MOLECULAR ANALYSIS OF PHENYLALANINE HYDROXYLASE GENE IN SOUTH ITALY PATIENTS AFFECTED BY PHENYLKETONURIA

Alessia Palmieri

Dottorato in Scienze Biotecnologiche – XXV ciclo
Indirizzo Biotecnologie Industriali e Molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXV ciclo
Indirizzo Biotecnologie Industriali e Molecolari Università di Napoli
Federico II

MOLECULAR ANALYSIS OF PHENYLALANINE HYDROXYLASE GENE IN SOUTH ITALY PATIENTS AFFECTED BY PHENYLKETONURIA

Alessia Palmieri

Dottoranda: Dott.ssa Alessia Palmieri

Relatore: Prof. Francesco Salvatore

Coordinatore: Prof. Giovanni Sannia

INDICE

Riassunto pag. 7

Summary pag. 13

1. Introduction pag. 15

1.1 Hyperphenylalaninemias pag. 16

1.2 Phenylalanine Hydroxylase Enzyme pag. 18

1.3 Catabolic Pathway of Phenylalanine pag. 21

1.4 Pathophysiology of Phenylketonuria pag. 23

1.5 Phenylalanine Hydroxylase Gene pag. 24

1.6 Causative mutations pag. 25

1.7 Genotype-Phenotype correlation pag. 26

1.8 Responsiveness to BH4 pag. 27

1.9 Molecular genetic testing pag. 28

1.10 Functional Test pag. 29

1.11 Prenatal diagnosis pag. 30

1.12 PKU Maternal Syndrome pag. 30

1.13 Management – Therapies under investigation pag. 31

2. Aim of the thesis pag. 37

3. Materials and Methods pag. 39

3.1 Patients pag. 39

3.2 Prenatal diagnosis pag. 39

3.3 DNA extraction and quantification pag. 40

3.4 PCR pag. 40

3.5 Primer design pag. 42

3.6 Direct sequencing pag. 43

3.7 Next Generation Sequencing pag. 43

3.8 Bioinformatic Analysis pag. 46

4. Results pag. 47

4.1 Molecular diagnosis of PAH mutations pag. 47

4.2 Prenatal Diagnosis pag. 48

4.3 Molecular epidemiology of PAH mutations pag. 49

4.4 Genotype-phenotype correlation pag. 59

4.5 BH_4 Responsiveness pag. 61

4.6 Mutations of novel identification pag. 62

4.7 Mutations identified by Next Generation Sequencing pag. 63

5. Discussion and Conclusions pag. 65

Bibliography pag. 71

RIASSUNTO

Con il termine di Iperfenilalaninemie (HPAs) si intende un gruppo di condizioni patologiche genetiche, a trasmissione autosomica recessiva. Le HPAs, caratterizzate da elevati livelli di L-Fenilalanina (L-Phe) nel sangue, rappresentano il più frequente disordine ereditario del metabolismo degli amminoacidi (Online Mendelian Inheritance in Man database 261600). Le HPAs, con una frequenza di 1/10.000 nuovi nati, sono dovute a parziale o totale deficit dell'enzima Fenilalanina Idrossilasi (PAH). Questo enzima, ad espressione epatica, è responsabile della conversione della L-Phe in L-Tirosina (L-Tyr) e richiede per il suo corretto funzionamento ossigeno molecolare ed il cofattore essenziale tetraidrobiopterina (BH4). I pazienti affetti da HPA presentano una ridotta attività di PAH con conseguente diminuzione o assenza di conversione della L-Phe in L-Tyr. Gli elevati livelli plasmatici di L-Phe, insieme con i metaboliti tossici che ne derivano, Fenil-Piruvato, Fenil-Lattato e Fenil-Acetato, determinano gravi danni cerebrali. I pazienti HPAs non trattati presentano un progressivo danno neurologico, convulsioni, comportamenti simil-autistici, microcefalia, rash cutanei ed ipopigmementazione.

Le HPAs presentano un alto grado di eterogeneità sia da un punto di vista genotipico sia per quanto riguarda il fenotipo biochimico; in base ai livelli di Phe riscontrati nel sangue si possono distinguere 3 classi:

- Fenilchetonuria (PKU): a questo gruppo appartengono i pazienti con valori ematici di Fenilalanina >20mg/dl (>1200µM). Tali valori sono responsabili del fenotipo più grave.
- Iperfenilalaninemia di tipo II (HPA II): a questo gruppo appartengono pazienti che presentano valori ematici di Fenilalanina compresi tra 10-20mg/dl (600-1200µM); questo corrisponde al fenotipo moderato e viene definita anche PKU lieve.
- Iperfenilalaninemia di tipo III (HPA III): a questo gruppo appartengono quei pazienti che presentano valori ematici di Fenilalanina compresi tra 2-10mg/dl (120-600µM); questa classe corrisponde al fenotipo lieve e viene definita anche Iperfenilalaninemia lieve.

Il gene PAH codificante per la Fenilalanina Idrossilasi si trova sul braccio lungo del cromosoma 12 (12q22-q24.1-PAH OMIM 261600) è lungo circa 174 kb ed è suddiviso in 13 esoni. Mutazioni nel gene PAH causano mancata o ridotta attività dell'enzima Fenilalanina Idrossilasi. L'enzima è un omotetramero in cui ciascuna subunità è costituita da 452 amminoacidi. Ciascun monomero è costituito inoltre da tre domini:

- Un dominio regolatorio N-terminale. In seguito al legame con la L-Phe, la PAH subisce un cambiamento conformazionale che ne determina l'attivazione.
- Un dominio centrale in cui è localizzato il sito catalitico dell'enzima con regioni di legame per il substrato L-Phe e regioni di legame per il cofattore (Cofactor Binding Regions, CBR) BH4, per l'ossigeno e un catione Fe 2+.
- Un dominio C-terminale responsabile della tetramerizzazione della proteina.

Il gene PAH produce un trascritto maturo di circa 2680 basi Il database del gene PAH include più di 600 diverse mutazioni causative di HPA. Il tipo più frequente è rappresentato da mutazioni tipo missense (62%). Le altre mutazioni sono meno frequenti (delezioni 14%; splicing 11%; silenti 6%; non-sense 5%; inserzioni 2%) (http://www.pahdb.mcgill.ca). La posizione e la natura delle mutazioni determinano un diverso effetto sull'attività enzimatica da cui spesso deriva l'eterogeneità fenotipica della PKU.

La Fenilalanina Idrossilasi (PAH) è una monoossigenasi a funzione mista che catalizza la reazione di idrossilazione della L-Phe in posizione 4 usando ossigeno molecolare e il cofattore essenziale BH4; nel corso della reazione, la BH4 è convertita nella forma ossidata a Quinonoide Diidrobiopterina (qBH2). Dalla L-Tyr, tramite vie biosintetiche, si ottengono le Catecolamine (Dopamina, Adrenalina e Noradrenalina), la Melanina e gli Ormoni Tiroidei (Triiodiotiroxina e Tetraiodiotiroxina). L'inadeguata attività di PAH determina la diminuzione o l'assenza della conversione della L-Phe in L-Tyr.

Il deficit o la non funzionalità di PAH comporta:

- l'aumento della concentrazione di L-Phe nel sangue;
- l'attivazione di pathway metabolici secondari con formazione di metaboliti tossici quali il Fenil-Piruvato, il Fenil-Lattato e il Fenil-Acetato.

Gli aumenti ematici dei livelli di Phe e la mancata conversione in Tyr possono determinare danni a livello del sistema nervoso centrale attraverso diversi meccanismi:

- deficit di Tyr: precursore delle Catecolammine, degli Ormoni Tiroidei e della Melanina, principale pigmento cutaneo. Questo giustifica gli attacchi di epilessia e l'ipopigmentazione cutanea spesso riscontrati nei soggetti HPAs.
- inibizione della Piruvato Decarbossilasi presente nel sistema nervoso centrale da parte dell'acido

Fenilipiruvico con conseguente interferenza nella formazione della mielina.

• competizione con il trasporto degli amminoacidi neutri (LNA) attraverso la barriera emato-encefalica: l'effetto finale potrebbe essere un blocco parziale del traffico amminoacidico. Quest'ultimo meccanismo sembrerebbe implicato nella patogenesi molecolare del danno cerebrale, infatti l'ingresso nel SNC della L-Phe è infatti mediato dal carrier degli aminoacidi neutri L-aminoacid trasporter 1 (LAT1).

Le alte concentrazioni plasmatiche di L-Phe potrebbero inibire LAT1 ed impedire il passaggio degli altri aminoacidi neutri nel cervello. Inoltre, polimorfismi nel gene che codifica per il sistema di trasporto degli LNAA (LAT1), potrebbero giocare un ruolo nel determinare la suscettibilità al danno cerebrale in soggetti HPAs. In definitiva, la discordanza esistente in alcuni pazienti tra fenotipo biochimico e clinico potrebbe anche essere spiegata dal ruolo svolto dalla BEE per il trasporto di L-Phe e dai meccanismi molecolari che determinano questo passaggio.

Inoltre, i bambini affetti da HPAs a causa dei deficit di L-Tyr e dei suoi derivati, spesso entro pochi anni dalla nascita manifestano sintomi quali ipereccitabilità e crisi convulsive che accompagnano il grave ritardo mentale. Quindi per prevenire il danno neurologico e le alterazioni comportamentali tipiche dei soggetti HPAs è fondamentale che la diagnosi e l'inizio della terapia vengano effettuati entro i primi giorni di vita.

