | chr5 | 138705417 | 138733307 | poly(A) binding protein interacting protein 2 | 51247 | 605604 |
| SLC23A1 | chr5 | 138730789 | 138746900 | solute carrier family 23 (nucleobase | 9963 | 603790 |
| MGC29506 | chr5 | 138751156 | 138753504 | hypothetical protein LOC51237 | 51237 | 609447 |
| DNAJC18 | chr5 | 138775279 | 138803038 | DnaJ (Hsp40) homolog, subfamily C, member 18 | 202052 | 0 |
| ECSM2 | chr5 | 138812143 | 138822504 | endothelial cell-specific molecule 2 | 641700 | 0 |
| TMEM173 | chr5 | 138835733 | 138842476 | transmembrane protein 173 | 340061 | 0 |
| UBE2D2 | chr5 | 138920934 | 138988200 | ubiquitin-conjugating enzyme E2D 2 isoform 1 | 7322 | 602962 |
| CXXC5 | chr5 | 139008484 | 139042864 | CXXC finger 5 | 51523 | 0 |
| PSD2 | chr5 | 139155589 | 139204232 | pleckstrin and Sec7 domain containing 2 | 84249 | 0 |
| NRG2 | chr5 | 139207443 | 139403063 | neuregulin 2 isoform 4 | 9542 | 603818 |
| DPYSL3 | chr5 | 146750565 | 146813453 | dihydropyrimidinase-like 3 | 1809 | 601168 |
| JAKMIP2 | chr5 | 146950898 | 147142445 | janus kinase and microtubule interacting protein | 9832 | 611197 |
| SPINK1 | chr5 | 147184338 | 147191453 | serine protease inhibitor, Kazal type 1 | 6690 | 167790 |
| SCGB3A2 | chr5 | 147238466 | 147241946 | secretoglobin, family 3A, member 2 | 117156 | 606531 |
| MGC23985 | chr5 | 147252463 | 147266294 | hypothetical protein LOC389336 | 389336 | 0 |
| SPINK5 | chr5 | 147423727 | 147484696 | serine peptidase inhibitor, Kazal type 5 isoform | 11005 | 605010 |
| SPINK5L2 | chr5 | 147529488 | 147535154 | Kazal type serine protease inhibitor 5-like 2 | 408187 | 0 |
| SLC36A3 | chr5 | 150636515 | 150663520 | solute carrier family 36 (proton/amino acid | 285641 | 608332 |
| SLC36A2 | chr5 | 150674732 | 150707312 | solute carrier family 36 (proton/amino acid | 153201 | 608331 |
| SLC36A1 | chr5 | 150807355 | 150852132 | solute carrier family 36 member 1 | 206358 | 606561 |
| FAT2 | chr5 | 150863845 | 150928698 | FAT tumor suppressor 2 precursor | 2196 | 604269 |
| SPARC | chr5 | 151021206 | 151046710 | secreted protein, acidic, cysteine-rich | 6678 | 182120 |
| ATOX1 | chr5 | 151102576 | 151118403 | antioxidant protein 1 | 475 | 602270 |
| G3BP1 | chr5 | 151131668 | 151165106 | Ras-GTPase-activating protein SH3-domain- | 10146 | 608431 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | binding | | |
| GLRA1 | chr5 | 151182360 | 151284596 | glycine receptor, alpha 1 | 2741 | 138491 |
| GRIA1 | chr5 | 152850276 | 153173622 | glutamate receptor, ionotropic, AMPA 1 isoform | 2890 | 138248 |
| SGCD | chr5 | 155686344 | 156127376 | delta-sarcoglycan isoform 1 | 6444 | 601411 |
| TIMD4 | chr5 | 156278948 | 156322844 | T-cell immunoglobulin and mucin domain | 91937 | 610096 |
| HAVCR1 | chr5 | 156389108 | 156418065 | hepatitis A virus cellular receptor 1 | 26762 | 606518 |
| HAVCR2 | chr5 | 156445420 | 156468716 | T cell immunoglobulin mucin 3 | 84868 | 606652 |
| MED7 | chr5 | 156498028 | 156502364 | mediator complex subunit 7 | 9443 | 605045 |
| FAM71B | chr5 | 156521928 | 156525853 | hypothetical protein LOC153745 | 153745 | 0 |
| ITK | chr5 | 156540484 | 156614687 | IL2-inducible T-cell kinase | 3702 | 186973 |
| CYFIP2 | chr5 | 156625668 | 156755182 | cytoplasmic FMR1 interacting protein 2 | 26999 | 606323 |
| C5orf40 | chr5 | 156701184 | 156705307 | hypothetical protein LOC408263 | 408263 | 0 |
| ICHTHYIN | chr5 | 156819604 | 156834308 | ichthyin protein | 348938 | 609383 |
| ADAM19 | chr5 | 156836890 | 156935346 | ADAM metallopeptidase domain 19 isoform 2 | 8728 | 603640 |
| SOX30 | chr5 | 156985264 | 157012006 | SRY (sex determining region Y)-box 30 isoform b | 11063 | 606698 |
| THG1L | chr5 | 157090900 | 157099349 | interphase cytoplasmic foci protein 45 | 54974 | 0 |
| LSM11 | chr5 | 157103332 | 157116324 | LSM11, U7 small nuclear RNA associated | 134353 | 0 |
| CLINT1 | chr5 | 157145874 | 157218746 | epsin 4 | 9685 | 607265 |
| PHACTR1 | chr6 | 12825818 | 13395507 | phosphatase and actin regulator 1 | 221692 | 608723 |
| CAP2 | chr6 | 17501714 | 17666000 | adenylyl cyclase-associated protein 2 | 10486 | 0 |
| FAM8A1 | chr6 | 17708564 | 17719928 | family with sequence similarity 8, member A1 | 51439 | 0 |
| NUP153 | chr6 | 17723247 | 17814797 | nucleoporin 153kDa | 9972 | 603948 |
| KIF13A | chr6 | 17868563 | 18095778 | kinesin family member 13A isoform d | 63971 | 605433 |
| NHLRC1 | chr6 | 18228696 | 18230830 | NHL repeat containing 1 | 378884 | 608072 |
| TPMT | chr6 | 18236525 | 18263353 | thiopurine S-methyltransferase | 7172 | 187680 |
| AOF1 | chr6 | 18263597 | 18332063 | amine oxidase (flavin containing) domain 1 | 221656 | 0 |
| DEK | chr6 | 18332380 | 18372778 | DEK oncogene | 7913 | 125264 |
| RNF144B | chr6 | 18495572 | 18576827 | IBR domain containing 2 | 255488 | 0 |
| ID4 | chr6 | 19945595 | 19948893 | inhibitor of DNA binding 4, dominant negative | 3400 | 600581 |
| MBOAT1 | chr6 | 20208913 | 20320649 | membrane bound O-acyltransferase domain | 154141 | 611732 |
| E2F3 | chr6 | 20510115 | 20601924 | E2F transcription factor 3 | 1871 | 600427 |
| CDKAL1 | chr6 | 20642666 | 21339743 | CDK5 regulatory subunit associated protein | 54901 | 611259 |
| SOX4 | chr6 | 21701950 | 21706826 | SRY (sex determining region Y)-box 4 | 6659 | 184430 |
| PRL | chr6 | 22395458 | 22405709 | prolactin | 5617 | 176760 |
| HDGFL1 | chr6 | 22677656 | 22678728 | hepatoma derived growth factor-like 1 | 154150 | 0 |
| NRSN1 | chr6 | 24234392 | 24255736 | neurensin 1 | 140767 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| DCDC2 | chr6 | 24279962 | 24466259 | doublecortin domain containing 2 | 51473 | 605755 |
| KAAG1 | chr6 | 24465109 | 24466491 | kidney associated antigen 1 | 353219 | 608211 |
| MRS2 | chr6 | 24511131 | 24533788 | MRS2-like, magnesium homeostasis factor | 57380 | 0 |
| GPLD1 | chr6 | 24536383 | 24597829 | glycosylphosphatidylinositol specific | 2822 | 602515 |
| ALDH5A1 | chr6 | 24603175 | 24645413 | aldehyde dehydrogenase 5A1 precursor, isoform 1 | 7915 | 610045 |
| KIAA0319 | chr6 | 24652310 | 24754362 | KIAA0319 | 9856 | 609269 |
| TTRAP | chr6 | 24758184 | 24775094 | TRAF and TNF receptor-associated protein | 51567 | 605764 |
| THEM2 | chr6 | 24775253 | 24809919 | thioesterase superfamily member 2 | 55856 | 0 |
| C6orf62 | chr6 | 24813145 | 24827382 | chromosome 6 open reading frame 62 | 81688 | 0 |
| GMNN | chr6 | 24883142 | 24894255 | geminin | 51053 | 602842 |
| FAM65B | chr6 | 24912491 | 25019174 | hypothetical protein LOC9750 isoform 1 | 9750 | 0 |
| LRRC16A | chr6 | 25387626 | 25728737 | leucine rich repeat containing 16 | 55604 | 0 |
| SCGN | chr6 | 25760407 | 25809985 | secretagogin precursor | 10590 | 609202 |
| HIST1H2AA | chr6 | 25834270 | 25834769 | histone cluster 1, H2aa | 221613 | 0 |
| HIST1H2BA | chr6 | 25835115 | 25835551 | histone cluster 1, H2ba | 255626 | 609904 |
| SLC17A4 | chr6 | 25862944 | 25888419 | solute carrier family 17 (sodium phosphate), | 10050 | 604216 |
| SLC17A1 | chr6 | 25891104 | 25940266 | solute carrier family 17 (sodium phosphate), | 6568 | 182308 |
| SLC17A3 | chr6 | 25953306 | 25982450 | solute carrier family 17 (sodium phosphate), | 10786 | 611034 |
| SLC17A2 | chr6 | 26020968 | 26038818 | solute carrier family 17 (sodium phosphate), | 10246 | 611049 |
| TRIM38 | chr6 | 26071049 | 26093329 | tripartite motif-containing 38 | 10475 | 0 |
| HIST1H1A | chr6 | 26125239 | 26126019 | histone cluster 1, H1a | 3024 | 142709 |
| HIST1H3A | chr6 | 26128696 | 26129165 | histone cluster 1, H3a | 8350 | 602810 |
| HIST1H4A | chr6 | 26129885 | 26130256 | histone cluster 1, H4a | 8359 | 602822 |
| HIST1H4B | chr6 | 26135103 | 26135459 | histone cluster 1, H4b | 8366 | 602829 |
| HIST1H3B | chr6 | 26139795 | 26140267 | histone cluster 1, H3b | 8358 | 602819 |
| HIST1H2AB | chr6 | 26141298 | 26141775 | histone cluster 1, H2ab | 8335 | 602795 |
| HIST1H2BB | chr6 | 26151434 | 26151864 | histone cluster 1, H2bb | 3018 | 602803 |
| HIST1H3C | chr6 | 26153617 | 26154075 | histone cluster 1, H3c | 8352 | 602812 |
| HIST1H1C | chr6 | 26163946 | 26164678 | histone cluster 1, H1c | 3006 | 142710 |
| HFE | chr6 | 26195487 | 26203448 | hemochromatosis protein isoform 11 precursor | 3077 | 235200 |
| HIST1H4C | chr6 | 26212154 | 26212543 | histone cluster 1, H4c | 8364 | 602827 |
| HIST1H1T | chr6 | 26215619 | 26216343 | histone cluster 1, H1t | 3010 | 142712 |
| HIST1H2BC | chr6 | 26231673 | 26232111 | histone cluster 1, H2bc | 8347 | 602847 |
| HIST1H2AC | chr6 | 26232351 | 26232896 | histone cluster 1, H2ac | 8334 | 602794 |
| HIST1H1E | chr6 | 26264537 | 26265321 | histone cluster 1, H1e | 3008 | 142220 |
| HIST1H2BD | chr6 | 26266327 | 26266848 | histone cluster 1, H2bd | 3017 | 602799 |

| | | | | | | |
|---|---|---|---|---|---|---|
| HIST1H2BE | chr6 | 26292002 | 26292436 | histone cluster 1, H2be | 8344 | 602805 |
| HIST1H4D | chr6 | 26296917 | 26297283 | histone cluster 1, H4d | 8360 | 602823 |
| HIST1H3D | chr6 | 26304991 | 26307443 | histone cluster 1, H3d | 8351 | 602811 |
| HIST1H2AD | chr6 | 26306990 | 26307450 | histone cluster 1, H2ad | 3013 | 602792 |
| HIST1H2BF | chr6 | 26307765 | 26308194 | histone cluster 1, H2bf | 8343 | 602804 |
| HIST1H4E | chr6 | 26312851 | 26313227 | histone cluster 1, H4e | 8367 | 602830 |
| HIST1H2BG | chr6 | 26324409 | 26324851 | histone cluster 1, H2bg | 8339 | 602798 |
| HIST1H2AE | chr6 | 26325126 | 26325690 | histone cluster 1, H2ae | 3012 | 602786 |
| HIST1H3E | chr6 | 26333361 | 26333822 | histone cluster 1, H3e | 8353 | 602813 |
| HIST1H1D | chr6 | 26342419 | 26343195 | histone cluster 1, H1d | 3007 | 142210 |
| HIST1H4F | chr6 | 26348632 | 26349000 | histone cluster 1, H4f | 8361 | 602824 |
| HIST1H4G | chr6 | 26354818 | 26355184 | histone cluster 1, H4g | 8369 | 602832 |
| HIST1H3F | chr6 | 26358348 | 26358814 | histone cluster 1, H3f | 8968 | 602816 |
| HIST1H2BH | chr6 | 26359857 | 26360281 | histone cluster 1, H2bh | 8345 | 602806 |
| HIST1H3G | chr6 | 26379125 | 26379591 | H3 histone family, member H | 8355 | 602815 |
| HIST1H2BI | chr6 | 26381182 | 26381618 | histone cluster 1, H2bi | 8346 | 602807 |
| HIST1H4H | chr6 | 26393333 | 26393706 | histone cluster 1, H4h | 8365 | 602828 |
| BTN3A2 | chr6 | 26473376 | 26486525 | butyrophilin, subfamily 3, member A2 precursor | 11118 | 0 |
| BTN2A2 | chr6 | 26491332 | 26503077 | butyrophilin, subfamily 2, member A2 isoform a | 10385 | 0 |
| BTN3A1 | chr6 | 26510459 | 26523421 | butyrophilin, subfamily 3, member A1 | 11119 | 0 |
| BTN2A3 | chr6 | 26530325 | 26539821 | butyrophilin, subfamily 2, member A3 | 54718 | 0 |
| BTN3A3 | chr6 | 26548741 | 26561621 | butyrophilin, subfamily 3, member A3 isoform a | 10384 | 0 |
| BTN2A1 | chr6 | 26566167 | 26577844 | butyrophilin, subfamily 2, member A1 isoform 1 | 11120 | 0 |
| BTN1A1 | chr6 | 26609473 | 26618631 | butyrophilin, subfamily 1, member A1 | 696 | 601610 |
| HMGN4 | chr6 | 26646550 | 26655141 | high mobility group nucleosomal binding domain | 10473 | 0 |
| ABT1 | chr6 | 26705158 | 26708254 | activator of basal transcription 1 | 29777 | 0 |
| ZNF322A | chr6 | 26742589 | 26767942 | zinc finger protein 322A | 79692 | 610847 |
| HIST1H2BJ | chr6 | 27208074 | 27208554 | histone cluster 1, H2bj | 8970 | 0 |
| HIST1H2AG | chr6 | 27208799 | 27211049 | histone cluster 1, H2ag | 8969 | 0 |
| HIST1H2BK | chr6 | 27214051 | 27222598 | histone cluster 1, H2bk | 85236 | 0 |
| HIST1H4I | chr6 | 27215066 | 27215436 | histone cluster 1, H4i | 8294 | 602833 |
| HIST1H2AH | chr6 | 27222886 | 27223324 | histone cluster 1, H2ah | 85235 | 0 |
| PRSS16 | chr6 | 27323486 | 27332228 | protease, serine, 16 | 10279 | 607169 |
| FKSG83 | chr6 | 27400556 | 27401720 | FKSG83 | 83954 | 0 |
| ZNF391 | chr6 | 27464502 | 27477206 | zinc finger protein 391 | 346157 | 0 |
| ZNF184 | chr6 | 27526505 | 27548858 | zinc finger protein 184 | 7738 | 602277 |

| | | | | | | |
|---|---|---|---|---|---|---|
| HIST1H2BL | chr6 | 27883235 | 27883688 | histone cluster 1, H2bl | 8340 | 602800 |
| HIST1H2AI | chr6 | 27883955 | 27884423 | histone cluster 1, H2ai | 8329 | 602787 |
| HIST1H3H | chr6 | 27885820 | 27886292 | histone cluster 1, H3h | 8357 | 602818 |
| HIST1H2AJ | chr6 | 27890059 | 27890497 | histone cluster 1, H2aj | 8331 | 602791 |
| HIST1H2BM | chr6 | 27890800 | 27891245 | histone cluster 1, H2bm | 8342 | 602802 |
| HIST1H4J | chr6 | 27899881 | 27900236 | histone cluster 1, H4j | 8363 | 602826 |
| HIST1H4K | chr6 | 27906931 | 27907284 | histone cluster 1, H4k | 8362 | 602825 |
| ZNF323 | chr6 | 28400493 | 28429951 | zinc finger protein 323 isoform 2 | 64288 | 610794 |
| ZKSCAN3 | chr6 | 28425669 | 28442503 | zinc finger with KRAB and SCAN domains 3 | 80317 | 0 |
| ZSCAN23 | chr6 | 28508410 | 28519258 | zinc finger protein 390 | 222696 | 0 |
| GPX6 | chr6 | 28579051 | 28591549 | glutathione peroxidase 6 | 257202 | 607913 |
| GPX5 | chr6 | 28601767 | 28609923 | glutathione peroxidase 5 precursor, isoform 1 | 2880 | 603435 |
| SCAND3 | chr6 | 28647386 | 28663091 | SCAN domain containing 3 | 114821 | 0 |
| TRIM27 | chr6 | 28978758 | 28999747 | ret finger protein | 5987 | 602165 |
| ZNF311 | chr6 | 29070572 | 29081016 | zinc finger protein 311 | 282890 | 0 |
| OR2W1 | chr6 | 29119968 | 29120931 | olfactory receptor, family 2, subfamily W, | 26692 | 0 |
| OR2B3P | chr6 | 29162062 | 29163004 | olfactory receptor, family 2, subfamily B, | 442184 | 0 |
| OR2J3 | chr6 | 29187646 | 29188582 | olfactory receptor, family 2, subfamily J, | 442186 | 0 |
| OR2J2 | chr6 | 29249289 | 29250330 | olfactory receptor, family 2, subfamily J, | 26707 | 0 |
| LOC651503 | chr6 | 29338458 | 29339835 | seven transmembrane helix receptor | 651503 | 0 |
| OR14J1 | chr6 | 29382445 | 29383410 | olfactory receptor, family 5, subfamily U member | 442191 | 0 |
| OR5V1 | chr6 | 29430985 | 29432033 | olfactory receptor, family 5, subfamily V, | 81696 | 0 |
| OR12D3 | chr6 | 29449179 | 29451047 | olfactory receptor, family 12, subfamily D, | 81797 | 0 |
| OR12D2 | chr6 | 29472394 | 29473427 | olfactory receptor, family 12, subfamily D, | 26529 | 0 |
| OR11A1 | chr6 | 29501260 | 29503488 | olfactory receptor, family 11, subfamily A, | 26531 | 0 |
| OR10C1 | chr6 | 29515771 | 29516709 | olfactory receptor, family 10, subfamily C, | 442194 | 0 |
| OR2H1 | chr6 | 29534208 | 29540076 | olfactory receptor, family 2, subfamily H, | 26716 | 0 |
| MAS1L | chr6 | 29562523 | 29563658 | MAS1 oncogene-like | 116511 | 607235 |
| UBD | chr6 | 29631390 | 29635681 | ubiquitin D | 10537 | 606050 |
| OR2H2 | chr6 | 29663661 | 29664724 | olfactory receptor, family 2, subfamily H, | 7932 | 600578 |
| GABBR1 | chr6 | 29677983 | 29708941 | gamma-aminobutyric acid (GABA) B receptor 1 | 2550 | 603540 |
| MOG | chr6 | 29732787 | 29748126 | myelin oligodendrocyte glycoprotein isoform | 4340 | 159465 |
| ZFP57 | chr6 | 29748238 | 29756866 | zinc finger protein 57 homolog | 346171 | 0 |
| HLA-F | chr6 | 29799095 | 29803052 | major histocompatibility complex, class I, F | 3134 | 143110 |
| HLA-G | chr6 | 29902734 | 29906878 | major histocompatibility complex, class I, G | 3135 | 142871 |
| HLA-A29.1 | chr6 | 30018304 | 30085130 | major histocompatibility complex class I | 649853 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| HLA-A | chr6 | 30018309 | 30021632 | major histocompatibility complex, class I, A | 3105 | 142800 |
| HCG9 | chr6 | 30050870 | 30054148 | HLA complex group 9 | 10255 | 0 |
| ZNRD1 | chr6 | 30137014 | 30140664 | zinc ribbon domain containing 1 | 30834 | 607525 |
| PPP1R11 | chr6 | 30142910 | 30146085 | protein phosphatase 1, regulatory (inhibitor) | 6992 | 606670 |
| RNF39 | chr6 | 30146022 | 30151607 | ring finger protein 39 isoform 1 | 80352 | 607524 |
| TRIM31 | chr6 | 30178654 | 30188846 | tripartite motif protein 31 | 11074 | 609316 |
| TRIM40 | chr6 | 30212488 | 30224491 | tripartite motif-containing 40 | 135644 | 0 |
| TRIM10 | chr6 | 30227702 | 30236690 | tripartite motif-containing 10 isoform 1 | 10107 | 605701 |
| TRIM15 | chr6 | 30238961 | 30248452 | tripartite motif protein 15 | 89870 | 0 |
| TRIM26 | chr6 | 30260212 | 30289132 | tripartite motif-containing 26 | 7726 | 600830 |
| FLJ45422 | chr6 | 30335352 | 30342707 | hypothetical protein LOC441140 | 441140 | 0 |
| TRIM39 | chr6 | 30403018 | 30419484 | tripartite motif-containing 39 isoform 1 | 56658 | 605700 |
| RPP21 | chr6 | 30420915 | 30422611 | ribonuclease P/MRP 21kDa subunit | 79897 | 0 |
| HLA-E | chr6 | 30565249 | 30569064 | major histocompatibility complex, class I, E | 3133 | 143010 |
| GNL1 | chr6 | 30621674 | 30632987 | guanine nucleotide binding protein-like 1 | 2794 | 143024 |
| PRR3 | chr6 | 30632734 | 30640159 | proline-rich protein 3 isoform b | 80742 | 0 |
| ABCF1 | chr6 | 30647148 | 30667286 | ATP-binding cassette, sub-family F, member 1 | 23 | 603429 |
| PPP1R10 | chr6 | 30676161 | 30692999 | protein phosphatase 1, regulatory subunit 10 | 5514 | 603771 |
| MRPS18B | chr6 | 30693464 | 30702153 | mitochondrial ribosomal protein S18B | 28973 | 611982 |
| C6orf134 | chr6 | 30702591 | 30720427 | hypothetical protein LOC79969 isoform 2 | 79969 | 0 |
| C6orf136 | chr6 | 30722779 | 30728961 | hypothetical protein LOC221545 isoform 1 | 221545 | 0 |
| DHX16 | chr6 | 30728885 | 30748736 | DEAH (Asp-Glu-Ala-His) box polypeptide 16 | 8449 | 603405 |
| KIAA1949 | chr6 | 30752144 | 30761963 | phostensin | 170954 | 610990 |
| NRM | chr6 | 30763805 | 30766748 | nurim | 11270 | 0 |
| MDC1 | chr6 | 30775562 | 30793437 | mediator of DNA damage checkpoint 1 | 9656 | 607593 |
| TUBB | chr6 | 30796135 | 30801172 | tubulin, beta | 203068 | 191130 |
| FLOT1 | chr6 | 30803491 | 30818432 | flotillin 1 | 10211 | 606998 |
| IER3 | chr6 | 30818954 | 30820306 | immediate early response 3 | 8870 | 602996 |
| DDR1 | chr6 | 30959839 | 30975910 | discoidin domain receptor family, member 1 | 780 | 600408 |
| GTF2H4 | chr6 | 30983955 | 30989858 | general transcription factor IIH, polypeptide 4, | 2968 | 601760 |
| VARS2 | chr6 | 30989960 | 31002212 | valyl-tRNA synthetase 2-like | 57176 | 0 |
| SFTPG | chr6 | 31007105 | 31007931 | surfactant associated protein G precursor | 389376 | 0 |
| DPCR1 | chr6 | 31027537 | 31029977 | diffuse panbronchiolitis critical region 1 | 135656 | 604809 |
| MUC21 | chr6 | 31059463 | 31065654 | mucin 21 | 394263 | 0 |
| C6orf15 | chr6 | 31186980 | 31188311 | STG protein | 29113 | 611401 |
| PSORS1C1 | chr6 | 31190601 | 31215816 | SEEK1 protein | 170679 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CDSN | chr6 | 31190848 | 31196202 | corneodesmosin precursor | 1041 | 602593 |
| PSORS1C2 | chr6 | 31213289 | 31215106 | SPR1 protein | 170680 | 0 |
| CCHCR1 | chr6 | 31218194 | 31233545 | coiled-coil alpha-helical rod protein 1 isoform | 54535 | 605310 |
| TCF19 | chr6 | 31234281 | 31239970 | transcription factor 19 | 6941 | 600912 |
| POU5F1 | chr6 | 31240093 | 31242006 | POU domain, class 5, transcription factor 1 | 5460 | 164177 |
| HCG27 | chr6 | 31273577 | 31279723 | HLA complex group 27 | 253018 | 0 |
| HLA-C | chr6 | 31344508 | 31347834 | major histocompatibility complex, class I, C | 3107 | 142840 |
| HCP5 | chr6 | 31538937 | 31541460 | HLA complex P5 | 10866 | 604676 |
| MICB | chr6 | 31573833 | 31586880 | MHC class I polypeptide-related sequence B | 4277 | 602436 |
| MCCD1 | chr6 | 31604717 | 31605987 | mitochondrial coiled-coil domain 1 | 401250 | 609624 |
| BAT1 | chr6 | 31605976 | 31618204 | HLA-B associated transcript 1 | 7919 | 142560 |
| ATP6V1G2 | chr6 | 31620220 | 31622606 | ATPase, H+ transporting, lysosomal, V1 subunit | 534 | 606853 |
| NFKBIL1 | chr6 | 31623350 | 31634584 | nuclear factor of kappa light polypeptide gene | 4795 | 601022 |
| LTA | chr6 | 31648071 | 31650077 | lymphotoxin alpha precursor | 4049 | 153440 |
| TNF | chr6 | 31651328 | 31654089 | tumor necrosis factor alpha | 7124 | 191160 |
| LTB | chr6 | 31656316 | 31658181 | lymphotoxin-beta isoform b | 4050 | 600978 |
| LST1 | chr6 | 31661949 | 31664664 | leukocyte specific transcript 1 isoform 3 | 7940 | 109170 |
| NCR3 | chr6 | 31664650 | 31668741 | natural cytotoxicity triggering receptor 3 | 259197 | 611550 |
| AIF1 | chr6 | 31691011 | 31692775 | allograft inflammatory factor 1 isoform 3 | 199 | 601833 |
| BAT2 | chr6 | 31696428 | 31713533 | HLA-B associated transcript-2 | 7916 | 142580 |
| BAT3 | chr6 | 31714783 | 31728456 | HLA-B associated transcript-3 isoform b | 7917 | 142590 |
| APOM | chr6 | 31731649 | 31733965 | apolipoprotein M | 55937 | 606907 |
| C6orf47 | chr6 | 31734053 | 31736528 | G4 protein | 57827 | 0 |
| BAT4 | chr6 | 31737841 | 31741142 | HLA-B associated transcript 4 | 7918 | 142610 |
| CSNK2B | chr6 | 31741635 | 31745822 | casein kinase 2, beta polypeptide | 1460 | 115441 |
| LY6G5B | chr6 | 31746706 | 31748206 | lymphocyte antigen 6 complex G5B | 58496 | 610433 |
| LY6G5C | chr6 | 31752439 | 31756120 | lymphocyte antigen 6 complex G5C | 80741 | 610434 |
| BAT5 | chr6 | 31762714 | 31779067 | HLA-B associated transcript 5 | 7920 | 142620 |
| LY6G6F | chr6 | 31782662 | 31786351 | G6f protein | 259215 | 611404 |
| LY6G6D | chr6 | 31791111 | 31793560 | lymphocyte antigen 6 complex G6D | 58530 | 606038 |
| LY6G6C | chr6 | 31794404 | 31797489 | lymphocyte antigen 6 complex G6C | 80740 | 610435 |
| C6orf25 | chr6 | 31799139 | 31800829 | G6B protein isoform G6b-A precursor | 80739 | 606520 |
| DDAH2 | chr6 | 31802796 | 31806018 | dimethylarginine dimethylaminohydrolase 2 | 23564 | 604744 |
| CLIC1 | chr6 | 31806338 | 31812320 | chloride intracellular channel 1 | 1192 | 602872 |
| MSH5 | chr6 | 31815752 | 31838429 | mutS homolog 5 isoform c | 4439 | 603382 |
| C6orf26 | chr6 | 31838751 | 31840601 | hypothetical protein LOC401251 | 401251 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| C6orf27 | chr6 | 31841349 | 31853087 | G7c protein | 80737 | 609693 |
| VARS | chr6 | 31853277 | 31871691 | valyl-tRNA synthetase | 7407 | 192150 |
| LSM2 | chr6 | 31873154 | 31882722 | LSM2 homolog, U6 small nuclear RNA associated | 57819 | 607282 |
| HSPA1L | chr6 | 31885374 | 31890814 | heat shock 70kDa protein 1-like | 3305 | 140559 |
| HSPA1A | chr6 | 31891269 | 31893698 | heat shock 70kDa protein 1A | 3303 | 140550 |
| HSPA1B | chr6 | 31903490 | 31906010 | heat shock 70kDa protein 1B | 3304 | 603012 |
| C6orf48 | chr6 | 31910671 | 31915518 | G8 protein | 50854 | 605447 |
| NEU1 | chr6 | 31934807 | 31938688 | neuraminidase precursor | 4758 | 608272 |
| SLC44A4 | chr6 | 31938948 | 31954802 | choline transporter-like protein 4 | 80736 | 606107 |
| EHMT2 | chr6 | 31955515 | 31973443 | euchromatic histone-lysine N-methyltransferase 2 | 10919 | 604599 |
| ZBTB12 | chr6 | 31975372 | 31977748 | zinc finger and BTB domain containing 12 | 221527 | 0 |
| C2 | chr6 | 32003472 | 32021427 | complement component 2 precursor | 717 | 217000 |
| CFB | chr6 | 32021699 | 32027840 | complement factor B preproprotein | 629 | 138470 |
| RDBP | chr6 | 32027843 | 32034843 | RD RNA-binding protein | 7936 | 154040 |
| SKIV2L | chr6 | 32034559 | 32045511 | superkiller viralicidic activity 2-like homolog | 6499 | 600478 |
| DOM3Z | chr6 | 32045566 | 32048011 | DOM-3 homolog Z | 1797 | 605996 |
| STK19 | chr6 | 32047624 | 32057201 | serine/threonine kinase 19 isoform 1 | 8859 | 604977 |
| C4B | chr6 | 32057812 | 32078436 | complement component 4B preproprotein | 721 | 120820 |
| C4A | chr6 | 32090549 | 32111173 | complement component 4A preproprotein | 720 | 120810 |
| CYP21A2 | chr6 | 32114060 | 32117398 | cytochrome P450, family 21, subfamily A, | 1589 | 201910 |
| TNXB | chr6 | 32116910 | 32121883 | tenascin XB isoform 2 | 7148 | 600985 |
| CREBL1 | chr6 | 32191022 | 32203995 | cAMP responsive element binding protein-like 1 | 1388 | 600984 |
| FKBPL | chr6 | 32204464 | 32206045 | WAF-1/CIP1 stabilizing protein 39 | 63943 | 0 |
| PRRT1 | chr6 | 32224118 | 32227698 | NG5 protein | 80863 | 0 |
| PPT2 | chr6 | 32229278 | 32239429 | palmitoyl-protein thioesterase 2 isoform b | 9374 | 603298 |
| EGFL8 | chr6 | 32240382 | 32244040 | NG3 protein | 80864 | 609897 |
| AGPAT1 | chr6 | 32243967 | 32251877 | 1-acylglycerol-3-phosphate O-acyltransferase 1 | 10554 | 603099 |
| RNF5 | chr6 | 32254139 | 32256548 | ring finger protein 5 | 6048 | 602677 |
| AGER | chr6 | 32256723 | 32260001 | advanced glycosylation end product-specific | 177 | 600214 |
| PBX2 | chr6 | 32260487 | 32265941 | pre-B-cell leukemia homeobox 2 | 5089 | 176311 |
| GPSM3 | chr6 | 32266521 | 32271278 | G-protein signaling modulator 3 (AGS3-like, C. | 63940 | 0 |
| NOTCH4 | chr6 | 32270598 | 32299822 | notch4 preproprotein | 4855 | 164951 |
| C6orf10 | chr6 | 32368452 | 32447634 | chromosome 6 open reading frame 10 | 10665 | 0 |
| BTNL2 | chr6 | 32470491 | 32482878 | butyrophilin-like 2 | 56244 | 606000 |
| HLA-DRA | chr6 | 32515624 | 32520799 | major histocompatibility complex, class II, DR | 3122 | 142860 |