Le HPAs possono essere tenute sotto controllo, infatti, mediante una dieta a basso contenuto proteico e quindi di L-Phe, con la supplementazione di L-Tyr e di integratori a base di miscele amminoacidiche. Attraverso un buon controllo delle concentrazioni di L-Phe, soprattutto durante la prima infanzia, la maggior parte dei pazienti presenta uno sviluppo neurologico e fisico nella norma. Da studi recenti è emerso che adulti che abbandonano il regime dietetico restrittivo riscontrano anomalie elettrofisiologiche del S.N.C. In accordo con questi dati le linee guida raccomandano ai pazienti HPAs un trattamento dietetico a vita; esso risulta essere estremamente difficile da seguire soprattutto in particolari epoche della vita come l'adolescenza e la gravidanza. Durante la gravidanza gli alti livelli plasmatici di L-Phe attraversano la placenta e risultano estremamente tossici e teratogeni per il feto determinando una grave embriopatia plurimalformativa (Sindrome da PKU materna o MSPKU). La MSPKU è lesiva della morfogenesi, nonché dello sviluppo fisico e cerebrale del prodotto del concepimento. Il fenotipo clinico che ne deriva è caratterizzato da ritardo mentale,

dismorfie facciali, microcefalia, ritardo di crescita intrauterino, difetti cardiaci congeniti.

Poiché la terapia dietetica risulta molto difficile da seguire, sono in via di sperimentazione terapie alternative e/o integrazioni farmacologiche alla dietoterapia al fine di rendere più facile la compliance al trattamento dietetico. Alternative terapeutiche comprendono la somministrazione di BH4 in pazienti responsivi, l'utilizzo di aminoacidi neutri, la terapia enzimatica sostitutiva con Phe Ammonia Liasi (PAL), la terapia genica, l'utilizzo di pluripotent stem cells.

Una significativa percentuale di pazienti con HPA rispondono alla somministrazione di BH4 con una riduzione significativa della L-Phe sierica. Poiché nella stragrande maggioranza dei pazienti HPA il genotipo correla con il fenotipo biochimico e clinico, l'analisi molecolare del gene PAH è molto importante perché contribuisce a confermare o predire il fenotipo biochimico, identifica i carriers (in genere asintomatici) e permette l'analisi prenatale nelle famiglie ad alto rischio per HPAs. Inoltre, l'ampia eterogeneità delle mutazioni a carico del gene PAH rende fondamentale definire l'epidemiologia molecolare delle mutazioni responsabili sia della PKU che della HPA (nei singoli gruppi etnici).

La diagnosi biochimica delle HPAs si basa sul dosaggio delle concentrazioni plasmatiche di L-Phe mediante l'utilizzo della spettrometria di massa. La diagnosi molecolare di PKU viene effettuata mediante l'analisi molecolare del gene codificante la PAH. Tale indagine, in Campania, viene effettuata presso i laboratori del CEINGE, e prevede tre fasi: estrazione del DNA genomico dal campione di sangue conservato in EDTA, amplificazione del gene PAH mediante Polymerase Chain Reaction (PCR) e suo sequenziamento diretto.

La diagnosi molecolare delle HPAs contribuisce a confermare o predire il fenotipo biochimico dei pazienti consentendo loro di intraprendere una terapia appropriata e personalizzata. Inoltre attraverso l'analisi molecolare del gene PAH è possibile identificare i carriers, in genere asintomatici. Nei casi in cui è richiesta, l'analisi molecolare del gene PAH permette di effettuare la diagnosi prenatale in casi di familiarità per HPAs.

Durante il periodo del mio lavoro di tesi, ho effettuato l'analisi molecolare del gene PAH di pazienti provenienti dalla regione Campania. Per effettuare la diagnosi molecolare ho utilizzato diverse metodiche di biologia molecolare come l'estrazione del DNA genomico, disegno e sintesi di primers specifici, reazione di amplificazione

PCR ed il sequenziamento diretto secondo il metodo di Sanger.

Sono stati analizzati 390 pazienti HPA del Sud Italia (età media 15 anni, range 2-25 anni, maschi: femmine rapporto 1.2:1). I pazienti sono stati classificati come segue: 121 (15,8%) PKU, 64 (8,4%) HPAII e 205 (26,9%) HPAIII. Abbiamo effettuato l'analisi molecolare per la rilevazione dello stato di portatore in 373 soggetti. Inoltre, abbiamo eseguito la diagnosi prenatale in tre famiglie ad alto rischio dal momento che era già presente un figlio affetto; l'indagine molecolare ha previsto l'analisi del gene PAH nei due genitori e nei villi coriali o negli amniociti prelevati alla sedicesima settimana di gestazione. In una delle famiglie analizzate, la madre del feto era affetta da HPA mentre nel padre non è stata riscontrata alcuna mutazione quindi il feto è risultato portatore di una singola mutazione. Nelle altre due famiglie analizzate, un feto è risultato affetto da HPA in quanto portatore in eterozigosi di entrambe le mutazioni dei genitori. L'altro feto è risultato non affetto per l'assenza di entrambe le mutazioni presenti nei genitori.

L'analisi molecolare ha permesso di identificare mutazioni in 763 su 780 alleli con una detection rate del 97.8%. In diciassette pazienti è stata individuata solo una mutazione causativa. La mancata identificazione di una mutazione in uno dei due alleli può essere dovuta a: 1) alterazioni presenti nel promotore o nella regione 3' non tradotta non sottoposti a sequenziamento; 2) difetti molecolari negli introni; 3) mutazioni all'interno di un enhancer; 4) grandi delezioni o duplicazioni. L'analisi ha permesso l'identificazione di 103 mutazioni diverse che si distribuiscono uniformemente lungo la sequenza del gene PAH senza hot spot mutazionali. Delle 103 varianti, 76 sono mutazioni missense (73%), 19 mutazioni a carico di siti di splicing (18,4%), 4 sono delezioni (3,8%) e 4 mutazioni frame-shift (3,8%). Alcune mutazioni sono state trovate con una maggiore frequenza: le mutazioni p.R261Q, c.1066-11G> A e p.A403V hanno frequenza superiore al 10% (frequenza cumulativa = 24.8); le mutazioni p.L48S e p.A300S hanno una frequenza compresa tra il 6 ed il 10%, (frequenza cumulativa = 13.6); 51 mutazioni hanno frequenza compresa tra lo 0,2% ed il 3,2% (frequenza cumulativa = 56.9); 47 mutazioni sono state trovate in un singolo allele con una frequenza dello 0,1 % (frequenza cumulativa = 4.7).

L'esone 7 è la regione in cui sono state trovate la maggior parte delle mutazioni; infatti esso contiene circa il 23,8% delle mutazioni totali (la percentuale aumenta al 33,7% considerando solo i pazienti PKU). Inoltre, l'esone 11 insieme all'esone 12 contiene circa il 34,4% delle

mutazioni totali (la percentuale diventa del 39% considerando solo i pazienti HPAIII e del 28% considerando i PKU). L'esone 2 contiene l'11,5% delle mutazioni totali. Gli esoni rimanenti (1,3,4,5,6,8,9 e 10) contengono tra lo 0,4 e l'8,3% del totale delle mutazioni. Inoltre, dall'analisi dei nostri pazienti è emerso che circa il 75% dei pazienti analizzati hanno mutazioni BH4-responsive. Infine, in questo lavoro di tesi, tra le mutazioni identificate nella popolazione del Sud Italia, 18 sono di nuova identificazione (p.E76X, p.I95T, p.I65M, p.N223Y, p.Q235X, c.707-2delA, p.R297L, p.Q301P, p.L321I, p.W326S, p.F382L, c.1199nt4A>C, c.1315nt20C>T, p.K328N, p.N426H, p.I406M, p.Q419R, p.T418I): 14 missense, 3 di splicing e 1 delezione.

Attraverso l'analisi del gene PAH diretto, abbiamo raggiunto una detection rate del 97.8%. Il nostro studio conferma l'eterogeneità delle mutazioni nei pazienti PKU anche nella regione Campania e che l'epidemiologia delle mutazioni PAH nel Sud Italia si differenzia da quella di altri gruppi italiani ed europei. Inoltre, i nostri dati indicano che esiste frequentemente una correlazione genotipo-fenotipo. In particolare, si conferma che p.R261Q, p.L48S e c.1066-11G> A sono più frequentemente associate con la PKU e fenotipi PKU lievi, mentre le mutazioni p.A403V e p.E390G sono peculiari dei fenotipi più lievi.

Nell'ultima fase dello studio, abbiamo ottimizzato e validato un protocollo sperimentale per l'analisi dell'intera sequenza del gene PAH, incluse le zone regolatorie e gli introni, basato su tecnologie di sequenziamento di nuova generazione, al fine di identificare mutazioni causative di PKU anche in quei pazienti risultati negativi dopo lo screening molecolare descritto in precedenza.
In conclusione, il nostro studio, ci ha permesso di definire:
- la causa molecolare alla base delle HPAs nel 98.8% dei pazienti, -l'epidemiologia delle mutazioni nel gene PAH nei pazienti del Sud Italia, in particolare in Campania, -il rischio residuo di essere portatori, -la sensibilità al trattamento con BH4 dei pazienti affetti da PKU, -la correlazione tra alcune mutazioni e fenotipi della malattia.

Questi dati possono aiutare a ricostruire, da un punto di vista storico e antropologico, i vari insediamenti che si sono verificati nel nostro paese.

SUMMARY

Hyperphenylalaninemias (HPAs), the most common inborn error of metabolism (1:10.000), is recessively inherited and caused by mutations in the gene encoding phenylalanine hydroxylase (PAH). PAH converts L-Phenylalanine into L-Tyrosine. PAH block results in accumulation of L-Phe and toxic metabolites. HPAs is classified according Phe serum levels as: 1.Classical PKU 2.Mild PKU 3.Mild HPA.