| | | | | | | |
|---|---|---|---|---|---|---|
| HLA-DRB5 | chr6 | 32593138 | 32605984 | major histocompatibility complex, class II, DR | 3127 | 604776 |
| HLA-DRB1 | chr6 | 32654524 | 32665540 | major histocompatibility complex, class II, DR | 3123 | 142857 |
| HLA-DQA1 | chr6 | 32713160 | 32719406 | major histocompatibility complex, class II, DQ | 3117 | 146880 |
| HLA-DQB1 | chr6 | 32735634 | 32742444 | major histocompatibility complex, class II, DQ | 3119 | 604305 |
| HLA-DQA2 | chr6 | 32817140 | 32823197 | major histocompatibility complex, class II, DQ | 3118 | 0 |
| HLA-DOB | chr6 | 32888518 | 32892803 | major histocompatibility complex, class II, DO | 3112 | 600629 |
| TAP2 | chr6 | 32897587 | 32914525 | transporter 2, ATP-binding cassette, sub-family | 6891 | 170261 |
| PSMB8 | chr6 | 32916476 | 32919794 | proteasome beta 8 subunit isoform E2 proprotein | 5696 | 177046 |
| TAP1 | chr6 | 32920964 | 32929726 | transporter 1, ATP-binding cassette, sub-family | 6890 | 170260 |
| PSMB9 | chr6 | 32929915 | 32935604 | proteasome beta 9 subunit isoform 1 proprotein | 5698 | 177045 |
| HLA-DMB | chr6 | 33010393 | 33016795 | major histocompatibility complex, class II, DM | 3109 | 142856 |
| HLA-DMA | chr6 | 33024373 | 33028831 | major histocompatibility complex, class II, DM | 3108 | 142855 |
| BRD2 | chr6 | 33044414 | 33057260 | bromodomain containing 2 | 6046 | 601540 |
| HLA-DOA | chr6 | 33079939 | 33085367 | major histocompatibility complex, class II, DO | 3111 | 142930 |
| HLA-DPA1 | chr6 | 33140771 | 33149356 | major histocompatibility complex, class II, DP | 3113 | 142880 |
| HLA-DPB1 | chr6 | 33151737 | 33162954 | major histocompatibility complex, class II, DP | 3115 | 142858 |
| COL11A2 | chr6 | 33238446 | 33268223 | collagen, type XI, alpha 2 isoform 2 | 1302 | 120290 |
| RXRB | chr6 | 33269343 | 33276410 | retinoid X receptor, beta | 6257 | 180246 |
| SLC39A7 | chr6 | 33276580 | 33280191 | solute carrier family 39, member 7 | 7922 | 601416 |
| HSD17B8 | chr6 | 33280396 | 33282585 | estradiol 17 beta-dehydrogenase 8 | 7923 | 601417 |
| RING1 | chr6 | 33284263 | 33288476 | ring finger protein 1 | 6015 | 602045 |
| VPS52 | chr6 | 33326027 | 33347640 | vacuolar protein sorting 52 | 6293 | 603443 |
| RPS18 | chr6 | 33347829 | 33352259 | ribosomal protein S18 | 6222 | 180473 |
| B3GALT4 | chr6 | 33352894 | 33354580 | UDP-Gal:betaGlcNAc beta | 8705 | 603095 |
| WDR46 | chr6 | 33354863 | 33364969 | WD repeat domain 46 | 9277 | 0 |
| PFDN6 | chr6 | 33365355 | 33366689 | HLA class II region expressed gene KE2 | 10471 | 605660 |
| RGL2 | chr6 | 33367415 | 33374716 | ral guanine nucleotide dissociation | 5863 | 602306 |
| TAPBP | chr6 | 33375451 | 33389967 | tapasin isoform 1 precursor | 6892 | 601962 |
| ZBTB22 | chr6 | 33390173 | 33393472 | zinc finger and BTB domain containing 22 | 9278 | 0 |
| DAXX | chr6 | 33394378 | 33398682 | death-associated protein 6 | 1616 | 603186 |
| KIFC1 | chr6 | 33467290 | 33485677 | kinesin family member C1 | 3833 | 603763 |
| PHF1 | chr6 | 33486771 | 33492191 | PHD finger protein 1 isoform b | 5252 | 602881 |
| CUTA | chr6 | 33492296 | 33494043 | cutA divalent cation tolerance homolog isoform | 51596 | 0 |
| SYNGAP1 | chr6 | 33495824 | 33529444 | synaptic Ras GTPase activating protein 1 isoform | 8831 | 603384 |
| ZBTB9 | chr6 | 33530333 | 33533296 | zinc finger and BTB domain containing 9 | 221504 | 0 |
| BAK1 | chr6 | 33648301 | 33656048 | BCL2-antagonist/killer 1 | 578 | 600516 |

| | | | | | | |
|---|---|---|---|---|---|---|
| FLJ43752 | chr6 | 33661860 | 33669093 | hypothetical protein LOC401253 | 401253 | 0 |
| ITPR3 | chr6 | 33697138 | 33772326 | inositol 1,4,5-triphosphate receptor, type 3 | 3710 | 147267 |
| C6orf125 | chr6 | 33773323 | 33787482 | hypothetical protein LOC84300 | 84300 | 0 |
| IHPK3 | chr6 | 33797421 | 33822660 | inositol hexaphosphate kinase 3 | 117283 | 606993 |
| LEMD2 | chr6 | 33846968 | 33864874 | LEM domain containing 2 | 221496 | 0 |
| MLN | chr6 | 33870426 | 33879771 | motilin isoform 1 preproprotein | 4295 | 158270 |
| GRM4 | chr6 | 34097606 | 34209421 | glutamate receptor, metabotropic 4 | 2914 | 604100 |
| HMGA1 | chr6 | 34312627 | 34321985 | high mobility group AT-hook 1 isoform b | 3159 | 600701 |
| C6orf1 | chr6 | 34322135 | 34324882 | hypothetical protein LOC221491 | 221491 | 611419 |
| NUDT3 | chr6 | 34363979 | 34468419 | nudix-type motif 3 | 11165 | 609228 |
| RPS10 | chr6 | 34493210 | 34501854 | ribosomal protein S10 | 6204 | 603632 |
| PACSIN1 | chr6 | 34541882 | 34610976 | protein kinase C and casein kinase substrate in | 29993 | 606512 |
| SPDEF | chr6 | 34613557 | 34632069 | SAM pointed domain containing ets transcription | 25803 | 608144 |
| C6orf106 | chr6 | 34663049 | 34772603 | chromosome 6 open reading frame 106 isoform a | 64771 | 0 |
| SNRPC | chr6 | 34833289 | 34849549 | small nuclear ribonucleoprotein polypeptide C | 6631 | 603522 |
| UHRF1BP1 | chr6 | 34867771 | 34953269 | ICBP90 binding protein 1 | 54887 | 0 |
| ZFAND3 | chr6 | 37895610 | 38230375 | zinc finger, AN1-type domain 3 | 60685 | 607455 |
| BTBD9 | chr6 | 38244204 | 38671821 | BTB (POZ) domain containing 9 isoform b | 114781 | 611237 |
| RIMS1 | chr6 | 72653370 | 73169228 | regulating synaptic membrane exocytosis 1 | 22999 | 606629 |
| ANKRD6 | chr6 | 90199615 | 90400124 | ankyrin repeat domain 6 | 22881 | 610583 |
| LYRM2 | chr6 | 90398665 | 90405195 | LYR motif containing 2 | 57226 | 0 |
| MDN1 | chr6 | 90409951 | 90586163 | MDN1, midasin homolog | 23195 | 0 |
| CASP8AP2 | chr6 | 90596378 | 90640875 | CASP8 associated protein 2 | 9994 | 606880 |
| GJA10 | chr6 | 90660908 | 90662538 | gap junction protein, alpha 10, 62kDa | 84694 | 611924 |
| BACH2 | chr6 | 90692968 | 91063182 | BTB and CNC homology 1, basic leucine zipper | 60468 | 605394 |
| MAP3K7 | chr6 | 91282073 | 91353628 | mitogen-activated protein kinase kinase kinase 7 | 6885 | 602614 |
| USP45 | chr6 | 99986905 | 100069973 | ubiquitin specific peptidase 45 | 85015 | 0 |
| CCNC | chr6 | 100096984 | 100123411 | cyclin C isoform b | 892 | 123838 |
| PRDM13 | chr6 | 100161370 | 100170175 | PR domain containing 13 | 59336 | 0 |
| PDSS2 | chr6 | 107580453 | 107887472 | prenyl diphosphate synthase, subunit 2 | 57107 | 610564 |
| SOBP | chr6 | 107918009 | 108089201 | sine oculis binding protein homolog | 55084 | 0 |
| SCML4 | chr6 | 108130056 | 108252214 | sex comb on midleg-like 4 | 256380 | 0 |
| REV3L | chr6 | 111726926 | 111911107 | REV3-like, catalytic subunit of DNA polymerase | 5980 | 602776 |
| TRAF3IP2 | chr6 | 111986837 | 112033755 | TRAF3 interacting protein 2 isoform 1 | 10758 | 607043 |
| FYN | chr6 | 112089179 | 112301320 | protein-tyrosine kinase fyn isoform a | 2534 | 137025 |
| MAD1L1 | chr7 | 1628668 | 2045824 | MAD1-like 1 protein | 8379 | 602686 |

| | | | | | | |
|---|---|---|---|---|---|---|
| FTSJ2 | chr7 | 2047168 | 2055074 | FtsJ homolog 2 | 29960 | 606906 |
| NUDT1 | chr7 | 2055097 | 2064021 | nudix-type motif 1 isoform p22 | 4521 | 600312 |
| SNX8 | chr7 | 2067878 | 2127340 | sorting nexin 8 | 29886 | 0 |
| EIF3B | chr7 | 2167714 | 2193616 | eukaryotic translation initiation factor 3, | 8662 | 603917 |
| CHST12 | chr7 | 2216499 | 2247455 | carbohydrate (chondroitin 4) sulfotransferase | 55501 | 610129 |
| GRIFIN | chr7 | 2288012 | 2330625 | galectin-related inter-fiber protein | 402635 | 0 |
| LFNG | chr7 | 2332719 | 2341302 | lunatic fringe isoform a | 3955 | 602576 |
| C7orf27 | chr7 | 2350754 | 2368387 | hypothetical protein LOC221927 | 221927 | 0 |
| IQCE | chr7 | 2371872 | 2427609 | IQ motif containing E isoform 1 | 23288 | 0 |
| TTYH3 | chr7 | 2444843 | 2477668 | tweety 3 | 80727 | 608919 |
| AMZ1 | chr7 | 2492403 | 2528310 | archaelysin family metallopeptidase 1 | 155185 | 0 |
| GNA12 | chr7 | 2540983 | 2657200 | guanine nucleotide binding protein (G protein) | 2768 | 604394 |
| CARD11 | chr7 | 2719009 | 2856820 | caspase recruitment domain family, member 11 | 84433 | 607210 |
| SDK1 | chr7 | 3114320 | 4081870 | sidekick 1 isoform 1 | 221935 | 607216 |
| NEUROD6 | chr7 | 31150321 | 31153778 | neurogenic differentiation 6 | 63974 | 611513 |
| VWC2 | chr7 | 49590517 | 49729399 | von Willebrand factor C domain containing 2 | 375567 | 611108 |
| DKFZp564N2472 | chr7 | 52877557 | 52878826 | hypothetical protein LOC285877 | 285877 | 0 |
| OR2A5 | chr7 | 143185142 | 143186077 | olfactory receptor, family 2, subfamily A, | 393046 | 0 |
| OR2A25 | chr7 | 143208960 | 143209892 | olfactory receptor, family 2, subfamily A, | 392138 | 0 |
| OR2A12 | chr7 | 143229848 | 143230780 | olfactory receptor, family 2, subfamily A, | 346525 | 0 |
| OR2A2 | chr7 | 143244323 | 143245280 | olfactory receptor, family 2, subfamily A, | 442361 | 0 |
| OR2A14 | chr7 | 143263853 | 143264785 | olfactory receptor, family 2, subfamily A, | 135941 | 0 |
| CTAGE4 | chr7 | 143318195 | 143320821 | CTAGE family, member 4 | 100128553 | 608910 |
| FLJ43692 | chr7 | 143321324 | 143330384 | hypothetical protein LOC445328 | 445328 | 0 |
| OR2A1 | chr7 | 143366653 | 143367584 | olfactory receptor, family 2, subfamily A, | 346528 | 0 |
| OR2A42 | chr7 | 143366653 | 143367566 | olfactory receptor, family 2, subfamily A, | 402317 | 0 |
| OR2A7 | chr7 | 143393437 | 143394369 | olfactory receptor, family 2, subfamily A, | 401427 | 0 |
| ARHGEF5 | chr7 | 143490136 | 143515372 | rho guanine nucleotide exchange factor 5 | 7984 | 600888 |
| NOBOX | chr7 | 143533686 | 143544968 | NOBOX oogenesis homeobox | 135935 | 610934 |
| TPK1 | chr7 | 143586690 | 143970794 | thiamin pyrophosphokinase 1 isoform b | 27010 | 606370 |
| ADAM7 | chr8 | 24354453 | 24422163 | a disintegrin and metalloproteinase domain 7 | 8756 | 607310 |
| PNMA2 | chr8 | 26418112 | 26427400 | paraneoplastic antigen MA2 | 10687 | 603970 |
| DPYSL2 | chr8 | 26491337 | 26571610 | dihydropyrimidinase-like 2 | 1808 | 602463 |
| ADRA1A | chr8 | 26661583 | 26778839 | alpha-1A-adrenergic receptor isoform 2 | 148 | 104221 |
| FZD3 | chr8 | 28407691 | 28477878 | frizzled 3 | 7976 | 606143 |
| EXTL3 | chr8 | 28615071 | 28667121 | Reg receptor | 2137 | 605744 |

| | | | | | | |
|---|---|---|---|---|---|---|
| INTS9 | chr8 | 28681098 | 28803398 | integrator complex subunit 9 | 55756 | 611352 |
| HMBOX1 | chr8 | 28804143 | 28965334 | homeobox containing 1 | 79618 | 0 |
| KIF13B | chr8 | 28980713 | 29176529 | kinesin family member 13B | 23303 | 607350 |
| DUSP4 | chr8 | 29249538 | 29262241 | dual specificity phosphatase 4 isoform 2 | 1846 | 602747 |
| KCNU1 | chr8 | 36760999 | 36912801 | potassium channel, subfamily U, member 1 | 157855 | 0 |
| FKSG2 | chr8 | 36865154 | 36865984 | apoptosis inhibitor | 59347 | 0 |
| A26A1 | chr8 | 43266741 | 43337485 | ANKRD26-like family A, member 1 isoform 2 | 340441 | 608915 |
| XKR4 | chr8 | 56177570 | 56601262 | XK, Kell blood group complex subunit-related | 114786 | 0 |
| CHD7 | chr8 | 61753892 | 61942017 | chromodomain helicase DNA binding protein 7 | 55636 | 608765 |
| RIMS2 | chr8 | 104582151 | 105334627 | regulating synaptic membrane exocytosis 2 | 9699 | 606630 |
| ZFPM2 | chr8 | 106400322 | 106885941 | zinc finger protein, multitype 2 | 23414 | 603693 |
| MTSS1 | chr8 | 125632208 | 125809911 | metastasis suppressor 1 | 9788 | 608486 |
| EFR3A | chr8 | 132985540 | 133094953 | EFR3 homolog A | 23167 | 611798 |
| OC90 | chr8 | 133105909 | 133131298 | otoconin 90 | 729330 | 601658 |
| LRRC6 | chr8 | 133653628 | 133756995 | leucine rich repeat containing 6 | 23639 | 0 |
| TMEM71 | chr8 | 133791375 | 133842010 | transmembrane protein 71 | 137835 | 0 |
| ADAMTSL1 | chr9 | 18464103 | 18674952 | ADAMTS-like 1 isoform 2 precursor | 92949 | 609198 |
| RORB | chr9 | 74341805 | 74531669 | RAR-related orphan receptor B | 6096 | 601972 |
| TRPM6 | chr9 | 74566964 | 74732564 | transient receptor potential cation channel, | 140803 | 607009 |
| C9orf40 | chr9 | 74791052 | 74797356 | hypothetical protein LOC55071 | 55071 | 0 |
| C9orf41 | chr9 | 74827428 | 74872864 | hypothetical protein LOC138199 | 138199 | 0 |
| C9orf95 | chr9 | 74905669 | 74932687 | nicotinamide riboside kinase 1 isoform 2 | 54981 | 608704 |
| OSTF1 | chr9 | 74932951 | 74991668 | osteoclast stimulating factor 1 | 26578 | 610180 |
| ZNF322B | chr9 | 97039091 | 97041465 | zinc finger protein 322B | 387328 | 0 |
| KIAA1529 | chr9 | 97080333 | 97219124 | hypothetical protein LOC57653 | 57653 | 0 |
| NEK6 | chr9 | 124099802 | 124194271 | putative serine-threonine protein kinase | 10783 | 604884 |
| PSMB7 | chr9 | 124195298 | 124257275 | proteasome beta 7 subunit proprotein | 5695 | 604030 |
| NR5A1 | chr9 | 124323068 | 124349253 | nuclear receptor subfamily 5, group A, member 1 | 2516 | 184757 |
| NR6A1 | chr9 | 124364256 | 124613130 | nuclear receptor subfamily 6, group A, member 1 | 2649 | 602778 |
| OLFML2A | chr9 | 124618990 | 124656711 | olfactomedin-like 2A | 169611 | 0 |
| WDR38 | chr9 | 124695308 | 124699713 | WD repeat domain 38 | 401551 | 0 |
| RPL35 | chr9 | 124699711 | 124703794 | ribosomal protein L35 | 11224 | 0 |
| ARPC5L | chr9 | 124711037 | 124719249 | actin related protein 2/3 complex, subunit | 81873 | 0 |
| GOLGA1 | chr9 | 124720182 | 124782911 | golgin 97 | 2800 | 602502 |
| C9orf126 | chr9 | 124794620 | 124985329 | hypothetical protein LOC286205 | 286205 | 0 |
| PPP6C | chr9 | 124988405 | 125031772 | protein phosphatase 6, catalytic subunit isoform | 5537 | 300141 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RABEPK | chr9 | 125042610 | 125075991 | Rab9 effector protein with kelch motifs | 10244 | 605962 |
| HSPA5 | chr9 | 125076680 | 125083215 | heat shock 70kDa protein 5 | 3309 | 138120 |
| GAPVD1 | chr9 | 125103664 | 125206843 | GTPase activating protein and VPS9 domains 1 | 26130 | 611714 |
| MAPKAP1 | chr9 | 125279227 | 125549067 | mitogen-activated protein kinase associated | 79109 | 610558 |
| PBX3 | chr9 | 125589219 | 125809207 | pre-B-cell leukemia homeobox 3 | 5090 | 176312 |
| KIF5B | chr10 | 32337943 | 32385377 | kinesin family member 5B | 3799 | 602809 |
| EPC1 | chr10 | 32597864 | 32676119 | enhancer of polycomb 1 | 80314 | 610999 |
| CCDC7 | chr10 | 32775046 | 32903496 | coiled-coil domain containing 7 | 221016 | 0 |
| C10orf68 | chr10 | 32896656 | 33211798 | chromosome 10 open reading frame 68 | 79741 | 0 |
| ZNF33B | chr10 | 42404560 | 42453998 | zinc finger protein 33B | 7582 | 194522 |
| ADO | chr10 | 64234521 | 64238244 | 2-aminoethanethiol (cysteamine) dioxygenase | 84890 | 611392 |
| EGR2 | chr10 | 64241762 | 64246133 | early growth response 2 protein | 1959 | 129010 |
| NRBF2 | chr10 | 64563012 | 64584792 | nuclear receptor binding factor 2 | 29982 | 0 |
| JMJD1C | chr10 | 64596995 | 64895728 | jumonji domain containing 1C isoform a | 221037 | 604503 |
| KIAA1274 | chr10 | 71908569 | 71998212 | KIAA1274 | 27143 | 0 |
| PRF1 | chr10 | 72027109 | 72032537 | perforin 1 precursor | 5551 | 170280 |
| ADAMTS14 | chr10 | 72102564 | 72192199 | ADAM metallopeptidase with thrombospondin type 1 | 140766 | 607506 |
| C10orf27 | chr10 | 72201000 | 72215163 | stromal protein associated with thymii and lymph | 219793 | 0 |
| SGPL1 | chr10 | 72245721 | 72309871 | sphingosine-1-phosphate lyase 1 | 8879 | 603729 |
| PCBD1 | chr10 | 72313272 | 72318547 | pterin-4 alpha-carbinolamine | 5092 | 126090 |
| BTRC | chr10 | 103103814 | 103307058 | beta-transducin repeat containing protein | 8945 | 603482 |
| POLL | chr10 | 103328630 | 103337963 | polymerase (DNA directed), lambda | 27343 | 606343 |
| RP11-529I10.4 | chr10 | 103338078 | 103359399 | DPCD protein | 25911 | 0 |
| FBXW4 | chr10 | 103360411 | 103444733 | F-box and WD repeat domain containing 4 | 6468 | 608071 |
| FGF8 | chr10 | 103519876 | 103525817 | fibroblast growth factor 8 isoform A precursor | 2253 | 600483 |
| NPM3 | chr10 | 103531072 | 103533148 | nucleophosmin/nucleoplasmin 3 | 10360 | 606456 |
| MGEA5 | chr10 | 103534198 | 103568165 | meningioma expressed antigen 5 (hyaluronidase) | 10724 | 604039 |
| KCNIP2 | chr10 | 103575721 | 103589601 | Kv channel interacting protein 2 isoform 5 | 30819 | 604661 |
| C10orf76 | chr10 | 103595345 | 103805922 | hypothetical protein LOC79591 | 79591 | 0 |
| HPS6 | chr10 | 103815136 | 103817782 | Hermansky-Pudlak syndrome-6 | 79803 | 607522 |
| LDB1 | chr10 | 103857314 | 103870200 | LIM domain binding 1 isoform 1 | 8861 | 603451 |
| PPRC1 | chr10 | 103882776 | 103900078 | peroxisome proliferator-activated receptor | 23082 | 0 |
| NOLC1 | chr10 | 103901922 | 103913617 | nucleolar and coiled-body phosphoprotein 1 | 9221 | 602394 |
| ELOVL3 | chr10 | 103976132 | 103979332 | elongation of very long chain fatty acids like | 83401 | 611815 |
| PITX3 | chr10 | 103979935 | 103991221 | paired-like homeodomain 3 | 5309 | 602669 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GBF1 | chr10 | 103995298 | 104132639 | golgi-specific brefeldin A resistance factor 1 | 8729 | 603698 |
| NFKB2 | chr10 | 104144218 | 104152270 | nuclear factor of kappa light polypeptide gene | 4791 | 164012 |
| PSD | chr10 | 104152365 | 104168891 | pleckstrin and Sec7 domain containing | 5662 | 602327 |
| FBXL15 | chr10 | 104169560 | 104172883 | F-box and leucine-rich repeat protein 15 | 79176 | 610287 |
| CUEDC2 | chr10 | 104172991 | 104182413 | CUE domain containing 2 | 79004 | 0 |
| C10orf95 | chr10 | 104199583 | 104201290 | hypothetical protein LOC79946 | 79946 | 0 |
| TMEM180 | chr10 | 104211159 | 104226790 | transmembrane protein 180 | 79847 | 0 |
| ACTR1A | chr10 | 104228976 | 104252452 | ARP1 actin-related protein 1 homolog A, | 10121 | 605143 |
| SUFU | chr10 | 104253753 | 104383199 | suppressor of fused | 51684 | 607035 |
| TRIM8 | chr10 | 104394241 | 104408066 | tripartite motif-containing 8 | 81603 | 606125 |
| ARL3 | chr10 | 104423478 | 104464180 | ADP-ribosylation factor-like 3 | 403 | 604695 |
| SFXN2 | chr10 | 104464287 | 104488932 | sideroflexin 2 | 118980 | 0 |
| C10orf26 | chr10 | 104493716 | 104566011 | hypothetical protein LOC54838 isoform 1 | 54838 | 611129 |
| CYP17A1 | chr10 | 104580279 | 104587280 | cytochrome P450, family 17 | 1586 | 609300 |
| C10orf32 | chr10 | 104604008 | 104613960 | hypothetical protein LOC119032 | 119032 | 0 |
| AS3MT | chr10 | 104619199 | 104651643 | arsenic (+3 oxidation state) methyltransferase | 57412 | 611806 |
| C10orf122 | chr10 | 127334252 | 127361703 | hypothetical protein LOC387718 | 387718 | 0 |
| C10orf137 | chr10 | 127398073 | 127442702 | erythroid differentiation-related factor 1 | 26098 | 0 |
| MMP21 | chr10 | 127445015 | 127454380 | matrix metalloproteinase 21 preproprotein | 118856 | 608416 |
| UROS | chr10 | 127467136 | 127501827 | uroporphyrinogen III synthase | 7390 | 606938 |
| BCCIP | chr10 | 127502093 | 127532254 | BRCA2 and CDKN1A-interacting protein isoform | 56647 | 611883 |
| DHX32 | chr10 | 127514898 | 127559874 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 32 | 55760 | 607960 |
| FANK1 | chr10 | 127575097 | 127688151 | fibronectin type III and ankyrin repeat domains | 92565 | 611640 |
| ADAM12 | chr10 | 127693414 | 128067055 | ADAM metallopeptidase domain 12 isoform 1 | 8038 | 602714 |
| MGMT | chr10 | 131155455 | 131455356 | O-6-methylguanine-DNA methyltransferase | 4255 | 156569 |
| GLRX3 | chr10 | 131824652 | 131867860 | glutaredoxin 3 | 10539 | 0 |
| SWAP70 | chr11 | 9642203 | 9731081 | SWAP-70 protein | 23075 | 604762 |
| SBF2 | chr11 | 9756789 | 10272330 | SET binding factor 2 | 81846 | 607697 |
| GALNTL4 | chr11 | 11249001 | 11600128 | UDP-N-acetyl-alpha-D-galactosamine:polypeptide | 374378 | 0 |
| USP47 | chr11 | 11819545 | 11937446 | ubiquitin specific protease 47 | 55031 | 0 |
| DKK3 | chr11 | 11941120 | 11987205 | dickkopf homolog 3 precursor | 27122 | 605416 |
| MICAL2 | chr11 | 12088713 | 12241905 | microtubule associated monoxygenase, calponin | 9645 | 608881 |
| MICALCL | chr11 | 12265022 | 12337267 | MICAL C-terminal like | 84953 | 0 |
| PARVA | chr11 | 12355721 | 12507986 | parvin, alpha | 55742 | 608120 |
| TEAD1 | chr11 | 12652544 | 12922875 | TEA domain family member 1 | 7003 | 189967 |

| | | | | | | |
|---|---|---|---|---|---|---|
| RASSF10 | chr11 | 12987271 | 12989223 | Ras association (RalGDS/AF-6) domain family | 644943 | 0 |
| ARNTL | chr11 | 13255900 | 13365386 | aryl hydrocarbon receptor nuclear | 406 | 602550 |
| BTBD10 | chr11 | 13366133 | 13441414 | K+ channel tetramerization protein | 84280 | 0 |
| PTH | chr11 | 13470176 | 13474143 | parathyroid hormone preproprotein | 5741 | 168450 |
| FAR1 | chr11 | 13646846 | 13710467 | fatty acyl CoA reductase 1 | 84188 | 0 |
| SPON1 | chr11 | 13940489 | 14246222 | spondin 1, extracellular matrix protein | 10418 | 604989 |
| RRAS2 | chr11 | 14256041 | 14336607 | related RAS viral (r-ras) oncogene homolog 2 | 22800 | 600098 |
| COPB1 | chr11 | 14435628 | 14477974 | coatomer protein complex, subunit beta | 1315 | 600959 |
| PSMA1 | chr11 | 14482998 | 14498524 | proteasome alpha 1 subunit isoform 2 | 5682 | 602854 |
| PDE3B | chr11 | 14621906 | 14848926 | phosphodiesterase 3B, cGMP-inhibited | 5140 | 602047 |
| CYP2R1 | chr11 | 14856132 | 14870327 | cytochrome P450, family 2, subfamily R, | 120227 | 608713 |
| CALCA | chr11 | 14944791 | 14950408 | calcitonin isoform CGRP preproprotein | 796 | 114130 |
| CALCB | chr11 | 15051721 | 15056753 | calcitonin-related polypeptide, beta | 797 | 114160 |
| INSC | chr11 | 15090545 | 15225328 | inscuteable isoform a | 387755 | 610668 |
| SOX6 | chr11 | 15948370 | 16380968 | SRY (sex determining region Y)-box 6 isoform 1 | 55553 | 607257 |
| C11orf58 | chr11 | 16716769 | 16734149 | small acidic protein | 10944 | 0 |
| PLEKHA7 | chr11 | 16765787 | 16992535 | pleckstrin homology domain containing, family A | 144100 | 0 |
| RPS13 | chr11 | 17052515 | 17055796 | ribosomal protein S13 | 6207 | 180476 |
| PIK3C2A | chr11 | 17064699 | 17147930 | phosphoinositide-3-kinase, class 2, alpha | 5286 | 603601 |
| NUCB2 | chr11 | 17254861 | 17309645 | nucleobindin 2 | 4925 | 608020 |
| KCNJ11 | chr11 | 17363373 | 17366782 | potassium inwardly-rectifying channel J11 | 3767 | 600937 |
| ABCC8 | chr11 | 17371007 | 17455025 | ATP-binding cassette, sub-family C, member 8 | 6833 | 600509 |
| USH1C | chr11 | 17472035 | 17522539 | harmonin isoform a | 10083 | 605242 |
| MYOD1 | chr11 | 17697685 | 17700253 | myogenic differentiation 1 | 4654 | 159970 |
| KCNC1 | chr11 | 17714070 | 17755953 | Shaw-related voltage-gated potassium channel | 3746 | 176258 |
| SERGEF | chr11 | 17766171 | 17991213 | deafness locus associated putative guanine | 26297 | 606051 |
| TPH1 | chr11 | 17999115 | 18018885 | tryptophan hydroxylase 1 | 7166 | 191060 |
| SAAL1 | chr11 | 18058466 | 18084214 | serum amyloid A-like 1 | 113174 | 0 |
| MRGPRX3 | chr11 | 18099077 | 18116601 | MAS-related GPR, member X3 | 117195 | 607229 |
| MRGPRX4 | chr11 | 18150959 | 18152403 | MAS-related GPR, member X4 | 117196 | 607230 |
| SAA4 | chr11 | 18209479 | 18214910 | serum amyloid A4, constitutive | 6291 | 104752 |
| SAA2 | chr11 | 18217163 | 18226758 | serum amyloid A2 isoform b | 6289 | 104751 |
| SAA1 | chr11 | 18244347 | 18248102 | serum amyloid A1 preproprotein | 6288 | 104750 |
| HPS5 | chr11 | 18256792 | 18300297 | Hermansky-Pudlak syndrome 5 isoform b | 11234 | 607521 |
| GTF2H1 | chr11 | 18300718 | 18345152 | general transcription factor IIH, polypeptide 1, | 2965 | 189972 |
| LDHA | chr11 | 18372557 | 18385974 | lactate dehydrogenase A | 3939 | 150000 |