PAH gene, located on chromosome 12, is composed of 171 Kb and 13 exons. To date more than 600 causative mutations for PKU are described in the entire PAH gene: 62% are missense mutations; 14% deletions; 11% splice site mutations; 6% nonsense mutations; 2% insertions.

PAH, hepatic monooxygenase of 452 amino acids that catalyzes conversion of L-Phe to L-Tyr using 6R-l-erythro-5,6,7,8-tetrahydrobiopterin (BH4) and molecular oxygen. PAH consists of three structural domains: 1.N-term regulatory domain 2.Large catalytic domain 3.C-term tetramerization domain.

HPA is also the first metabolic disorder in which treatment (a Phe-restricted diet) was found to prevent clinical symptoms. The aim of my thesis was to perform molecular analysis of the PAH gene in 390 PKU patients and 373 carriers from the Campania Region Screening Center. The analysis of PAH mutations allowed to define the molecular epidemiology of PAH gene in Campania and compare the obtained data with those from other regions of Italy. The analysis of the PAH gene have been performed through the extraction of genomic DNA from peripheral blood, followed by amplification and direct sequencing of the entire gene.

In the last phase of the study, we optimized and validated a next generation sequencing protocol for the analysis of the entire PAH gene sequence, including introns and regulatory regions to identify PKU causative mutations in patients negative.

In conclusion, our study, has allowed us to define: - molecular diagnosis of HPA patients achieving a detection rate of 98.8%, - epidemiology of mutations in PAH gene of patients from Campania, - the residual risk to be carriers, and prenatal diagnosis in PKU family, - the responsiveness of BH4 in PKU patients, - good correlation between mutations and phenotypes of the disease.

These data may help to reconstruct the historical and anthropological point of view, the various settlements that have occurred in our country.

1. Introduction

Inborn Errors of Metabolism (IEM) comprise disorders in which a single gene defect causes a clinically significant block in a metabolic pathway resulting either in accumulation of substrate behind the block or deficiency of the product. All IEMs are all genetically transmitted typically in an autosomal recessive or X-linked recessive fashion. The major categories are:

- Organic acidemias (e.g., methylmalonic or propionic acidemia, multiple carboxylase deficiency) are caused by abnormal metabolism of proteins, fats or carbohydrate.
- Fatty acid oxidation defects (e.g., short, medium, and long- chain acyl-CoA dehydrogenase deficiencies) also known as Beta-oxidation defects, are a distinct type of organic acid disorder.
- Primary Lactic Acidoses (e.g., pyruvate dehydrogenase, pyruvate carboxylase and cytochrome oxidase deficiencies) present with severe lactic acidosis.
- Urea cycle defects (e.g., citrullinemia, ornithine transcarbamylase deficiency, and arginosuccinic aciduria).
- Disorders of carbohydrate metabolism (e.g., galactosemia, hereditary fructose intolerance, fructose 1,6-diphosphatase deficiency and the glycogen storage diseases) are a heterogeneous group caused by inability to metabolize specific sugars, aberrant glycogen synthesis, or disorders of gluconeogenesis.
- Lysosomal storage disorders (e.g., mucopolysaccharidosis, Tay-Sachs, Niemann-Pick disease, Gaucher's disease) are caused by accumulation of glycoproteins, glycolipids, or glycosaminoglycans within lysosomes in various tissues.
- Peroxisomal disorders (e.g., Zellweger syndrome and neonatal adrenoleukodystrophy) result from failure of the peroxisomal enzymes.
- Aminoacidopathies (e.g., phenylketonuria, hereditary tyrosinemia, nonketotic
- hyperglycinemia, maple syrup urine disease [MSUD] and homocystinuria) are a very heterogeneous group ofdisorders.

Phenylketonuria (PKU) is an Inborn Error of Metabolism in which the normal conversion of the dietary amino acid L-phenylalanine (L-Phe) to tyrosine is blocked. The resulting

build-up of phenylalanine and its metabolites in young patients produces a number of severe effects including intellectual impairment and cutaneous changes.

1.1 Hyperphenylalaninemias

Hyperphenylalaninemias (HPAs) are a group of inherited diseases due to the defective phenylalanine hydroxylase (PAH) activity resulting in accumulation of phenylalanine in blood and other tissues. The mean frequency of HPAs is 1:10.000 in Caucasian population, in most cases (98% of subjects), HPAs result from mutations in the phenylalanine hydroxylase gene (20). Human phenylalanine hydroxylase (PAH) converts the essential amino acid L-Phe into L-tyrosine (L-Tyr), essential for the synthesis of important neurotransmitters like dopamine, epinephrine, nor-epinephrine and melanin (28). L-Phe is an essential amino acid present in the various proteins contained in the diet. Normally, a small L-Phe amount is used for protein synthesis; the remainder is hydroxylated to L-tyrosine, which is used for synthesis of protein and several compounds or is degraded to produce energy. The hydroxylation of phenylalanine requires phenylalanine hydroxylase (PAH) enzyme, (6R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH4) cofactor, and molecular oxygen (22, 29). The catalysis by this iron-dependent enzyme is the major pathway for catabolic degradation of dietary L-Phe and accounts for approximately 75% of the L-Phe disposal (34). The autosomal recessive disorder PKU is the result of a deficiency of PAH enzymatic activity. When the conversion of L-Phe to L-tyrosine is blocked, L-Phe accumulates in body fluids or is converted to other metabolites toxic for the central nervous system (5).

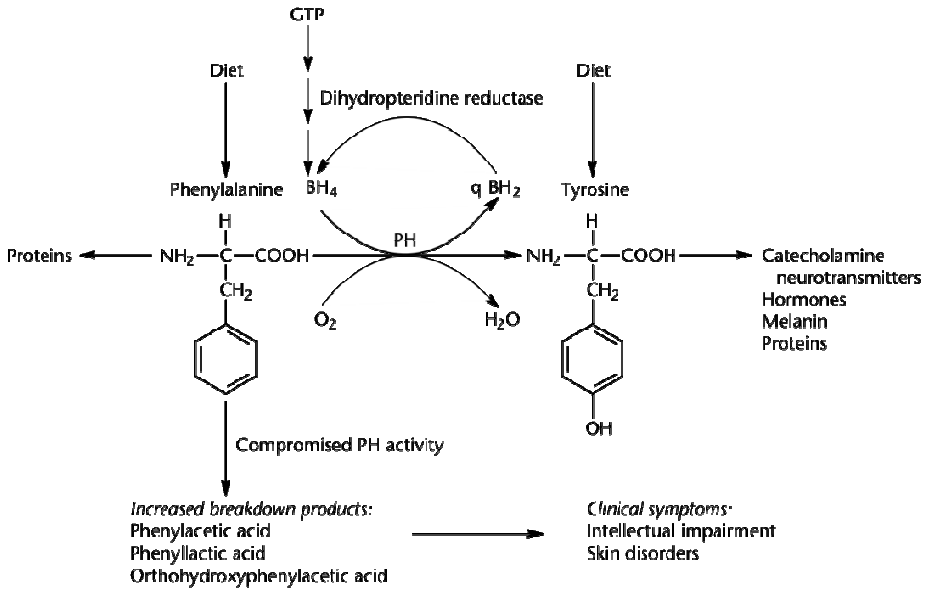

Fig. 1: Phenylalanine metabolism. Phenylalanine hydroxylase catalyses the conversion of phenylalanine into tyrosine. Deficiencies in the activity of this enzyme result in incomplete phenylalanine metabolism and build-up of toxic waste products.

The associated phenotypes range in severity from classic PKU to mild HPA. HPAs appear as a highly heterogeneous trait with a broad continuum of phenotypes (6). The term PKU is reserved to the most severe form of the disease. Based on plasma L-Phe levels, 3 different phenotypes of HPAs due to PAH deficiency have been described:

1. Classical PKU
2. Mild PKU
3. Mild HPA

The patients with the classical or typical form of PKU have L-Phe levels above >20mg/dl (>1200µM). Mild PKU patients have L-Phe levels between 10-20mg/dl (600-1200µM). Mild HPA patients have L-Phe levels lower than 10mg/dl (120-600µM). With a routine paediatric screening, the diagnosis of PKU is usually established soon after birth. The newborn screening test involves the measurement of blood L-Phe levels by the Guthrie test, a bacterial inhibition assay or my mass spectrometry. Normally, blood L-Phe levels are below 2mg/dl (120 µM).

If left untreated, classical PKU patients can lead to severe intellectual impairment. Increased levels of L-Phe and its metabolites cause irreparable damage to the developing central nervous system. Furthermore, this biochemical

defect can result in a variety of cutaneous abnormalities, including diffuse hypo-pigmentation, eczema and photosensitivity (5). These cutaneous changes may result from the toxic effects of L-Phe and its decomposition products in the skin. However, the clinical picture includes delayed psychomotor development, mental retardation, mousy odour in the urine, irritability, light cutaneous pigmentation, eczema, and epilepsy. Among PKU patients, a remarkably wide variation is observed in both clinical manifestation and therapeutic response (5).

1.2 PAH Protein

(PAH: EC 1.14.16.1). PAH is a hepatic monooxygenase that catalyses the conversion of L-Phe to L-Tyr using 6R-l-erythro-5,6,7,8-tetrahydrobiopterin (BH4) as a coenzyme.

PAH consists of three structural and functional domains:

- N-terminal regulatory domain (residues 1-142),
- Large catalytic domain (residues 143-410),
- Small C-terminal tetramerization domain (residues 411-452) (Figs. 2 and 4).