| | | | | | | |
|---|---|---|---|---|---|---|
| LDHC | chr11 | 18390433 | 18429369 | lactate dehydrogenase C | 3948 | 150150 |
| LDHAL6A | chr11 | 18434006 | 18457723 | lactate dehydrogenase A-like 6A | 160287 | 0 |
| TSG101 | chr11 | 18458434 | 18505065 | tumor susceptibility gene 101 | 7251 | 601387 |
| UEVLD | chr11 | 18509820 | 18566857 | ubiquitin-conjugating enzyme E2-like isoform a | 55293 | 610985 |
| SPTY2D1 | chr11 | 18584524 | 18612596 | SPT2, Suppressor of Ty, domain containing 1 | 144108 | 0 |
| TMEM86A | chr11 | 18676926 | 18682908 | transmembrane protein 86A | 144110 | 0 |
| IGSF22 | chr11 | 18682427 | 18704353 | immunoglobulin superfamily, member 22 | 283284 | 0 |
| PTPN5 | chr11 | 18706053 | 18769965 | protein tyrosine phosphatase, non-receptor type | 84867 | 176879 |
| MRGPRX2 | chr11 | 19032580 | 19038804 | MAS-related GPR, member X2 | 117194 | 607228 |
| ZDHHC13 | chr11 | 19095267 | 19154542 | zinc finger, DHHC domain containing 13 isoform | 54503 | 0 |
| CSRP3 | chr11 | 19160153 | 19180165 | cysteine and glycine-rich protein 3 | 8048 | 600824 |
| E2F8 | chr11 | 19202193 | 19219083 | E2F family member 8 | 79733 | 612047 |
| NAV2 | chr11 | 19328846 | 20099723 | neuron navigator 2 isoform 3 | 89797 | 607026 |
| P53AIP1 | chr11 | 128310479 | 128318032 | p53-regulated apoptosis-inducing protein 1 | 63970 | 605426 |
| RICS | chr11 | 128343051 | 128399222 | Rho GTPase-activating protein | 9743 | 608541 |
| BARX2 | chr11 | 128751090 | 128827384 | BarH-like homeobox 2 | 8538 | 604823 |
| TAS2R7 | chr12 | 10845398 | 10846493 | taste receptor, type 2, member 7 | 50837 | 604793 |
| TAS2R8 | chr12 | 10849917 | 10850846 | taste receptor, type 2, member 8 | 50836 | 604794 |
| TAS2R9 | chr12 | 10852960 | 10854034 | taste receptor, type 2, member 9 | 50835 | 604795 |
| TAS2R10 | chr12 | 10869211 | 10870135 | taste receptor, type 2, member 10 | 50839 | 604791 |
| PRR4 | chr12 | 10889714 | 11215480 | proline rich 4 (lacrimal) isoform 1 | 11272 | 605359 |
| PRH1 | chr12 | 10924826 | 11215477 | proline-rich protein HaeIII subfamily 1 | 5554 | 168730 |
| TAS2R13 | chr12 | 10951791 | 10953428 | taste receptor, type 2, member 13 | 50838 | 604792 |
| PRH2 | chr12 | 10973100 | 10978711 | proline-rich protein HaeIII subfamily 2 | 5555 | 168790 |
| TAS2R14 | chr12 | 10982120 | 10983073 | taste receptor, type 2, member 14 | 50840 | 604790 |
| TAS2R50 | chr12 | 11029778 | 11030778 | taste receptor, type 2, member 50 | 259296 | 609627 |
| TAS2R49 | chr12 | 11039827 | 11041741 | taste receptor, type 2, member 49 | 259295 | 0 |
| TAS2R48 | chr12 | 11065538 | 11066437 | taste receptor, type 2, member 48 | 259294 | 0 |
| TAS2R44 | chr12 | 11074252 | 11075273 | taste receptor, type 2, member 44 | 259290 | 0 |
| TAS2R46 | chr12 | 11105230 | 11106160 | taste receptor, type 2, member 46 | 259292 | 0 |
| TAS2R43 | chr12 | 11135152 | 11136179 | taste receptor, type 2, member 43 | 259289 | 0 |
| TAS2R47 | chr12 | 11177150 | 11178110 | type 2 taste receptor member 47 | 259293 | 0 |
| TAS2R42 | chr12 | 11229865 | 11230810 | taste receptor, type 2, member 42 | 353164 | 0 |
| PRB3 | chr12 | 11310124 | 11313908 | proline-rich protein BstNI subfamily 3 | 5544 | 168840 |
| PRB4 | chr12 | 11351283 | 11354633 | proline-rich protein BstNI subfamily 4 | 5545 | 180990 |
| PRB1 | chr12 | 11396023 | 11399791 | proline-rich protein BstNI subfamily 1 isoform 1 | 5542 | 180989 |

| | | | | | | |
|---|---|---|---|---|---|---|
| PRB2 | chr12 | 11435742 | 11439768 | proline-rich protein BstNI subfamily 2 | 653247 | 168810 |
| ETV6 | chr12 | 11694054 | 11939592 | ets variant gene 6 | 2120 | 600618 |
| GRIN2B | chr12 | 13605676 | 14024289 | N-methyl-D-aspartate receptor subunit 2B | 2904 | 138252 |
| VDR | chr12 | 46521588 | 46585081 | vitamin D (1,25-dihydroxyvitamin D3) receptor | 7421 | 601769 |
| TMEM106C | chr12 | 46644258 | 46648927 | transmembrane protein 106C | 79022 | 0 |
| COL2A1 | chr12 | 46653015 | 46684552 | collagen, type II, alpha 1 isoform 2 precursor | 1280 | 120140 |
| SENP1 | chr12 | 46723023 | 46785908 | sentrin/SUMO-specific protease 1 | 29843 | 612157 |
| PFKM | chr12 | 46799294 | 46826154 | phosphofructokinase, muscle | 5213 | 610681 |
| ASB8 | chr12 | 46827841 | 46837644 | ankyrin repeat and SOCS box-containing 8 | 140461 | 0 |
| LOC387856 | chr12 | 46863665 | 46865976 | hypothetical protein LOC387856 | 387856 | 0 |
| OR10AD1 | chr12 | 46882388 | 46883342 | olfactory receptor, family 10, subfamily AD, | 121275 | 0 |
| H1FNT | chr12 | 47009029 | 47010328 | H1 histone family, member N, testis-specific | 341567 | 0 |
| ZNF641 | chr12 | 47022178 | 47030268 | zinc finger protein 641 | 121274 | 0 |
| ANP32D | chr12 | 47152714 | 47153109 | acidic nuclear phosphoprotein 32D | 23519 | 606878 |
| C12orf54 | chr12 | 47163289 | 47176432 | hypothetical protein LOC121273 | 121273 | 0 |
| OR8S1 | chr12 | 47205681 | 47208153 | olfactory receptor, family 8, subfamily S, | 341568 | 0 |
| KRT5 | chr12 | 51194627 | 51200510 | keratin 5 | 3852 | 148040 |
| KRT71 | chr12 | 51223960 | 51233170 | keratin 71 | 112802 | 608245 |
| KRT74 | chr12 | 51245869 | 51253876 | keratin 6 irs4 | 121391 | 608248 |
| KRT72 | chr12 | 51265639 | 51281559 | keratin 72 | 140807 | 608246 |
| KRT73 | chr12 | 51287621 | 51298610 | keratin 73 | 319101 | 608247 |
| KRT2 | chr12 | 51324609 | 51332226 | keratin 2 | 3849 | 600194 |
| KRT1 | chr12 | 51354786 | 51360458 | keratin 1 | 3848 | 139350 |
| KRT77 | chr12 | 51369676 | 51383514 | keratin 77 | 374454 | 611158 |
| NTS | chr12 | 84770542 | 84779231 | neurotensin/neuromedin N preproprotein | 4922 | 162650 |
| MGAT4C | chr12 | 84875504 | 85735149 | UDP-N-acetylglucosamine: alpha-1,3-D-mannoside | 25834 | 607385 |
| FRY | chr13 | 31503436 | 31768775 | furry homolog | 10129 | 0 |
| BRCA2 | chr13 | 31787616 | 31871809 | breast cancer 2, early onset | 675 | 600185 |
| N4BP2L1 | chr13 | 31872860 | 31900315 | NEDD4 binding protein 2-like 1 isoform 2 | 90634 | 0 |
| N4BP2L2 | chr13 | 31904929 | 32010936 | phosphonoformate immuno-associated protein 5 | 10443 | 0 |
| POU4F1 | chr13 | 78071232 | 78075696 | POU domain, class 4, transcription factor 1 | 5457 | 601632 |
| RNF219 | chr13 | 78086421 | 78131315 | ring finger protein 219 | 79596 | 0 |
| AKAP6 | chr14 | 31868273 | 32372018 | A-kinase anchor protein 6 | 9472 | 604691 |
| NPAS3 | chr14 | 32478209 | 33340702 | neuronal PAS domain protein 3 isoform 1 | 64067 | 609430 |
| PAX9 | chr14 | 36196532 | 36216762 | paired box 9 | 5083 | 167416 |

| | | | | | | |
|---|---|---|---|---|---|---|
| SLC25A21 | chr14 | 36218828 | 36711616 | solute carrier family 25 (mitochondrial | 89874 | 607571 |
| SEC23A | chr14 | 38570873 | 38642188 | SEC23-related protein A | 10484 | 610511 |
| SIP1 | chr14 | 38653238 | 38675926 | SMN-interacting protein 1 isoform alpha | 8487 | 602595 |
| TRAPPC6B | chr14 | 38686767 | 38709385 | trafficking protein particle complex 6B isoform | 122553 | 610397 |
| PNN | chr14 | 38714137 | 38722173 | pinin, desmosome associated protein | 5411 | 603154 |
| MIA2 | chr14 | 38772875 | 38792324 | melanoma inhibitory activity 2 | 117153 | 608001 |
| CTAGE5 | chr14 | 38804226 | 38890148 | CTAGE family, member 5 isoform 2 | 4253 | 602132 |
| FBXO33 | chr14 | 38936714 | 38971371 | F-box protein 33 | 254170 | 609103 |
| RPL10L | chr14 | 46189971 | 46190778 | ribosomal protein L10-like protein | 140801 | 0 |
| MDGA2 | chr14 | 46378577 | 47213738 | MAM domain containing 1 isoform 1 | 161357 | 611128 |
| NID2 | chr14 | 51541269 | 51605696 | nidogen 2 precursor | 22795 | 605399 |
| PTGDR | chr14 | 51804180 | 51813191 | prostaglandin D2 receptor | 5729 | 604687 |
| C14orf37 | chr14 | 57540561 | 57688600 | hypothetical protein LOC145407 | 145407 | 0 |
| ACTR10 | chr14 | 57736585 | 57772106 | uncharacterized hypothalamus protein HARP11 | 55860 | 0 |
| PSMA3 | chr14 | 57781345 | 57808479 | proteasome alpha 3 subunit isoform 1 | 5684 | 176843 |
| ARID4A | chr14 | 57834974 | 57910201 | retinoblastoma-binding protein 1 isoform II | 5926 | 180201 |
| UNQ9438 | chr14 | 57932396 | 57945172 | hypothetical protein LOC387990 | 387990 | 0 |
| TIMM9 | chr14 | 57945123 | 57963985 | translocase of inner mitochondrial membrane 9 | 26520 | 607384 |
| KIAA0586 | chr14 | 57964462 | 58085302 | talpid3 protein | 9786 | 610178 |
| DACT1 | chr14 | 58174509 | 58184789 | dapper 1 isoform 1 | 51339 | 607861 |
| C14orf83 | chr14 | 67006736 | 67051774 | hypothetical protein LOC161145 | 161145 | 0 |
| PLEKHH1 | chr14 | 67069760 | 67126008 | pleckstrin homology domain containing, family H | 57475 | 0 |
| PIGH | chr14 | 67125775 | 67136770 | phosphatidylinositol glycan anchor biosynthesis, | 5283 | 600154 |
| ARG2 | chr14 | 67156331 | 67188187 | arginase, type II precursor | 384 | 107830 |
| VTI1B | chr14 | 67187618 | 67211301 | vesicle transport through interaction with | 10490 | 603207 |
| RDH11 | chr14 | 67213270 | 67232263 | retinol dehydrogenase 11 | 51109 | 607849 |
| RDH12 | chr14 | 67238355 | 67270921 | retinol dehydrogenase 12 | 145226 | 608830 |
| ZFYVE26 | chr14 | 67282991 | 67353055 | zinc finger, FYVE domain containing 26 | 23503 | 612012 |
| RAD51L1 | chr14 | 67356261 | 68014562 | RAD51-like 1 isoform 2 | 5890 | 602948 |
| POMT2 | chr14 | 76811051 | 76856970 | protein O-mannosyltransferase | 29954 | 607439 |
| GSTZ1 | chr14 | 76857106 | 76867692 | glutathione transferase zeta 1 isoform 2 | 2954 | 603758 |
| TMED8 | chr14 | 76877867 | 76913149 | transmembrane emp24 protein transport domain | 283578 | 0 |
| C14orf174 | chr14 | 76913514 | 76927340 | hypothetical protein LOC161394 | 161394 | 0 |
| C14orf148 | chr14 | 76930176 | 76959132 | hypothetical protein LOC122945 isoform 1 | 122945 | 0 |
| C14orf133 | chr14 | 76962770 | 76993657 | hypothetical protein LOC63894 | 63894 | 0 |
| AHSA1 | chr14 | 76994154 | 77005560 | AHA1, activator of heat shock 90kDa protein | 10598 | 608466 |

| | | | | | | |
|---|---|---|---|---|---|---|
| THSD3 | chr14 | 77010492 | 77034912 | thrombospondin, type I domain containing 3 | 145501 | 0 |
| SPTLC2 | chr14 | 77043024 | 77152863 | serine palmitoyltransferase, long chain base | 9517 | 605713 |
| ALKBH1 | chr14 | 77208501 | 77244109 | alkylated DNA repair protein alkB homolog | 8846 | 605345 |
| C14orf156 | chr14 | 77244177 | 77253696 | SRA stem-loop-interacting RNA-binding protein | 81892 | 610211 |
| SNW1 | chr14 | 77253696 | 77297250 | SKI-interacting protein | 22938 | 603055 |
| C14orf178 | chr14 | 77296645 | 77305838 | hypothetical protein LOC283579 | 283579 | 0 |
| ADCK1 | chr14 | 77336226 | 77470047 | aarF domain containing kinase 1 | 57143 | 0 |
| SPRED1 | chr15 | 36332343 | 36436742 | sprouty-related protein 1 with EVH-1 domain | 161742 | 609291 |
| FAM98B | chr15 | 36533619 | 36561121 | family with sequence similarity 98, member B | 283742 | 0 |
| RASGRP1 | chr15 | 36567593 | 36644299 | RAS guanyl releasing protein 1 isoform b | 10125 | 603962 |
| MEX3B | chr15 | 80121182 | 80125416 | mex-3 homolog B | 84206 | 611008 |
| AKAP13 | chr15 | 83724874 | 84093590 | A-kinase anchor protein 13 isoform 2 | 11214 | 604686 |
| AGBL1 | chr15 | 84486245 | 85373287 | ATP/GTP binding protein-like 1 | 123624 | 0 |
| IQGAP1 | chr15 | 88732476 | 88846479 | IQ motif containing GTPase activating protein 1 | 8826 | 603379 |
| CRTC3 | chr15 | 88874201 | 88989580 | transducer of regulated CREB protein 3 isoform | 64784 | 608986 |
| BLM | chr15 | 89061582 | 89159688 | Bloom syndrome protein | 641 | 604610 |
| FURIN | chr15 | 89212888 | 89227691 | furin preproprotein | 5045 | 136950 |
| FES | chr15 | 89228712 | 89240008 | V-FES feline sarcoma viral/V-FPS fujinami avian | 2242 | 190030 |
| MAN2A2 | chr15 | 89248423 | 89266818 | mannosidase, alpha, class 2A, member 2 | 4122 | 600988 |
| UNC45A | chr15 | 89274413 | 89298326 | smooth muscle cell associated protein-1 isoform | 55898 | 611219 |
| HDDC3 | chr15 | 89275158 | 89276778 | HD domain containing 3 | 374659 | 0 |
| RCCD1 | chr15 | 89299109 | 89307357 | RCC1 domain containing 1 | 91433 | 0 |
| PRC1 | chr15 | 89310276 | 89338808 | protein regulator of cytokinesis 1 isoform 1 | 9055 | 603484 |
| VPS33B | chr15 | 89342777 | 89366837 | vacuolar protein sorting 33B (yeast homolog)) | 26276 | 608552 |
| HYDIN | chr16 | 69398789 | 69822070 | hydrocephalus inducing isoform a | 54768 | 610812 |
| FLJ11171 | chr16 | 69873703 | 69880816 | hypothetical protein LOC55783 | 55783 | 0 |
| CALB2 | chr16 | 69950126 | 69981843 | calbindin 2 isoform 1 | 794 | 114051 |
| ZNF23 | chr16 | 70039013 | 70053618 | zinc finger protein 23 | 7571 | 194527 |
| ZNF19 | chr16 | 70065476 | 70080755 | zinc finger protein 19 | 7567 | 194525 |
| CHST4 | chr16 | 70117561 | 70129914 | carbohydrate (N-acetylglucosamine 6-O) | 10164 | 0 |
| TAT | chr16 | 70158254 | 70168499 | tyrosine aminotransferase | 6898 | 276600 |
| MARVELD3 | chr16 | 70217570 | 70226732 | MARVEL domain containing 3 isoform 2 | 91862 | 0 |
| PHLPPL | chr16 | 70236354 | 70306205 | PH domain and leucine rich repeat protein | 23035 | 611066 |
| CA10 | chr17 | 47062672 | 47592160 | carbonic anhydrase X | 56934 | 604642 |
| CTAGE1 | chr18 | 18247561 | 18251876 | cutaneous T-cell lymphoma-associated antigen 1 | 64693 | 608856 |
| RBBP8 | chr18 | 18767292 | 18860443 | retinoblastoma binding protein 8 isoform a | 5932 | 604124 |