The tetramer model is a dimer of two conformationally different dimers, and its dimensions are of about 85 Å × 100 Å × 75 Å (5, 6) (Fig.2):

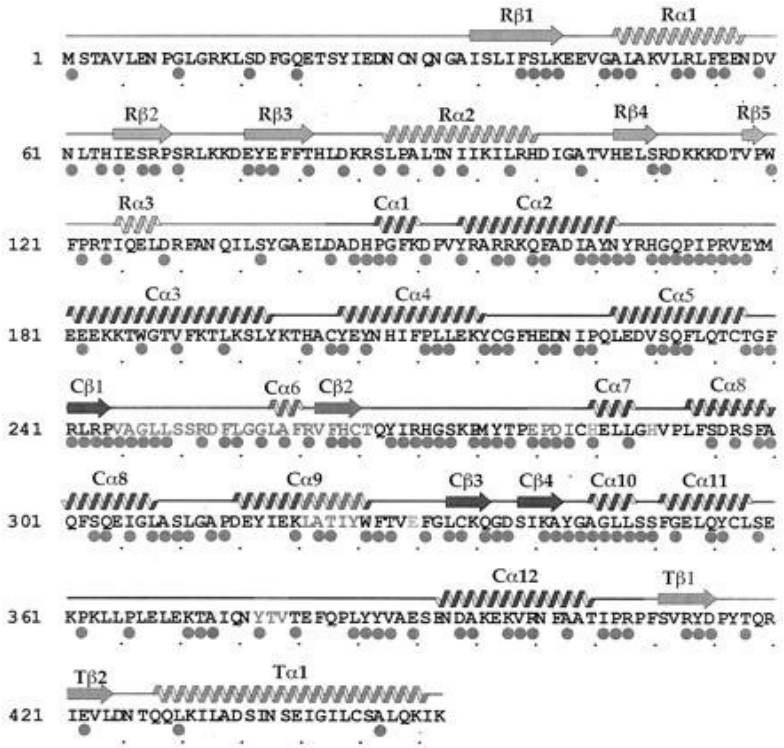

Fig. 2: Scheme showing the secondary structure assignment of human Phenylalanine Hydroxylase sequence (SWISS-PROT P00439).

The regulatory domain of PAH contains an α-β sandwich with an interlocking double βαβ motif (Fig 3). The N-terminal autoregulatory sequence (ARS; residues 19–33) extends over the active site in the catalytic domain. A key residue in the regulation of this enzyme is the serine 16, phosphorylated by the cAMP-dependent protein kinase (PKA). Although L-Phe is the primary factor in the activation of PAH, phosphorylation mediates the enzyme reaction by decreasing the substrate concentration required to activate PAH (20, 28).

The catalytic domain has a basket-like arrangement composed of 12 α-helices (45%) and 6 β-strands (16%). The active site is located in the center of the catalytic domain (18). Adjacent to the active site is a channel that may provide substrate access to the active site (Fig. 3). As isolated, the PAH contains a Fe (III) ion in the active site (34), coordinated to His285, His290 and Glu330. Both His285 and His290 have been shown by site-directed mutagenesis to be required for iron binding (23). A set of 27 amino acids (His 263 to His 289) forms a region that

was thought to be responsible for BH4 binding (29).

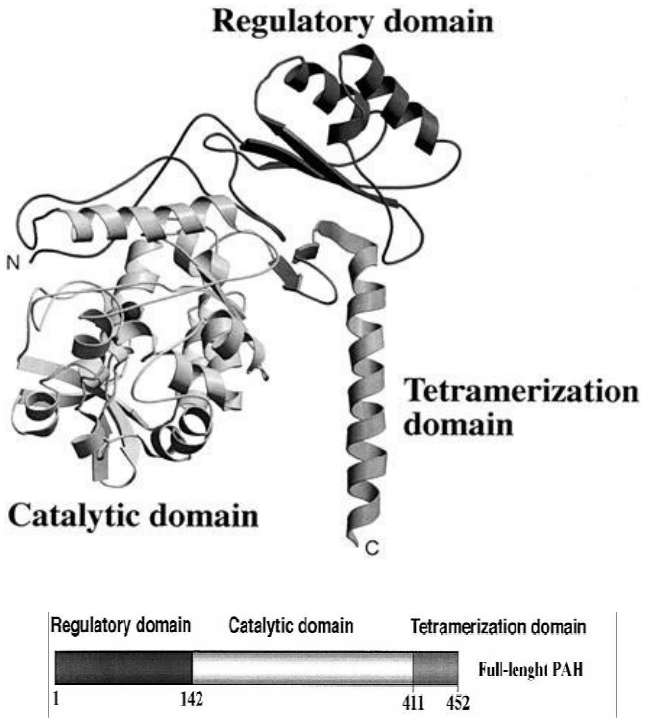

Fig. 3: Structure of a monomer of human Phenylalanine Hydroxylase full-length composite model. In figure are shown the regulatory domain (residues 19–142), the catalytic domain (residues 143–410) and the tetramerization domain.

The tetramerization domain, consisting of a C-terminal arm, contains two antiparallel β-strands, forming a β-sheet (the dimerization motif), and α-helix (the tetramerization motif) (Fig. 3). The crystal structure of PAH has revealed that the tetrameric oligomers are dimers of dimers in which the interaction between the two dimers is mediated by the C-terminal 'arm'. The tetrameric and dimeric forms of PAH are in equilibrium and have different catalytic properties because the tetramer, but not the dimer, demonstrates a positive kinetic cooperativity with respect to L-Phe. In particular, substrate activation results in conformational changes involving the tertiary as well as the quaternary structure and drives the tetramer–dimer equilibrium toward the tetrameric form (9).

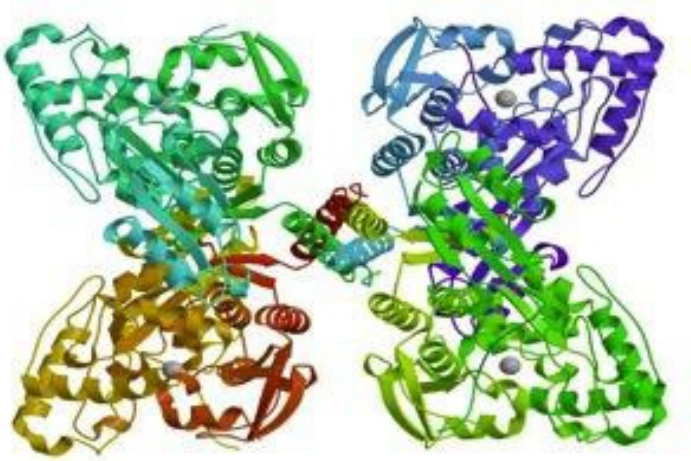

Fig. 4: Structure of a tetramer of human Phenylalanine Hydroxylase full-lengtht. The monomers interact through the tetramerization domain.

1.3 Catabolic Pathway of Phenylalanine

The L-Phe amino acid is hydroxylated by the enzyme PAH to form Tyr; the PAH is a mixed function monooxygenase which catalyzes the hydroxylation reaction of the L-Phe at position 4 using molecular O2 and the essential cofactor BH4 (Fig.5); in the course of the reaction, the BH4 is converted into the oxidized form to Quinonoide dihydrobiopterin (QBH2). The cofactor is regenerated by the action of the enzyme dihydropteridine reductase (DHPR) that uses NADPH as a source of reducing power. Tyr, via secondary biosynthetic pathways, is then converted in several products: catecholamines (dopamine, adrenaline and noradrenaline), melanin and thyroid hormones. The inadequate activity of PAH determines the decrease or absence of the conversion of L-Phe in L-Tyr (5, 16).

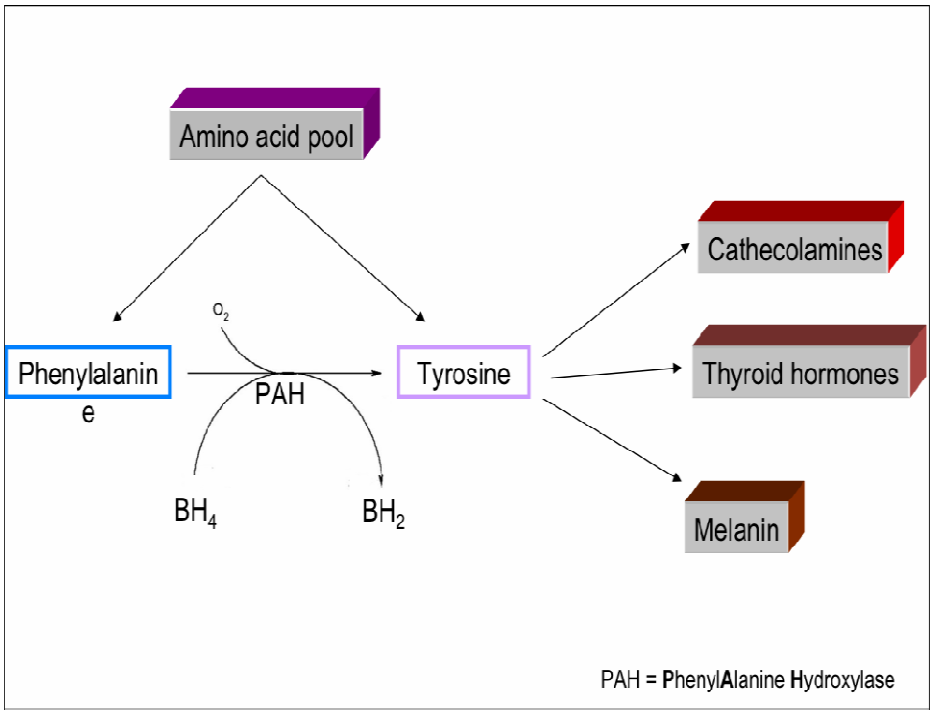

Fig. 5: Normal Phenylalanine Metabolism.

The deficit or non-functionality of PAH involves:

• The increase of blood L-Phe;

• The activation of secondary metabolic pathways with the formation of toxic metabolites such as phenyl-pyruvate, lactate and the phenyl-acetate (Fig. 6).

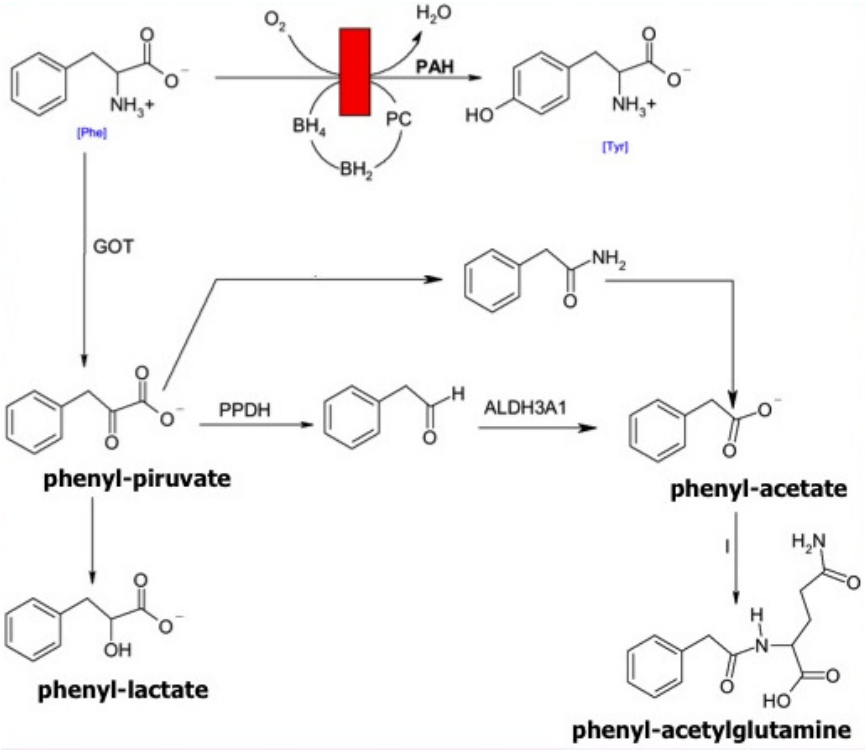

Fig. 6: Alterated Phenylalanine metabolism. The inhibition of Phenylalanine Hydrolase leads to the accumulation of Phenylalanine and of toxic metabolites.

1.4 Pathophysiology of Phenylketonuria

The increased levels of L-Phe in blood can cause damage in the central nervous system through several mechanisms:

- Tyr deficiency, which is the precursor of catecholamines, thyroid hormones and melanin. This justifies the attacks of epilepsy and skin hypo-pigmentation often seen in PKU patients.
- Inhibition of pyruvate decarboxylase in the central nervous system by acid Phenyl-piruvic; it interferes with the myelin formation (5).
- Competition with the transport of neutral amino acids through the blood-brain barrier: the final effect may be a partial blockage of the amino acid traffic.
- Albeit not yet completely clear, the latter mechanism would seem implicated in the pathogenesis of brain damage from PKU. From a molecular point of view, the L-Phe entry into the CNS is mediated by the carrier of the neutral amino acids (L-aminoacid transporter 1-LAT1).

Two other neutral amino acids, L-Tyr and L-Typtophan, are transported into the brain through the same carrier LAT1. The high plasma concentrations of L-Phe may inhibit and prevent the passage of the other neutral amino acids in the brain. This will lead to a dysfunction of neurotransmitters (30).

In addition, polymorphisms in the gene encoding for the LNAA transport system (LAT1) may play a role in determining the susceptibility to cerebral damage in PKU. Ultimately, the discrepancy that exists in some patients between clinical and biochemical phenotype could also be explained by the role played by BEE in the transport of L-Phe.

In addition, children with PKU, because of the lack of Tyr and its derivatives, often within a few years after the birth, have symptoms such as hyperexcitability. To prevent neurological damage and behavioural changes is essential that diagnosis and therapy are carried out within the first few days of life.

1.5 Phenylalanine Hydroxylase Gene

The PAH gene is located on the long (q) arm of chromosome 12 between positions 22 and 24.2. More precisely, the PAH gene is located from base pair 103,232,103 to base pair 103,311,380 on chromosome 12. The Human PAH gene is composed of 171 Kb organized into 13 exons. The full-length PAH cDNA (GenBank cDNA Reference Sequence U 49897) encodes a protein of 452 amino acids (about 52 kDa for each subunit).

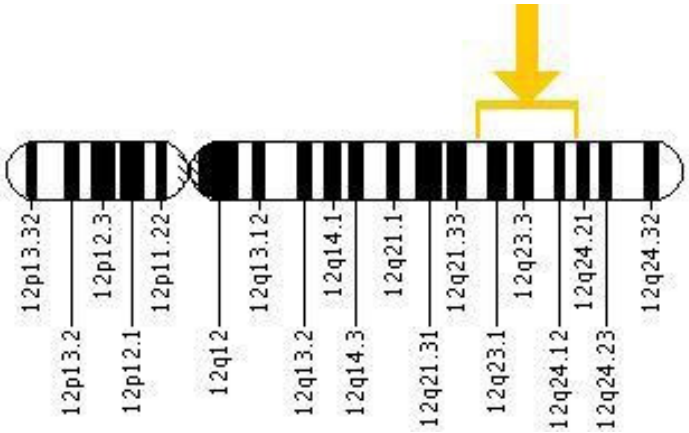

Fig 7: Cytogenetic Location of Phenylalanine Hydroxylase gene: 12q22-q24.2. Molecular Location on chromosome 12

1.6 Causative mutations

To date more than 600 causative mutations for PKU have been described in the entire PAH gene.

The pathogenic mutations in the PAH gene of PKU patients were identified after the cloning of the gene in 1983. Subsequently, the 'Phenylalanine Hydroxylase Locus Knowledgebase' PAHdb was generated and curated at McGill University (http://www.pahdb.mcgill.ca/); in this way, the mutations and the variants present in PAH gene are catalogued together with different information about PAH alleles and mutations; in fact, clinicians and laboratories from around the world report in this database the features of all mutations. In the 'Phenylalanine Hydroxylase Locus Knowledgebase' PAHdb database at this time are present more than 600 PAH mutations identified world wide. Furthermore, in the database are listed the notices about the nucleotide alterations as the *in vitro* residual enzyme activities of about 200 mutations.

Of the 600 mutations:

- 60.5% are missense mutations;
- 13.5% deletions;
- 11% splice site mutations;
- 5% nonsense mutations;
- 1.8% insertions.

Most of the PAH mutations are missense mutations that not avoiding transcription or translation. The majority of the patients are compound heterozygotes (10, 12, 13, 14, 24, 25, 26, 50, 51). The nature and the position in the gene of the mutation determines the effect on the structural, biochemical and biophysical properties of the PAH as the enzymatic activity which correlates with the phenotype of the patient.

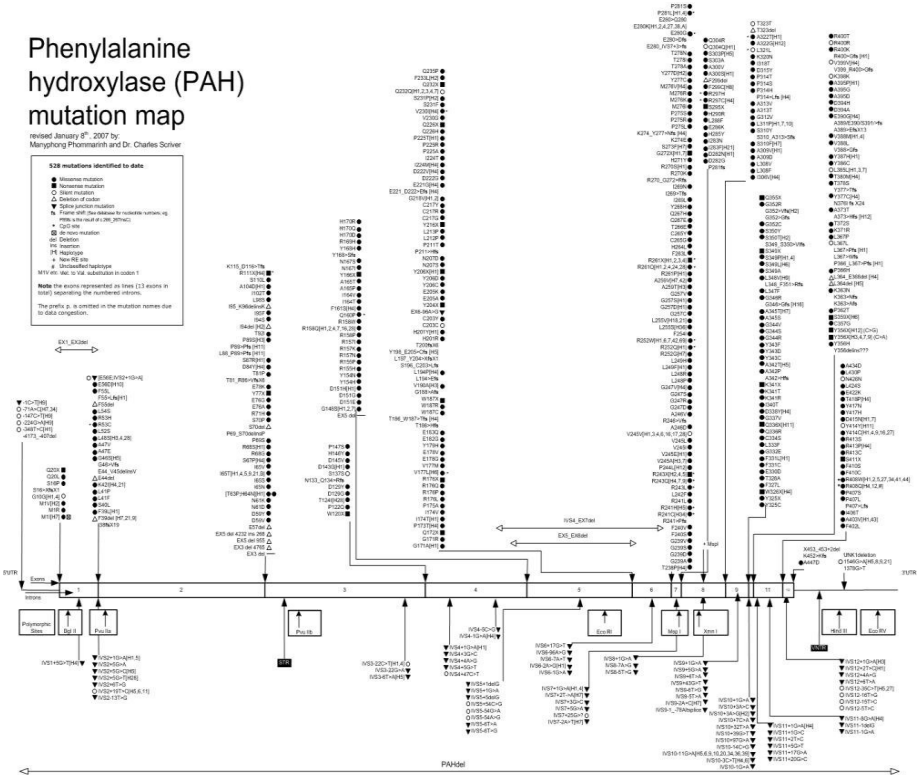

Fig. 8: Mutation map of Phenylalaine Hydroxylase gene. In figure are reported causative mutations described until August 2007.