| | | | | | | |
|---|---|---|---|---|---|---|
| CABLES1 | chr18 | 18969724 | 19094432 | Cdk5 and Abl enzyme substrate 1 isoform 2 | 91768 | 609194 |
| C18orf45 | chr18 | 19129978 | 19271923 | hypothetical protein LOC85019 | 85019 | 0 |
| RIOK3 | chr18 | 19286784 | 19317097 | sudD suppressor of bimD6 homolog | 8780 | 603579 |
| C18orf8 | chr18 | 19337459 | 19365740 | colon cancer-associated protein Mic1 | 29919 | 0 |
| NPC1 | chr18 | 19365460 | 19420468 | Niemann-Pick disease, type C1 precursor | 4864 | 607623 |
| ANKRD29 | chr18 | 19433975 | 19496847 | ankyrin repeat domain 29 | 147463 | 0 |
| LAMA3 | chr18 | 19523559 | 19789025 | laminin alpha 3 subunit isoform 1 | 3909 | 600805 |
| TTC39C | chr18 | 19826734 | 19966989 | tetratricopeptide repeat domain 39C | 125488 | 0 |
| CABYR | chr18 | 19972952 | 19995561 | calcium-binding tyrosine | 26256 | 612135 |
| OSBPL1A | chr18 | 19996009 | 20231788 | oxysterol-binding protein-like 1A isoform B | 114876 | 606730 |
| IMPACT | chr18 | 20260606 | 20287492 | Impact homolog | 55364 | 0 |
| HRH4 | chr18 | 20294590 | 20313918 | histamine H4 receptor | 59340 | 606792 |
| ZNF521 | chr18 | 20895888 | 21186114 | zinc finger protein 521 | 25925 | 610974 |
| AQP4 | chr18 | 22686004 | 22699714 | aquaporin 4 isoform a | 361 | 600308 |
| CHST9 | chr18 | 22749613 | 23019279 | GalNAc-4-sulfotransferase 2 | 83539 | 610191 |
| CDH2 | chr18 | 23784933 | 24011189 | cadherin 2, type 1 preproprotein | 1000 | 114020 |
| DSC3 | chr18 | 26824049 | 26876779 | desmocollin 3 isoform Dsc3b preproprotein | 1825 | 600271 |
| DSC2 | chr18 | 26899939 | 26936386 | desmocollin 2 isoform Dsc2b preproprotein | 1824 | 125645 |
| DSC1 | chr18 | 26963213 | 26996817 | desmocollin 1 isoform Dsc1b preproprotein | 1823 | 125643 |
| DSG1 | chr18 | 27152049 | 27191391 | desmoglein 1 preproprotein | 1828 | 125670 |
| DSG4 | chr18 | 27210737 | 27247877 | desmoglein 4 | 147409 | 607892 |
| DSG3 | chr18 | 27281729 | 27312663 | desmoglein 3 preproprotein | 1830 | 169615 |
| DSG2 | chr18 | 27332024 | 27382812 | desmoglein 2 preproprotein | 1829 | 125671 |
| TTR | chr18 | 27425837 | 27432794 | transthyretin | 7276 | 176300 |
| B4GALT6 | chr18 | 27456206 | 27518684 | beta-1,4-galactosyltransferase 6 | 9331 | 604017 |
| MCART2 | chr18 | 27593656 | 27594841 | mitochondrial carrier triple repeat 2 | 147407 | 0 |
| KIAA1012 | chr18 | 27663133 | 27777089 | hypothetical protein LOC22878 | 22878 | 0 |
| RNF125 | chr18 | 27852442 | 27907152 | ring finger protein 125 | 54941 | 610432 |
| RNF138 | chr18 | 27925834 | 27965521 | ring finger protein 138 isoform 1 | 51444 | 0 |
| MEP1B | chr18 | 28023984 | 28054364 | meprin A, beta | 4225 | 600389 |
| FAM59A | chr18 | 28101476 | 28304445 | hypothetical protein LOC64762 | 64762 | 0 |
| C18orf34 | chr18 | 28771363 | 29274043 | hypothetical protein LOC374864 isoform 1 | 374864 | 0 |
| ASXL3 | chr18 | 29412538 | 29581397 | additional sex combs like 3 | 80816 | 0 |
| NOL4 | chr18 | 29685061 | 30057444 | nucleolar protein 4 | 8715 | 603577 |
| DTNA | chr18 | 30327251 | 30725806 | dystrobrevin alpha isoform 2 | 1837 | 601239 |
| MAPRE2 | chr18 | 30875388 | 30975813 | microtubule-associated protein, RP/EB family, | 10982 | 605789 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ZNF397 | chr18 | 31075016 | 31092356 | zinc finger protein 397 | 84307 | 609601 |
| ZNF397OS | chr18 | 31086816 | 31124170 | zinc finger protein 397 opposite strand | 100101467 | 0 |
| ZNF271 | chr18 | 31124264 | 31144728 | zinc finger protein 271 | 10778 | 604754 |
| ZNF24 | chr18 | 31166175 | 31178424 | zinc finger protein 24 | 7572 | 194534 |
| ZNF396 | chr18 | 31200658 | 31211299 | zinc finger protein 396 | 252884 | 609600 |
| C18orf37 | chr18 | 31302288 | 31331953 | Ies6-similar protein isoform 1 | 125476 | 0 |
| GALNT1 | chr18 | 31488593 | 31545792 | polypeptide N-acetylgalactosaminyltransferase 1 | 2589 | 602273 |
| C18orf21 | chr18 | 31806585 | 31813238 | chromosome 18 open reading frame 21 | 83608 | 0 |
| P15RS | chr18 | 31823789 | 31901371 | cyclin-dependent kinase 2B-inhibitor-related | 55197 | 610347 |
| SLC39A6 | chr18 | 31942491 | 31963355 | solute carrier family 39 (zinc transporter), | 25800 | 608731 |
| ELP2 | chr18 | 31963884 | 32008605 | elongator protein 2 | 55250 | 0 |
| MOCOS | chr18 | 32021477 | 32102683 | molybdenum cofactor sulfurase | 55034 | 0 |
| FHOD3 | chr18 | 32131699 | 32614016 | formin homology 2 domain containing 3 | 80206 | 609691 |
| C18orf10 | chr18 | 32630032 | 32663156 | tubulin polyglutamylase complex subunit 2 | 25941 | 0 |
| KIAA1328 | chr18 | 32663077 | 33059285 | hypothetical protein LOC57536 | 57536 | 0 |
| BRUNOL4 | chr18 | 33077005 | 33399998 | bruno-like 4, RNA binding protein isoform 4 | 56853 | 0 |
| RIT2 | chr18 | 38577189 | 38949655 | Ras-like without CAAX 2 | 6014 | 609592 |
| SYT4 | chr18 | 39101854 | 39111342 | synaptotagmin IV | 6860 | 600103 |
| SLC14A2 | chr18 | 41448763 | 41517058 | solute carrier family 14 (urea transporter), | 8170 | 601611 |
| PSG2 | chr19 | 48260201 | 48278654 | pregnancy specific beta-1-glycoprotein 2 | 5670 | 176391 |
| PSG5 | chr19 | 48363734 | 48382528 | pregnancy specific beta-1-glycoprotein 5 | 5673 | 176394 |
| PSG4 | chr19 | 48388695 | 48401630 | pregnancy specific beta-1-glycoprotein 4 isoform | 5672 | 176393 |
| PSG9 | chr19 | 48449274 | 48465522 | pregnancy specific beta-1-glycoprotein 9 | 5678 | 176398 |
| CD177 | chr19 | 48549664 | 48559319 | CD177 molecule | 57126 | 162860 |
| TEX101 | chr19 | 48584602 | 48614607 | testis expressed 101 isoform 1 | 83639 | 0 |
| LYPD3 | chr19 | 48656785 | 48661671 | GPI-anchored metastasis-associated protein | 27076 | 609484 |
| PHLDB3 | chr19 | 48671094 | 48700825 | pleckstrin homology-like domain, family B, | 653583 | 0 |
| ETHE1 | chr19 | 48702711 | 48723236 | ETHE1 protein | 23474 | 608451 |
| ZNF575 | chr19 | 48729168 | 48732122 | zinc finger protein 575 | 284346 | 0 |
| XRCC1 | chr19 | 48739303 | 48771570 | X-ray repair cross complementing protein 1 | 7515 | 194360 |
| IRGQ | chr19 | 48787865 | 48791516 | immunity-related GTPase family, Q | 126298 | 0 |
| ZNF576 | chr19 | 48792383 | 48795995 | zinc finger protein 576 | 79177 | 0 |
| ZNF428 | chr19 | 48803216 | 48815846 | zinc finger protein 428 | 126299 | 0 |
| CADM4 | chr19 | 48818361 | 48835831 | cell adhesion molecule 4 | 199731 | 609744 |
| PLAUR | chr19 | 48842087 | 48866342 | plasminogen activator, urokinase receptor | 5329 | 173391 |
| IRGC | chr19 | 48912077 | 48916007 | immunity-related GTPase family, cinema | 56269 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| C19orf61 | chr19 | 48927140 | 48950982 | hypothetical protein LOC56006 | 56006 | 0 |
| KCNN4 | chr19 | 48962526 | 48977249 | intermediate conductance calcium-activated | 3783 | 602754 |
| LYPD5 | chr19 | 48991918 | 48998453 | LY6/PLAUR domain containing 5 isoform A | 284348 | 0 |
| ZNF283 | chr19 | 49023312 | 49044890 | zinc finger protein 283 | 284349 | 0 |
| ZNF404 | chr19 | 49068354 | 49076128 | zinc finger protein 404 | 342908 | 0 |
| ZNF45 | chr19 | 49108620 | 49121398 | zinc finger protein 45 | 7596 | 194554 |
| TOB2 | chr22 | 40153991 | 40167527 | transducer of ERBB2, 2 | 10766 | 607396 |
| PHF5A | chr22 | 40180221 | 40189208 | PHD-finger 5A | 84844 | 0 |
| ACO2 | chr22 | 40189628 | 40249492 | aconitase 2 precursor | 50 | 100850 |
| POLR3H | chr22 | 40246307 | 40264964 | polymerase (RNA) III (DNA directed) polypeptide | 171568 | 0 |
| CSDC2 | chr22 | 40281513 | 40297170 | RNA-binding protein pippin | 27254 | 0 |
| PMM1 | chr22 | 40297389 | 40310371 | phosphomannomutase 1 | 5372 | 601786 |
| FAM152B | chr22 | 40321435 | 40341416 | hypothetical protein LOC27351 | 27351 | 0 |
| XRCC6 | chr22 | 40341794 | 40384552 | ATP-dependent DNA helicase II, 70 kDa subunit | 2547 | 152690 |
| NHP2L1 | chr22 | 40394437 | 40403056 | NHP2 non-histone chromosome protein 2-like 1 | 4809 | 601304 |
| MEI1 | chr22 | 40420017 | 40519959 | meiosis defective 1 | 150365 | 608797 |
| CCDC134 | chr22 | 40521177 | 40546803 | coiled-coil domain containing 134 | 79879 | 0 |
| SREBF2 | chr22 | 40553605 | 40626873 | sterol regulatory element-binding transcription | 6721 | 600481 |
| TNFRSF13C | chr22 | 40645535 | 40647321 | BAFF receptor | 115650 | 606269 |
| CENPM | chr22 | 40659240 | 40660723 | centromere protein M isoform c | 79019 | 610152 |
| 03-set | chr22 | 40697430 | 40718725 | septin 3 isoform A | 55964 | 608314 |
| WBP2NL | chr22 | 40719291 | 40748975 | WBP2 N-terminal like | 164684 | 610981 |
| NAGA | chr22 | 40778837 | 40791346 | alpha-N-acetylgalactosaminidase precursor | 4668 | 104170 |
| FAM109B | chr22 | 40794754 | 40799941 | hypothetical protein LOC150368 | 150368 | 0 |
| C22orf32 | chr22 | 40800198 | 40804787 | hypothetical protein LOC91689 | 91689 | 0 |
| NDUFA6 | chr22 | 40806030 | 40811388 | NADH dehydrogenase (ubiquinone) 1 alpha | 4700 | 602138 |
| CYP2D6 | chr22 | 40847000 | 40851381 | cytochrome P450, family 2, subfamily D, | 1565 | 124030 |
| TCF20 | chr22 | 40880516 | 40935943 | transcription factor 20 isoform 1 | 6942 | 603107 |
| NFAM1 | chr22 | 41100911 | 41152899 | NFAT activation molecule 1 precursor | 150372 | 608740 |
| SERHL | chr22 | 41221093 | 41233063 | serine hydrolase-like | 94009 | 607979 |
| CGI-96 | chr22 | 41233072 | 41240306 | hypothetical protein LOC27341 | 27341 | 0 |
| SERHL2 | chr22 | 41274422 | 41294885 | serine hydrolase-like 2 | 253190 | 0 |
| POLDIP3 | chr22 | 41304225 | 41335460 | DNA polymerase delta interacting protein 3 | 84271 | 611520 |
| CYB5R3 | chr22 | 41339311 | 41369903 | cytochrome b5 reductase isoform m | 1727 | 250800 |
| MPPED1 | chr22 | 42132517 | 42227696 | metallophosphoesterase domain containing 1 | 758 | 602112 |

| | | | | | | |
|---|---|---|---|---|---|---|
| EFCAB6 | chr22 | 42249554 | 42533019 | CAP-binding protein complex interacting protein | 64800 | 0 |
| SULT4A1 | chr22 | 42545287 | 42583279 | sulfotransferase family 4A, member 1 | 25830 | 608359 |

**Table 16. Fingerprint ROHs gene annotation.** The list of genes associated to ROHs was obtained using the UCSC Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables) and by querying the hg17 database.

# DISCUSSION

Current genome-wide high density mapping technologies have improved the number of feasible applications of SNPs genotyping in the identification of new disease associated loci using genotype calls not only for association studies but also for loss of heterozygosity analysis and the subsequence LOH traits classification as structural variations (deletion, UPD, autozygosity), thanks to the contemporary analysis of copy number. This methodological approach to the analysis of SNPs genotyping data has been called "homozygosity mapping", which was useful in the identification of disease associated loci in probably inbred population like Aborigines (Li, Ho et al. 2006) as well as in outbreed population like that represented by Hap Map samples (Gibson, Morton et al. 2006). These studies underlined that these long and contiguous traits of homozygosity were not generally due to uniparental dysomy (that is a rare event and associated to particular syndromes in which region of imprinted region or cancer predisposing genes are involved), but they were autozygosity traits found in region of broad linkage disequilibrium and low frequency of recombination, probably due to presence of a common ancestor, or to the nature of particular chromosome traits allowing them to remain unchanged .

Homozygosity mapping focused on individuation of regions of autozygosity which were in common between affected individuals of the same family or between affected unrelated probands with the same disease which had high probability to harbor new mutations , and could underline regions that could lead to a new gene or to a number of candidate loci as responsible of the genetic basis of the disease. In order to identify new disease susceptibility loci in a cohort of 34 CHI consanguineous and non consanguineous families, we performed genome-wide homozygosity mapping using 250K NspI Gene Chip Affymetrix SNP microarrays.

The first step of our homozygosity mapping method was single sample LOH analysis in all the 110 components of the 34 CHI families, which not only produced a list of homozygosity regions in this disease cohort butled us to investigate the nature of this LOH traits. Copy number and aplotype analysis in LOH traits found in all CHI families' affected members excluded deletions or UPD to be at the molecular basis of the disease and has confirmed a relevant presence of autozygosity traits in CHI patients and relatives too. The autozygosity nature of these homozygosity regions was confirmed by the presence of declared inbreeding in at least 6 of the analyzed families. In other families the high degree of homozygosity was probably due to restricted geographical provenience or distant unknown parental relationship.

We focused genome-wide SNP analysis performed using Affymetrix GeneChip® 250K Nsp I Human Mapping on the presence of the largest (or second largest) homozygous segments of the patient's genotype which have the major probability to bear the disease causing mutation (Knight, Maclean et al. 2008) ) and

we evidenced in one of the 35 CHI proband a homozygous region of 79Mb on chromosome 4 (from 4q13.2 to 4q31.21) that included 312 known genes in the region. This stretch contains the causative HADHSC gene , which is genotyped on the microarray by 7 SNPs covering 45 kb. The child was born from blood-related parents (double first cousins) and at 2 months he presented convulsions with deep hypoglycemia; CHI with diffuse disease was diagnosed but mutational analysis by direct sequencing of ABCC8 and KCNJ11 genes was negative for disease causing variations in coding regions of both genes. Diagnostic evalutation of urine metabolites showed an elevated excretion of 3-OH butyryc, piruvic, succinic, etilmalonic, fumaric, glutaric and 2- ketoglutaric acid; this metabolic profile was compatible with a slow down of β -oxidation at level of SCHAD , leading to increased accumulation of medium and long chain L-3-OHCoA species, promoting a reverse flux through SCHAD converting acetoacetylCoA into L-3-OHbutyrylCoA. Direct sequencing of the coding region of HADHSC gene was performed and showed a homozygous C to T transition in exon 6 (c.706C>T) causing a premature stop of the synthesized protein at codon 236 , resulting in the not previously described nonsense mutation R236X. The finding of HADHSC mutation confirmed the power of SNPs genome-wide genotyping technologies to underline loci with high mutational probability and to elucidate the molecular basis of autosomic recessive diseases; mutation analysis of exon 6 of the HADHSC gene in the parents of the proband showed they were heterozygous carriers of the same mutation.

SCHAD protein exhibits a two-domain topology, with the N- terminal domain of the monomer (residues 12-201) similar to a NAD(P)+ -binding enzyme and the C- terminal domain ( residues 207-302 ) involved in subunit dimerization and, presumably, binding of acylCoA substrate. Homology modeling bioinformatic analysis of the protein bearing the nonsense mutation R236X led to a truncated protein lacking the most part of the C- terminal domain, particularly relevant to the closing of the active site cleft, and probably occurring in a unfunctional protein. Enzymatic assays would be necessary to elucidate the function of the truncated protein. We further investigated if SCHAD truncated protein mRNA was expressed in the patient's blood lymphocytes or recognized and suppressed by RNA nonsense-mediated decay, and which was the level of expression in thehomozygous proband and in heterozygous mother, comparing them with healthy controls .

We confirmed the amplification of SCHAD cDNA in both homozygous and etherozygous R236X carriers, even if with slightly decrease levels of expression compared with the healthy controls. The transcrition/ translation of the truncated protein is still under study.

So far 3 CHI patients with mutations in the HADH gene have been reported (Clayton, Eaton et al. 2001; Molven, Matre et al. 2004; Hussain, Clayton et al. 2005; Knight, Maclean et al. 2008): a missense mutation P258L, a six base pair deletion at the start of exon 5 and a splice site mutation IV S6-2.The acylcarnitine profile in all reported patients has demonstrated raised levels, and urinary organic acids showed raised 3-hydroxyglutarate concentrations together with highly elevated levels of 3-hydroxybutyrate, and 3-hydroxydicarbonylic acids of chain lengths $C_6$-$C_{14.}$ These loss of function of SCHAD proteins promoted an

increase in basal insulin secretion through a $K_{ATP}$ channel independent mechanism (Hardy, Hohmeier et al. 2007), probably by inhibition of carnitine palmitoyltrasferase I due to accumulation of short chain acyl CoA esters. This is the one of the supposed mechanisms of how defects in SCHAD enzyme could modulate β-oxidation metabolites production and insulin regulation and release (Palladino, Bennett et al. 2008). Our patient is the first case of a CHI proband bearing a truncated coded SCHAD enzyme and in this case hypoglycaemic screens neither evidenced specific increase of plasma hydroxybutyrylcarnitine nor urine 3-hydroxyglutarate, while elevated levels of 3-hydroxybutyrate, and 3-hydroxydicarbonylic acids was evidenced like in the previously described CHI patients. This metabolic profile is compatible with a slow down of β oxidation at level of HADH, leading to increased accumulation of medium and long chain L-3-OHCoA species, promoting a reverse flux through HADHSC converting acetoacetylCoA into L-3-OhbutyrylCoA. This case highlights the opportunity of HADH gene mutation screening in those diffuse CHI probands when KCNJ11 and ABCC88 genes screening is negative.

Additional strong indication came from a patient bearing a long autozygosity trait on chromosome 1 (25.63 Mb) including MCAD gene, a medium chain acyl-CoA dehydrogenase which catalyzes the rate-determining step of fatty acid β- oxidation in mitochondria and identified as the commonest inborn error of metabolism (Kompare and Rizzo 2008). Mutational screening of this gene in the proband revealed a previously undescribed homozygous synonymous variation (P195P) which could bear to a non canonical acceptor splice site and lead to a new alternative spliced isoform of MCAD protein. Biochemical screening of the proband highlighted hypoketotic hypoglicemia, organic urinary acid increased escretion (2- ketoglutaric acid two folds the normal value) and a plasmatic ratio lactate/piruvate of 36 (instead of 10), clearly identified a fatty acid oxidation unbalance. Even though acylcarnitines analysis were not available for this patient to confirm MCAD enzyme to be responsible for the metabolic disease , we could assume in this proband a correlation between disregulated insulin secretion and β-oxidation metabolic impairment. Further investigations on P195P mRNA transcrition and translation would determined the impact of this MCAD gene defect on the enzyme availability and functional studies would elucidate if this patient could be the first description of MCAD- HI .

Although each patient showed an homozygosity profile strictly peculiar, our ROHs computational approach was successful to reveal the number and the nature of homozygosity traits in common between probands , which could underline chromosomal regions containing candidate genes/loci involved in the etiopathogenesis of CHI. When we looked at the most involved chromosomal loci, considering those in which the larger amount of patients were homozygous (5 probands on chromosome 11 in our CHI dataset), we confirmed 11 p 15.2 to be the most involved in the disease .

Furthermore, 3 of these 5 probands had a common ROHs overlapping the 2 genes more deeply involved in CHI and subsequent screening for mutations in ABCC8 and KCNJ11 known causatives genes in the 2 CHI probands who shared a common homozygosity traits on chromosome 11p15.1 led to the discovery of the

disease causing mutation . These findings also supported the strong capability of homozygosity mapping methodology to elucidate disease causing mutations in autosomic recessive disease , especially when the homozygous trait is present in a consistent part of the clinical dataset (more then 10 %).

In order to realize a genomic fingerprint of our CHI dataset, we used dChip LOH visualization option to match the single patient whole –genome ROH pattern with the 164 homozygous regions in common between at least two patients, evidencing which proband was bearing the CHI related regions and on which chromosomal location . This genomic pattern analysis brought to light that the more homozygous and inbreeding probands visually clustered together and that their longest and contiguous traits on chromosome 1, 4, 6, 11, 13, 14 and 18 had shorter regions in common with al lest another patient , indicating that this loci are the more involved in the disease. When we further characterize the genomic fingerprint plotting on the single patient whole –genome ROH pattern the list of 35 homozygous region in common between 3 patients , we could redefine the putative CHI chromosomal loci as 1p31.2-31.1, 4q27-28.1, 6p22.3-22.31, 11p15.4-15.2, and 18q12.2.

Considering genes which were present in homozygosity traits in common between in more then 2 CHI patients we have evidenced numerous genes involved in β-oxidation metabolism which could support a relationship between disregulated insulin secretion and fatty acid metabolis; this pathway was present with 7 genes including MCAD and ACSL6 (acyl CoA synthetase long chain, family member 6), the second one responsible for activation of long-chain fatty acids for both synthesis and degradation of cellular lipids via beta-oxidation.

A second methabolic pathway related to the CHI was type I diabetes mellitus, which was present with mayor histocompatibility complex class I and II genes , tumor necrosis factor 2 (or TNF-□) and perforine 1.TNF2 gene encodes a multifunctional proinflammatory cytokine that belongs to the tumor necrosis factor (TNF) superfamily. This cytokine is mainly secreted by macrophages and is involved in the regulation of a wide spectrum of biological processes including cell proliferation, differentiation, apoptosis, lipid metabolism, and coagulation. This cytokine has been implicated in a variety of diseases, including autoimmune diseases, insulin resistance, and cancer. There is substantial evidence that TNF-□ contributes to insulin resistance and thus to type 2 diabetes. Long-term exposure of cultured cells to TNF-□ induces insulin resistance. TNF-□ inhibits insulin receptor signaling by decreasing autophosphorylation of insulin receptor and promoting serine phosphorylation of insulin receptor substrate proteins. TNF-□ impairs insulin secretion and in particular in pancreatic ß-cell lines TNF-□ decreased glucose-stimulated insulin secretion (Kubaszek, Pihlajamaki et al. 2003); upregulation of this gene could lead to enhnaced insulin secretion , which is the major characteristic of CHI. A strong relationship between insulin secretion and glycolysis /gluconeogenesis was evidenced even when we considered genes involved in the glycolysis /gluconeogenesis pathway, that were the alcohol dehydrogenase family (ADH), phosphofructokinase (PFKM) and lactate dehydrogenase (LDH).

In particular, lactate dehydrogenase activity in pancreatic β-cells is 100-fold lower than other mammalian cell type because enhanced glycolytic flux and lactate output observed in other mammalian tissues during anaerobiosis should be avoided in β-cells because responsible for glucose- stimulated gene expression. In pancreatic β-cells an increase in LDH expression may lead to reduced secretory response to glucose and such event might be involved in the pathogenesis of same forms of diabetes mellitus (Sekine, Cirulli et al. 1994); upregulation of LDH gene could even result in an enhanced insulin secretion and thus be related to the pathogenesis of CHI.

Homozygosity mapping analysis in 35 CHI probands also produced a complete list of genes included in ROHs in common between 4, 3 patients and in at least 2 patients in our dataset, and these genes would be the more characteristic of the disease. Mutational screening is strongly indicated for that genes which function implies a correlation between fatty-acid β- oxidation energy contribution and insulin secretion and release, or for that genes that are known to be associated with diabetes and to a disregulated glucose sensitivity and response. We can underline insulin-like 5 (INSL5) gene which encodes a protein with a classical signature of the insulin superfamily and Inositol 1,4,5-Triphosphate Receptor 3 (ITPR3) gene which mediates second messenger signaling by releasing $Ca^{++}$ from intracellular stores in response to inositol triphosphate. There are three ITPR subtypes, and these are widely expressed in many tissues, consistent with their role in mediating second messenger signaling. Each of the subtypes has a unique tissue-distribution profile. *ITPR3* mRNA and ITPR3 protein are rapidly upregulated in pancreatic β cells following stimulation with glucose, and the protein is then rapidly degraded in proteasomes (Qu, Marchand et al. 2008). Inositol 1,4,5-Triphosphate Receptor 3 was recently underlined as a risk factor for Type 1 Diabetes in Sweden. (Roach, Deutsch et al. 2006) Another candidate gene is BDH2 (3-hydroxybutyrate dehydrogenase, type 2), which has possible physiological role in cytosolic ketone body utilization, either as a secondary system for energy supply in starvation or to generate precursors for lipid and sterol synthesis (Guo, Lukacik et al. 2006). All these genes are good candidates for mutational screening in those CHI families in which the molecular basis of the disease was not clarified.

# CONCLUSIONS AND FUTURE PROSPECTS

To our knowledge this is the first study in which molecular mechanism of CHI disease was investigated with genome-wide homozygosity mapping method in consanguineous and not consanguineous CHI families. This methodology was highly successful in revealing the recessive disease causing mutation in four of the 35 patients; two of these mutations were in ABCC8 and KCNJ11 known causative genes , one was a novel mutation in HADH gene which could describe the fourth case of SCHAD-HI, and the last one was a putative splicing variant of MCAD gene that could indicate the first described case of MCAD-HI. These results confirmed how hyperinsulinism in CHI is strongly related to glycolisis/gluconeogenesis energy supply, to fatty acids β-oxidation balance and his relationship with insulin levels regulations. Furthermore, bioinformatics pathway analysis underlined how a huge number of genes in autozygosity traits shared by our 35 CHI probands covered the two major metabolic pathways of glycolisis/gluconeogenesis and fatty acid metabolism, indicating how it would be important for a correct diagnosis to consider CHI has as metabolic disease and to describe these patients with a metabolic profile which includes at least organic acids investigations and acilcarnitines characterizations, as suggested by Hussain K. Thanks to the increasing number of described SCHAD-HI, our study also indicated that mutations of HADH gene should be sought in hyperinsulinemic patients in whom diffuse form of CHI and autosomal recessive inheritance can be presumed, when KCNJ11 and ABCC8 genes mutational screening is negative.

Genome-wide homozygosity mapping in the present study has clearly evidenced 1p31.2-31.1, 4q27-28.1, 6p22.3-22.31, 11p15.4-15.2, and 18q12.2 chromosomal loci to be the most related to CHI disease. To better redefine that regions and to single out potential associated genes, we are planning to use SNPs genotypes calls generated from these 34 CHI family microarrays data to perform family based transmission disequilibrium analysis. On the other hand , we will increase the number of CHI families enrolled in the study in order to evidence by homozygosity mapping ROHs which are strictly peculiar of the disease and not common in the normal population.

This study and future works will be possible thanks to the collaboration within clinicians who enrol and follows the patients, paediatric endocrinology laboratory diagnostic mutational screening of the candidate genes, our microarray laboratory experimental and computational analysis, and of course thanks to the co-operation of patients and their families.

# BIBLIOGRAPHY

Ashcroft, F. M. (2005). "ATP-sensitive potassium channelopathies: focus on insulin secretion." J Clin Invest **115**(8): 2047-58.

Ashcroft, F. M. (2006). "K(ATP) channels and insulin secretion: a key role in health and disease." Biochem Soc Trans **34**(Pt 2): 243-6.

Bahi-Buisson, N., S. El Sabbagh, et al. (2008). "Myoclonic absence epilepsy with photosensitivity and a gain of function mutation in glutamate dehydrogenase." Seizure **17**(7): 658-664.

Barycki, J. J., L. K. O'Brien, et al. (1999). "Biochemical characterization and crystal structure determination of human heart short chain L-3-hydroxyacyl-CoA dehydrogenase provide insights into catalytic mechanism." Biochemistry **38**(18): 5786-98.

Barycki, J. J., L. K. O'Brien, et al. (2000). "Sequestration of the active site by interdomain shifting. Crystallographic and spectroscopic evidence for distinct conformations of L-3-hydroxyacyl-CoA dehydrogenase." J Biol Chem **275**(35): 27186-96.

Chiang, A. P., J. S. Beck, et al. (2006). "Homozygosity mapping with SNP arrays identifies TRIM32, an E3 ubiquitin ligase, as a Bardet-Biedl syndrome gene (BBS11)." Proc Natl Acad Sci U S A **103**(16): 6287-92.

Christesen, H. B., N. D. Tribble, et al. (2008). "Activating glucokinase (GCK) mutations as a cause of medically responsive congenital hyperinsulinism: prevalence in children and characterisation of a novel GCK mutation." Eur J Endocrinol **159**(1): 27-34.

Clayton, P. T., S. Eaton, et al. (2001). "Hyperinsulinism in short-chain L-3-hydroxyacyl-CoA dehydrogenase deficiency reveals the importance of beta-oxidation in insulin secretion." J Clin Invest **108**(3): 457-65.