1.7 Genotype-Phenotype correlation

In 1998 Guldberg at al. identified the complete genotype in 686 patients from 7 European centers. Based on the L-Phe tolerance of 297 functionally hemizygous patients (patients carrying a null allele and the uncharacterized mutant allele), an arbitrary phenotype class was assigned to each mutations for which the residual enzyme activity was not known. In 79% of the analyzed patients, the predicted phenotype tallied with the observed phenotype. Moreover, this study led to the conclusion that the genotype is the main determinant of the biochemical phenotype in most PKU patients. Many studies reported genotype/phenotype correlations in different populations (12, 13, 14, 24, 25, 26, 50, 51); some publications reported also inconsistencies, i.e. genotypes associated with several different phenotypes. Numerous studies reported that the combination of the two mutant alleles is important for the residual PAH enzyme activity. In fact, PAH enzyme, being a homotetramer, consists of arrangements of different mutant monomers and its residual activity is not simply the mean of each subunit activity (9). However, there is a highly significant

correlation between the genotype and the biochemical phenotype of the patients although the genotype is expressed as the predicted residual enzymatic activity of the patient and calculated as the mean of the combined in vitro residual enzyme activities of both mutant alleles.

In the population of Campania, genotype–phenotype analysis in compound heterozygous patients was performed according with the 'quasi-dominant' theory of Guldberg et al. (26), in which the milder mutation between the two mutations influences the outcome of phenotype. This simple method predicted a correct correlation genotype and phenotype in 76% of Campania cases. In the remaining patients, there was discordance between genotype and phenotype. In addition to our data, several other studies have reported inconsistencies between genotype–phenotype (12, 13, 14). Different factors may contribute to this discrepancy: possible phenotypic misclassifications, incorrect tolerance assessment, the unpredictable result of allelic complementation in heterozygous patients, and the role of modifier genes, including cellular quality control systems (15, 19, 42).

1.8 Responsiveness to BH4

A pilot study in the United States demonstrated that many patients with PKU respond to an oral loading dose of 10 mg/kg BH4 (43). The clinical trial was performed on 37 PKU patients, 20 females and 17 males, including 5 children ages 5–14 years who participated. Twenty-four patients were classical or severe PKU (blood Phe >1200 µmol/L), 10 had atypical or mild PKU (blood Phe 360–1199 µmol/L), and 3 had mild HPA (blood Phe <360 µmol/L). Today, the responsiveness to BH4 administration in PKU is defined as a decrease of L-Phe in serum of more than 30% of the value before the BH4 challenge (20mg/kg body weight), within approximately 24h post-load. Furthermore, BH4 deficiency should be excluded and pathologic PAH mutations should be identified. However, in literature BH4-responsiveness has been assessed in different ways (1, 3). Newborns are usually tested using the standard BH4 loading test. In few cases a BH4 treatment trial was performed. BH4-responsiveness is different among patients with different genotypes (3). In some BH4-responsive patients, L-Phe levels decrease significantly, but do not reach physiologic concentrations.

In addition, different PAH mutations have been studied in vitro by expression analysis, resulting in expression of mutant PAH protein with different residual enzymatic activity (9, 12). BH4-responsive PAH mutations are listed

elsewhere (http://www. bh4.org/biopku.html). It was reported that residual PAH activity (in vivo) is a prerequisite for BH4-responsiveness; on the other hand, mutations, causing severe structural changes (truncation) in the expressed protein and with undetectable PAH activity, are not expected to be stimulated by BH4. However, in vitro expression of mutant PAH protein with residual activity does not guarantee BH4–responsiveness.

Different hypotheses have been proposed for explaining BH4 responsiveness (4, 17, 45):

1) PAH mutant enzymes have a reduced binding affinity for BH4, which can be surmounted at increased BH4 concentrations;
2) BH4 may stabilize PAH by protecting the protein against proteolytic cleavage or degradation, as a chaperon;
3) BH4 helps stabilize PAH mRNA, and
4) BH4 contributes to regulation of PAH gene expression (18, 44, 4). However, it has been indicated that BH4 influences the conformational stability of PAH and also its activity (46).

1.9 Molecular genetic testing

Clinical testing: targeted mutation analysis. A panel of 1–15 more common point mutations and very small deletions have a detection rate of approximately 30–50%. Alleles may be population related (40).

Mutation scanning: with this technology it is virtually possible to indentify all point mutations in the entire PAH gene. Mutation scanning by denaturing high performance liquid chromatography, a fast and very efficient method to detect locus-specific point mutations, has shown its relevance for detection of the pathologic alleles of PKU (40).

Duplication/deletion analysis: comparative multiplex dosage analysis is helpful to detect large duplications or deletions in the rare cases in which no mutations have been detected. This technique has been successfully utilized to identify abnormal dosage in 20% of uncharacterized PKU mutations and therefore, duplications and/or deletions may account for up to 3% of mutations (40).

Linkage analysis: in families at high risk for PKU in which only one or neither PAH allele has been identified, linkage analysis may be an option for carrier testing and prenatal diagnosis. Linkage studies are based on accurate genetic counselling and clinical diagnosis of HPA in the affected

family member(s) and accurate understanding of the genetic relationships in the family. To carry out a linkage analysis it needs to obtain samples from multiple family members, including that from at least one affected individual. The markers (STR and VNTR) used for linkage analysis are highly informative because they are both intragenic and flanking to the PAH locus; thus, they can be used with 99% accuracy in PKU families (7).

Direct DNA sequencing: this is a routine procedure that in specialized clinical laboratories permits the discovery of disease-causing genes and underlying mutations. Single nucleotide substitutions leading to amino acid changes (missense mutations), translation stop signals (nonsense mutations), or aberrant exon/intron splicing, responsible for most disease-causing alterations of genes can be detect with a mutation detection rate of about 97% (13, 14).

Next Generation Sequencing: novel sequencing technology featured by high productivity and sensitivity in mutations detection. Their application to the study of the molecular basis of several human inherited diseases has demonstrated the effectiveness of the massively parallel sequencing approaches in generating high-quality data (20b). Recent findings also showed the quantitative nature of PCR amplification/high throughput sequencing approach and its ability in the simultaneous detection of both SNPs and CNVs in target genes, thus increasing the spectrum of detected variations (>99.5%) (25b).

1.10 Functional Test

To date, more than 600 gene variants have been identified in the PAH gene that contains 13 exons spread over 80 kb. The majority of PAH mutations are correlated to a specific biochemical/metabolic phenotype. For example, some “mild” PAH mutations reduce the affinity of the enzyme for the substrate Phe, other reduce the affinity of the enzyme for the BH4, other mutations cause instability of the structure of the enzyme. Data regarding the phenotypic impact of some mutations are either lacking or controversial (http://www.pahdb.mcgill.ca). This is particularly the case of mutations of new identification, because of the lack of functional analyses. To study the effects of mutations on PAH activity, the most reliable possibility is to perform expression studies using wild-type and mutant cDNA to evaluate the enzymatic activity of the PAH proteins. It is possible, for example, to measure the 14C radiolabeled L-Phe converted to L-Tyr. The residual activities of the mutant PAH enzymes can be then

estimated as the percentage of wild-type enzyme activity.

1.11 Prenatal diagnosis

Information of molecular diagnosis has helped families of HPA to decide different opinions to continue pregnancy.

Prenatal diagnosis in families at high risk permits an early genotype characterization allowing the prognosis of PAH deficiency. The molecular analysis of the PAH gene is conducted in both parents and in chorionic villi or amniocytes taken at the sixteenth week of gestation. Sequencing of PAH gene permits the elucidation of mutations causing HPAs. The analysis is performed on amniotic cells obtained at 16th week of gestation. Previous procedures for prenatal diagnosis and carrier screening for PKU have been based on the assessment of different RFLPs within the PAH gene, with a 30% of informativity in Chinese and Japanese PKU families and with a 80% of informativity in Caucasian PKU families. Successively, has been added VNTR system increasing the informativity level to 90%. Mutation detection in HPA disorder by PCR and direct sequencing of the 13 specific exons is a good diagnostic method. In addition, for the prenatal molecular diagnosis has advantages that it can be performed earlier, quicker and leads to an accurate diagnosis in a very short time. Finally, mutation analysis is a useful method for prenatal diagnosis, when the disease-causing mutations have been identified in the proband and ideally also confirmed in the parents.

1.12 PKU maternal Syndrome

In 1956, Charles Dent has been one of the first physicians to identify the teratogenic effects of maternal L-Phe on the foetus. These effects include mental retardation, facial dysmorphism, microcephaly, intrauterine growth retardation (IUGR), developmental delay and congenital heart disease (CHD) (31). In untreated PKU pregnancies with L-Phe levels ≥1200 μmol/l, more than 90% of the offspring have microcephaly and mental retardation, 40% have IUGR and 12 to 15% have CHD.

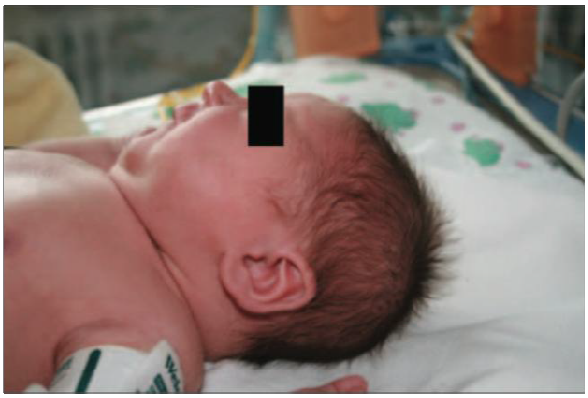

Fig. 9: The newborn suffering from maternal PKU syndrome.

The teratogenic effects are less recurrent when L-Phe is between 600 and 1200 µmol/l; in addition, when dietary treatment lowers the blood L-Phe level to 120 to 360 µmol/l, the offspring may be normal. Then there is a dose-response relation. The most important way to prevent teratogenic effects on the developing foetus is a good metabolic state achieved before conception. However, it is still advantageous to the foetus to decrease L-Phe levels within the first trimester in case of an unexpected pregnancy. Recently, it was suggest that fluctuations of L-Phe serum levels have significant impact on the offspring neuropsychometric outcome. However, few cases of untreated or poorly treated PKU pregnancies result in normal offspring, while some apparently well-treated pregnancies are not successful. The precise mechanism remains unclear.

1.13 Therapy

Conventional diet: worldwide, PKU diagnosis is performed in newborn screening programs, because a low-Phe diet is started early, mental retardation can be prevented (2, 33, 48). Difference in L-Phe tolerance is related not only to metabolic status but also to genetic background and to general health situation. Physical activity, growth, pregnancy, and infections may alter the needs for Phe. Therefore, the diet must be standardized for each patient so that:

1) Phe and its metabolites do not reach toxic concentrations
2) Intake of other amino acids must be adequate to provide the patient's metabolic needs.

The PKU diet composition has modified very little since its introduction in the 1950s; in fact, it is a low-protein diet

supplemented with a Phe-free mixture of amino acids, plus added minerals, vitamins and other nutrients. PKU diet is extremely restrictive and difficult to follow. Moreover, the Phe-free mixtures of amino acids have an unpleasant flavor and smell and must be consumed in relatively large amounts. The quality of life under the PKU diet is severely compromised. Therefore, alternatives to the PKU diet have been actively sought.

Nutritional deficiencies: The PKU patient management seeks to optimize the growth, the development, and the dietary compliance (33). In fact, the limitation of Phe determines many nutritional problems: the quality of the available amino acid integration, the neurotrophic and neuroprotective effects of added long-chain polyunsaturated fatty acids (LCPUFA), micronutrient deficiencies, bone disease, and antioxidant status. Then, the nutritional status of PKU patients should be regularly checked. Furthermore, a further benefit of BH4 supplementation is that it may reduce some of the nutritional deficits by increasing the natural protein intake. Adequate calcium and vitamin D integration is an important component of care. In addition, a vitamin B12 deficiency can occur when PKU patients rest the diet in adolescence. Finally, some of the PKU neuropsychological problems may be due to a deficiency of the amino acid Tyr. Furthermore, PKU subjects have to avoid aspartame, an artificial sweetener containing Phe.

Therapies under investigation

1. BH4 supplementation: There is debate in literature regarding the helpfulness of BH4 supplementation in patients bearing of some specific mutations or genotypes (27, 32). In fact, a relevant number of PKU patients show reduced plasma Phe levels after a BH4 loading test then, in these patients, an important alternative to diet is based on the BH4 supplementation. In fact, about 60% of the patients with plasma Phe levels between 400 and 800 µmol/L respond to the BH4. However, some patients with higher Phe levels also respond to BH4 therapy. A few hours after BH4 administration, blood Phe concentrations decrease considerably (1, 3, 44, 38). Nevertheless, a large group of patients with PAH mutations do respond to treatment with BH4 (3, 32, 36). Furthermore, it has been shown that long-term BH4 treatment may also raise the Phe tolerance in a lot of patients with severe PKU phenotype (27). In addition, it has been hypothesized that supplementation with BH4 combined with the diet would be sufficient also for the prevention of the maternal PKU syndrome (47).

2. Large neutral amino acids supplementation: Phe, as well as other large neutral amino acids are transported across the blood-brain barrier by L-type amino acid carrier. High Phe levels reduce the brain uptake of other LNAAs (Fig. 9) (30). Some LNAAs such as tyrosine and tryptophan are precursors of neurotransmitters, and then an impaired synthesis of neurotransmitter would be an further component contributing to the PKU cognitive dysfunction (35).

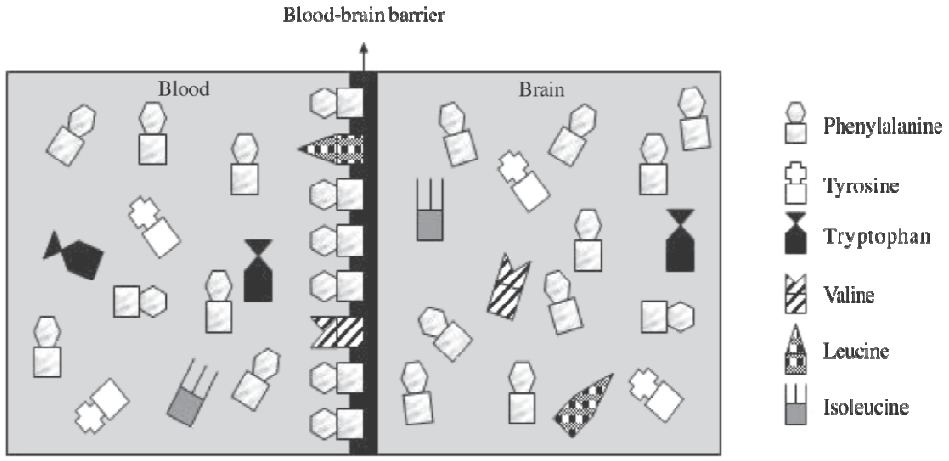

Fig. 10: Large neutral amino acids (LNAAs): asparagine, cysteine, glutamine, histidine, isoleucine, leucine, methionine, serine, threonine, tyrosine, tryptophan, and valine) are transported across the blood-brain barrier by means of L-type amino acid carrier.

3. Enzyme Replacement: Enzyme substitution therapy with phenylalanine ammonia lyase (PAL; E.C.4.3.1.5) is under intensive investigation as a possible alternative PKU treatment (2). PAL deaminates L-Phe to free ammonia and trans-cinnamic acid (Fig. 10). In humans, trans-cinnamic acid is safely and rapidly converted to hippuric acid excreted in the urine. If compared to PAH, PAL is structurally and catalytically less complex than PAH, physically more stable, and does not require a cofactor. Simple subcutaneous injection of the PAL into Pahenu2 mice, a model of human PKU, yielded complete correction of blood phenylalanine concentrations. In a Phase 1 trial designed to evaluate security of subcutaneous rAvPAL-PEG injection and carried out at seven centers in the US.

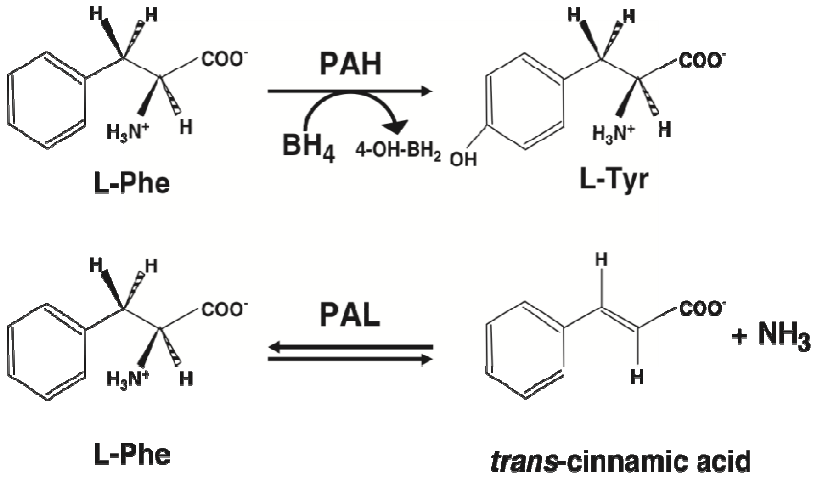

Fig. 11: Differences between PAH and PAL enzymes. PAH requires the presence of its cofactor, BH4, as well as a Fe+ ion atom and oxygen. PAL does not need a cofactor.

4. Gene therapy: Helper Dependent-Adenovirus vector-mediated gene therapy is a promising approach to PKU treatment. This approach resulted in a complete normalization of Phe and Tyr levels and reversal of coat pigmentation effects permanent throughout six months period.

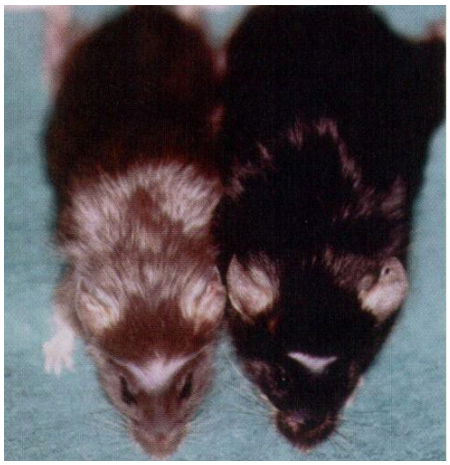

Fig. 12: Changes in coat color in PKU mice after gene therapy.

The spatial learning deficits observed in PKU mice were also reversed and hippocampus levels of the N-methyl-D-Aspartate and 2-amino-3-(5-methyl-3-oxo-1,2- oxazol-4-yl) propanoic acid receptor subunits returned to normal. Long-term potentiation, which is impaired in PKU mice, was also restored by treatment (8, 21).

In addition to gene therapy, some studies with PAH expression in tissues other than liver have been developed for PKU. Epidermal keratinocytes and dermal fibroblasts were infected with retroviral vectors expressing PAH and/or GTP-cyclohydrolase (one of the enzymes involved in BH4 synthesis) gene obtaining high Phe clearance (11).

5. Therapies under investigation: different attempts have been performed to find other effective treatment modalities for PKU. Glycomacropeptide (GMP) is a low Phe protein derived from cheese. Recently, has been reported improved nutritional management of PKU by using a diet with GMP, if compared with conventional amino acid formulas. GMP may represent an alternative safe therapy in the nutritional management of PKU (2).