Conrad, D. F., M. Jakobsson, et al. (2006). "A worldwide survey of haplotype variation and linkage disequilibrium in the human genome." Nat Genet **38**(11): 1251-60.

Curtis, D. (2007). "Extended homozygosity is not usually due to cytogenetic abnormality." BMC Genet **8**: 67.

Damaj, L., M. le Lorch, et al. (2008). "Chromosome 11p15 paternal isodisomy in focal forms of neonatal hyperinsulinism." J Clin Endocrinol Metab.

De Leon, D. D. and C. A. Stanley (2007). "Mechanisms of Disease: advances in diagnosis and treatment of hyperinsulinism in neonates." Nat Clin Pract Endocrinol Metab **3**(1): 57-68.

de Lonlay, P., V. Cormier-Daire, et al. (2002). "Facial appearance in persistent hyperinsulinemic hypoglycemia." Am J Med Genet **111**(2): 130-3.

de Lonlay, P., J. C. Fournet, et al. (2002). "Heterogeneity of persistent hyperinsulinaemic hypoglycaemia. A series of 175 cases." Eur J Pediatr **161**(1): 37-48.

de Lonlay, P., I. Giurgea, et al. (2004). "Hyperinsulinemic hypoglycemia in children." Ann Endocrinol (Paris) **65**(1): 96-8.

de Lonlay, P., I. Giurgea, et al. (2004). "Neonatal hypoglycaemia: aetiologies." Semin Neonatol **9**(1): 49-58.

de Lonlay, P., A. Simon-Carre, et al. (2006). "Congenital hyperinsulinism: pancreatic [18F]fluoro-L-dihydroxyphenylalanine (DOPA) positron emission tomography and immunohistochemistry study of DOPA decarboxylase and insulin secretion." J Clin Endocrinol Metab **91**(3): 933-40.

de Lonlay, P., G. Touati, et al. (2002). "Persistent hyperinsulinaemic hypoglycaemia." Semin Neonatol **7**(1): 95-100.

Delonlay, P., A. Simon, et al. (2007). "Neonatal hyperinsulinism: clinicopathologic correlation." Hum Pathol **38**(3): 387-99.

Derks, T. G., T. S. Boer, et al. (2008). "Neonatal screening for medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in The Netherlands: the importance of enzyme analysis to ascertain true MCAD deficiency." J Inherit Metab Dis **31**(1): 88-96.

Downing, M., N. J. Manning, et al. (2008). "Detection of urinary hexanoylglycine in the diagnosis of MCAD deficiency from newborn screening." J Inherit Metab Dis **31**(4): 550.

Fournet, J. C., C. Mayaud, et al. (2001). "Unbalanced expression of 11p15 imprinted genes in focal forms of congenital hyperinsulinism: association with a reduction to homozygosity of a mutation in ABCC8 or KCNJ11." Am J Pathol **158**(6): 2177-84.

Fournet, J. C., C. Mayaud, et al. (2000). "Loss of imprinted genes and paternal SUR1 mutations lead to focal form of congenital hyperinsulinism." Horm Res **53 Suppl 1**: 2-6.

Gibson, J., N. E. Morton, et al. (2006). "Extended tracts of homozygosity in outbred human populations." Hum Mol Genet **15**(5): 789-95.

Giurgea, I., C. Bellanne-Chantelot, et al. (2006). "Molecular mechanisms of neonatal hyperinsulinism." Horm Res **66**(6): 289-96.

Giurgea, I., D. Sanlaville, et al. (2006). "Congenital hyperinsulinism and mosaic abnormalities of the ploidy." J Med Genet **43**(3): 248-54.

Gloyn, A. L., E. R. Pearson, et al. (2004). "Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes." N Engl J Med **350**(18): 1838-49.

Gregersen, N., B. S. Andresen, et al. (2008). "Mitochondrial fatty acid oxidation defects-remaining challenges." J Inherit Metab Dis.

Grosse, S. D., M. J. Khoury, et al. (2006). "The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update." Genet Med **8**(4): 205-12.

Guo, K., P. Lukacik, et al. (2006). "Characterization of human DHRS6, an orphan short chain dehydrogenase/reductase enzyme: a novel, cytosolic type 2 R-beta-hydroxybutyrate dehydrogenase."

J Biol Chem **281**(15): 10291-7.

Hardy, O. T., M. Hernandez-Pampaloni, et al. (2007). "Accuracy of [18F]fluorodopa positron emission tomography for diagnosing and localizing focal congenital hyperinsulinism." J Clin Endocrinol Metab **92**(12): 4706-11.

Hardy, O. T., M. Hernandez-Pampaloni, et al. (2007). "Diagnosis and localization of focal congenital hyperinsulinism by 18F-fluorodopa PET scan." J Pediatr **150**(2): 140-5.

Hardy, O. T., H. E. Hohmeier, et al. (2007). "Functional genomics of the beta-cell: short-chain 3-hydroxyacyl-coenzyme A dehydrogenase regulates insulin secretion independent of K+ currents." Mol Endocrinol **21**(3): 765-73.

He, X. Y. and S. Y. Yang (2007). "3-hydroxyacyl-CoA dehydrogenase (HAD) deficiency replaces short-chain hydroxyacyl-CoA dehydrogenase (SCHAD) deficiency as well as medium- and short-chain hydroxyacyl-CoA dehydrogenase (M/SCHAD) deficiency as the consensus name of this fatty acid oxidation disorder." Mol Genet Metab **91**(2): 205-6.

Huqun, S. Izumi, et al. (2007). "Mutations in the SLC34A2 gene are associated with pulmonary alveolar microlithiasis." Am J Respir Crit Care Med **175**(3): 263-8.

Hussain, K. (2007). "Insights in congenital hyperinsulinism." Endocr Dev **11**: 106-21.

Hussain, K. (2008). "Diagnosis and management of hyperinsulinaemic hypoglycaemia of infancy." Horm Res **69**(1): 2-13.

Hussain, K. and A. Aynsley-Green (2003). "Hyperinsulinism in infancy: understanding the pathophysiology." Int J Biochem Cell Biol **35**(9): 1312-7.

Hussain, K., P. T. Clayton, et al. (2005). "Hyperinsulinism of infancy associated with a novel splice site mutation in the SCHAD gene." J Pediatr **146**(5): 706-8.

Hussain, K. and K. E. Cosgrove (2005). "From congenital hyperinsulinism to diabetes mellitus: the role of pancreatic beta-cell KATP channels." Pediatr Diabetes **6**(2): 103-13.

Hussain, K., K. E. Cosgrove, et al. (2005). "Hyperinsulinemic hypoglycemia in Beckwith-Wiedemann syndrome due to defects in the function of pancreatic beta-cell adenosine triphosphate-sensitive potassium channels." J Clin Endocrinol Metab **90**(7): 4376-82.

Hussain, K., M. Seppanen, et al. (2006). "The diagnosis of ectopic focal hyperinsulinism of infancy with [18F]-dopa positron emission tomography." J Clin Endocrinol Metab **91**(8): 2839-42.

Iafrate, A. J., L. Feuk, et al. (2004). "Detection of large-scale variation in the human genome." Nat Genet **36**(9): 949-51.

Kapoor, R. R., C. Gilbert, et al. (2008). "Congenital hyperinsulinism: [18F]DOPA PET/CT scan of a focal lesion in the head of the pancreas." Arch Dis Child Fetal Neonatal Ed **93**(2): F166.

Knight, H. M., A. Maclean, et al. (2008). "Homozygosity mapping in a family presenting with schizophrenia, epilepsy and hearing impairment." Eur J Hum Genet **16**(6): 750-8.

Kompare, M. and W. B. Rizzo (2008). "Mitochondrial fatty-acid oxidation disorders." Semin Pediatr Neurol **15**(3): 140-9.

Kubaszek, A., J. Pihlajamaki, et al. (2003). "Promoter polymorphisms of the TNF-alpha (G-308A) and IL-6 (C-174G) genes predict the conversion from impaired glucose tolerance to type 2 diabetes: the Finnish Diabetes Prevention Study." Diabetes **52**(7): 1872-6.

Lencz, T., C. Lambert, et al. (2007). "Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia." Proc Natl Acad Sci U S A **104**(50): 19942-7.

Li, L. H., S. F. Ho, et al. (2006). "Long contiguous stretches of homozygosity in the human genome." Hum Mutat **27**(11): 1115-21.

Maier, E. M., B. Liebl, et al. (2005). "Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency." Hum Mutat **25**(5): 443-52.

McCarroll, S. A., T. N. Hadnott, et al. (2006). "Common deletion polymorphisms in the human genome." Nat Genet **38**(1): 86-92.

Mitchell, G. A., N. Gauthier, et al. (2008). "Hereditary and acquired diseases of acyl-coenzyme A metabolism." Mol Genet Metab **94**(1): 4-15.

Molven, A., G. E. Matre, et al. (2004). "Familial hyperinsulinemic hypoglycemia caused by a defect in the SCHAD enzyme of mitochondrial fatty acid oxidation." Diabetes **53**(1): 221-7.

Otonkoski, T., C. Ammala, et al. (1999). "A point mutation inactivating the sulfonylurea receptor causes the severe form of persistent hyperinsulinemic hypoglycemia of infancy in Finland." Diabetes **48**(2): 408-15.

Palladino, A. A., M. J. Bennett, et al. (2008). "Hyperinsulinism in infancy and childhood: when an insulin level is not always enough." Clin Chem **54**(2): 256-63.

Qu, H. Q., L. Marchand, et al. (2008). "The association between type 1 diabetes and the ITPR3 gene polymorphism due to linkage disequilibrium with HLA class II." Genes Immun **9**(3): 264-6.

Rahier, J., Y. Guiot, et al. (2000). "Persistent hyperinsulinaemic hypoglycaemia of infancy: a heterogeneous syndrome unrelated to nesidioblastosis." Arch Dis Child Fetal Neonatal Ed **82**(2): F108-12.

Ramprasad, V. L., N. Soumittra, et al. (2008). "Identification of a novel splice-site mutation in the Lebercilin (LCA5) gene causing Leber congenital amaurosis." Mol Vis **14**: 481-6.

Ribeiro, M. J., N. Boddaert, et al. (2007). "The added value of [18F]fluoro-L-DOPA PET in the diagnosis of hyperinsulinism of infancy: a retrospective study involving 49 children." Eur J Nucl Med Mol Imaging **34**(12): 2120-8.

Ribeiro, M. J., N. Boddaert, et al. (2007). "Functional imaging of the pancreas: the role of [18F]fluoro-L-DOPA PET in the diagnosis of hyperinsulinism of infancy." Endocr Dev **12**: 55-66.

Roach, J. C., K. Deutsch, et al. (2006). "Genetic mapping at 3-kilobase resolution reveals inositol 1,4,5-

triphosphate receptor 3 as a risk factor for type 1 diabetes in Sweden." Am J Hum Genet **79**(4): 614-27.

Saudubray, J. M., P. de Lonlay, et al. (2000). "Genetic hypoglycaemia in infancy and childhood: pathophysiology and diagnosis." J Inherit Metab Dis **23**(3): 197-214.

Saudubray, J. M., M. C. Nassogne, et al. (2002). "Clinical approach to inherited metabolic disorders in neonates: an overview." Semin Neonatol **7**(1): 3-15.

Sebat, J., B. Lakshmi, et al. (2004). "Large-scale copy number polymorphism in the human genome." Science **305**(5683): 525-8.

Sekine, N., V. Cirulli, et al. (1994). "Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing." J Biol Chem **269**(7): 4895-902.

Shanik, M. H., Y. Xu, et al. (2008). "Insulin resistance and hyperinsulinemia: is hyperinsulinemia the cart or the horse?" Diabetes Care **31 Suppl 2**: S262-8.

Simon-Sanchez, J., S. Scholz, et al. (2007). "Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals." Hum Mol Genet **16**(1): 1-14.

Tuzun, E., A. J. Sharp, et al. (2005). "Fine-scale structural variation of the human genome." Nat Genet **37**(7): 727-32.

Valayannopoulos, V., S. Romano, et al. (2008). "What's new in metabolic and genetic hypoglycaemias: diagnosis and management." Eur J Pediatr **167**(3): 257-65.

Valayannopoulos, V., M. Vaxillaire, et al. (2007). "Coexistence in the same family of both focal and diffuse forms of hyperinsulinism." Diabetes Care **30**(6): 1590-2.

van de Bunt, M., E. L. Edghill, et al. (2008). "Gene duplications resulting in over expression of glucokinase are not a common cause of hypoglycaemia of infancy in humans." Mol Genet Metab **94**(2): 268-9.

van Hove, E. C., T. Hansen, et al. (2006). "The HADHSC gene encoding short-chain L-3-hydroxyacyl-CoA dehydrogenase (SCHAD) and type 2 diabetes susceptibility: the DAMAGE study." Diabetes **55**(11): 3193-6.

Waddell, L., V. Wiley, et al. (2006). "Medium-chain acyl-CoA dehydrogenase deficiency: genotype-biochemical phenotype correlations." Mol Genet Metab **87**(1): 32-9.

Woods, C. G., J. Cox, et al. (2006). "Quantification of homozygosity in consanguineous individuals with autosomal recessive disease." Am J Hum Genet **78**(5): 889-96.

Yang, S. Y., X. Y. He, et al. (2005). "3-Hydroxyacyl-CoA dehydrogenase and short chain 3-hydroxyacyl-CoA dehydrogenase in human health and disease." Febs J **272**(19): 4874-83.

# PRODUCTS OF THE PHD ACTIVITY

**1 Homozygousity mapping of congenital hyperinsulinism of infancy (CHI) in Italian patients.**
Alessandra Gessi,[1] Maria Carla Proverbio[1,] Eleonora Mangano[1] Roberta Spinelli[2], Maddalena Bove[3], Paola Sogno Valin[3], Stefania Di Candia[3], Ilaria Zamproni[4], Stefano Mora[4], Laura Bosio[3], Manuela Caruso[5], Alessandro Salvatoni[6], Giuseppe Chiumello[3], Cristina Battaglia[1],
European human Genetics Conference 2008 Barcelona Spain May 31-June 3 2008

**2 Iperinsulinismo congenito dell'infanzia: valutazione clinica e metabolica, analisi genetica e correlazione fenotipica, creazione registro nazionale.**
Laura Bosio [3], Maria Carla Proverbio[1], Alessandra Gessi[1], Cristina Battaglia[1], Eleonora Mangano[1]
Roberta Spinelli [2], Stefania Di Candia [3], Paola Sogno Valin [3], Maddalena Bove [3], Ilaria Zamproni [3], Stefano Mora [3], Manuela Caruso [4] e Alessandro Salvatoni [5] e Giuseppe Chiumello
XVI congresso nazionale società italiana di endocrinologia e diabetologia pediatrica SIEDP/ISPED
Parma 11-13 Ottobre 2007

**3 Congenital hyperinsulinism of infancy (CHI): clinical and metabolic evaluation, molecular analysis and phenotype correlation. Institution of the Italian National Registry.**

Maddalena Bove[1], Paola Sogno Valin[1], Stefania Di Candia[1], Ilaria Zamproni[2], Stefano Mora[2], Maria Carla Proverbio[3], Alessandra Gessi[3], Eleonora Mangano[3], Roberta Spinelli [6], Cristina Battaglia[3], Laura Bosio[1], Manuela Caruso[4], Alessandro Salvatoni[5], Giuseppe Chiumello[1].
Retreat HSR 2008 Stresa

**4 Defect in β-oxidation SCHAD enzyme in familial congenital hyperinsulinism of infancy (CHI).**
Alessandra Gessi,[1] Maria Carla Proverbio[1,] Eleonora Mangano[1,] Roberta Spinelli[2], Maddalena Bove[3], Paola Sogno Valin[3], Stefania Di Candia[3], Ilaria Zamproni[4], Stefano Mora[4], Laura Bosio[3], Manuela Caruso[5], Alessandro Salvatoni[6], Giuseppe Chiumello[3], Cristina Battaglia[1],
SIB LLP 2008 LITA Segrate 31 Maggio 2008

**5 The first mutation in HADHSC gene in a Italian infant with congenital hyperinsulinism**
S. Di Candia [1], M. Bove [1], P. Sogno Valin [1], I. Zamproni [2], S. Mora [2], M. C. Proverbio [3], A. Gessi [3], E. Mangano [3], R. Spinelli [6], C. Battaglia [3], L. Bosio [1], M. Caruso [4], A. Salvatoni [5], G.Chiumello [1].
ESPE 2008