6. Somatic gene therapy: this approach it is being tested in animal models and is promising as a potential treatment (11).

7. Induced pluripotent stem cells: hepatocyte or hematopoietic stem-cell transplanted with normal PAH cDNA and liver repopulation with these cells, represent a further treatment approach under investigation. In an animal model, this approach has effectively correct HPA phenotype (2).

2. Aim of the thesis

Molecular analysis of the PAH gene is useful for several reasons:

1. Identification of PAH mutations.
2. Identification of PKU carriers (usually asymptomatic).
3. Prenatal diagnosis for the pregnancies at high risk.
4. Prediction of the biochemical phenotype.
5. Identification of BH4-sensitive mutations that may lead to a patient-rational treatment.
6. Sequencing of the entire PAH gene (including 22000 bp at the 5' end of the transcription start site, exon regions, intron regions and 1000 bp at the 3' end).

Moreover, due to the wide heterogeneity of mutations in the PAH gene, it makes essential to define the molecular epidemiology of the mutations responsible for both PKU and HPA patients (in individual ethnic groups).

The aim of my thesis work was to perform molecular analysis of the PAH gene in 390 PKU patients and 373 carriers that come from the Campania Region Screening Center ASL 1. The analysis of PAH mutations allowed to define the molecular epidemiology of mutations in the PAH gene in Campania and compare the obtained data with those from other regions of Italy (Southern Italy, i.e. Calabria, Sicilia, Puglia and Basilicata, Northern Italy; i.e. Piemonte) and Europe (Spain and Portugal).

An additional aim of the present thesis was to develop and validate a diagnostic approach, based on next generation sequencing (NGS), able to identify novel causing-disease mutations in PAH patients with undetermined alleles, thus increasing our detection rate and improving our ability to achieve a genetically based etiologic diagnosis of PAH.

3. Materials and methods

3.1 Patients

We studied 390 Caucasian HPA unrelated patients from Southern Italy (98% from the Campania region; median age 15 years, range 2–25 years; male:female ratio 1.2:1). Moreover, we performed the molecular analysis for the detection of carrier status in 373 subjects.

Patients were classified on the basis of pre-treatment plasma Phe concentrations and Phe tolerance into HPAI or 'classic PKU' (pre-treatment Phe levels > 1200 mmol/L, Phe tolerance: 250–350 mg/L); HPAII (pre-treatment Phe levels in the range 600–1200 mmol/L, Phe tolerance: 350–600 mg/L); HPAIII (pre-treatment Phe levels < 600 mmol/L, Phe tolerance: > 600 mg/L). The HPAIII category included five patients whose Phe levels were < 360 mmol/L under a Phe unrestricted diet. Tolerance was defined in patients > 2 years of age as the highest Phe intake that was able to maintain plasma Phe levels within the safe range (120–360 mmol/L). In the case of discrepancies between pre-treatment plasma Phe concentrations and Phe tolerance, the phenotypic class was assigned according to Phe tolerance data. 390 patients were identified by a neonatal screening program and two patients who were born in the pre-screening era were diagnosed after the identification of mental retardation.

The patients were classified: 121 (15.8%) PKU, 64 (8.4%) HPAII and 205 (26.9%) HPAIII

The parents of each patient were informed about the aims, methods and possible results of DNA analysis and they gave the informed consent to the study.

3.2 Prenatal diagnosis

We performed the prenatal diagnosis in three families at high risk for HPA. Amniotic cells or corionic villi of the feti were obtained at 16th week of gestation, tripsinized, washed and centrifuged. The obtained pellets were used for genomic DNA extraction. Contamination of maternal material in amniotic fluid has been excluded through methods based on DNA polymorphisms by analysis of polymorphisms in the human genome repeat (STR). The sixteen STRs used for the survey are listed in the following Table:

	STR	Cromosom
1	D8S1179	8
2	D21S11	21q11.2-q21
3	D7S820	7q
4	CSF1PO	5q33.3-34
5	D3S1358	3p
6	THO1	11p15.5
7	D13S317	13q22-31
8	D16S539	16q24-qter
9	D2S1338	2q35-37.1
10	D19S433	19q12-13.1
11	vWA	12p12-pter
12	TPOX	2p23-2pter
13	D18S51	18q21.3
14	D5S818	5q21-31
15	FGA	4q28
16	Amel	X, Y

Successively, the genomic DNA was used for the amplification reaction (PCR) and finally for the direct sequencing of the 13 exons of PAH gene.

3.3 DNA extraction and quantification

A blood sample (1 mL) was collected by venipuncture in EDTA (ethylene-diamine-tetra-acetate). The genomic DNA was extracted from white blood cells, after cell lysis and deproteinization using sodium perchlorate by the commercial kit “Nucleon BACC2” of the Amersham Life Science. For each sample, the extracted DNA was quantified by spectrophotometric reading at 260nm. The quality of the extraction was evaluated by checking the relationship between the reading at 260nm and 280nm (optimal ratio ≥1.8). DNA samples were prepared in the work concentration of 60 ng/µl.

3.4 PCR

All coding regions and exon-flanking intronic sequence of the PAH gene of each DNA sample was analyzed by PCR, with which were amplified all the coding regions of the PAH gene. The PCR products are monitored by electrophoretic analysis on agarose gel at 1.5% and subsequently analyzed by direct sequencing. The PCR reaction is carried out in a final volume of 50 .00| on Thermal Cycler 2720 of "Applied Biosystems". The protocol used is the "Touchdown", with which the annealing temperature is progressively lowered by 0.5 ° C at each cycle, so that at the beginning the primers anneal only with sequences perfectly complementary. The reactions of chain polymerization were carried out using the amplification conditions summarized in Table 1 The

optimal program for the PCR amplification is summarized in Table 2.

Reagent	Initial Concentration	Final Concentration
PCR Buffer	10x	1x
$MgCl_2$	25mM	1.5 mM
dNTPs	2.5 mM	100 µM
Oligo Forward	15µM	0.6 µM
Oligo Reverse	15µM	0.6 µM
Taq Polimerase	5U/µl	0.04 U/µl
DNA	60 ng/ µl	5ng/ µl
H_2O	q.b	q.b

Table 1: amplification conditions for PCR

Temperature	Time	Cycle
94° C	2 min.	1
94° C	20 sec.	14
T_A℃ + 7℃	40 sec. ↓ 0.5℃/cycle	
65℃	45 sec.	
94° C	20 sec.	25
T_A℃	40 sec.	
65℃	45 sec.	
65℃	7 min.	1

Table 2: amplification program for PCR

T_A is the annealing temperature of primers, calculated according to the following formula:

$$T_A = \{[(A+T)\bullet 2+(G+C)\bullet 4]-5\}℃$$

A+T = Adenosine + Timidine
G+C = Guanidine + Citidine

3.5 Primer design

PCR primers were designed to generate amplicons to cover the coding regions and exon-flanking intronic sequences of the PAH gene. Specific primers were generated using the web-based program Primer3 (http://frodo.wi.mit.edu/primer3/) (v.0.4). The criteria for primers were: 20 base pairs in length; melting temperature (Tm) of 57-60 ℃; and G/C content 45-60%. Primers were also checked by BLASTn against the NCBI genome data bank for specificity (http://www.ncbi.nlm.nich.gov/tools/primer-blast/) and were tested for loop formation by using the DINA Melt web server (http://mfold.rna.albani.edu/?q=DINAMelt/Two-state-folding). The primer sequences are reported in Table 3.

EXONS	FORWARD	REVERSE
1	**5'**TACCTGAGGCCCTAAAAAGC**3'**	**5'**CTTCGGATCTCTTTCTCTGGA**3'**
2	**5'**ATGCTTGCTTTGTCCATGG**3'**	**5'**CAAATCTGCCTGTTCCAGATC**3'**
3	**5'**GGACGTTGCCTTCTCTGTGT**3'**"	**5'**TGTTGCAAAATTCCTCTAATTCTT**3'**
4	**5'**GTTATCTGGAAGCCAGCCC**3'**	**5'**GCCCTCGTGTAAATAGGAACA**3'**
5	**5'**ACCAGGCCTCTTCCTATGAA**3'**	**5'**TCTTCCCCTCAACAAGCAA**3'**
6	**5'**GCCCTGCTTGAGACACCTAT**3'**	**5'**CATTGACCCTGATGTGGACTT**3'**
7	**5'**CTCAGTGGTGATGAGCTTTGA**3'**	**5'**TCATTCTTGCAGCAGGAAAA**3'**
8	**5'**TCTGGCTTGGCTTAAACCTC**3'**	**5'**CTCCCTGGGCTCAACTCAT**3'**
9	**5'**GGTTGGTTCTGTGGTTCCA**3'**	**5'**CCTATAGCACTCCACCATCCA**3'**
10	**5'**TCCCTTCATCCAGTCAAGGT**3'**	**5'**CTGGAGAATGAGTTCCCAGG**3'**
11	**5'**GCAGCAGGGAATACTGATCC**3'**	**5'**CCAGAGCTAGTGGCTCACCT**3'**
12	**5'**TTCTCCAAATGGTGCCCTT**3'**	**5'**TGAGAAACCGAGTGGCCTC**3'**
13	**5'**TGCTTTGCACTGAGGACACT**3'**	**5'**TGCTTTTCGGACTTTTTCTGA**3'**

Table 3: Primer sequences for the amplification of the 13 exons of PAH gene.

3.6 Direct sequencing

The direct sequencing was performed using the protocol of Sanger; the fragments obtained were subsequently analyzed on both strands with an automated procedure using the 3100 Genetic Analyzer (Applied Biosystem). All PCR fragments were sequenced using the same primers used for PCR amplification. The results are represented by electropherograms (Fig. 13).

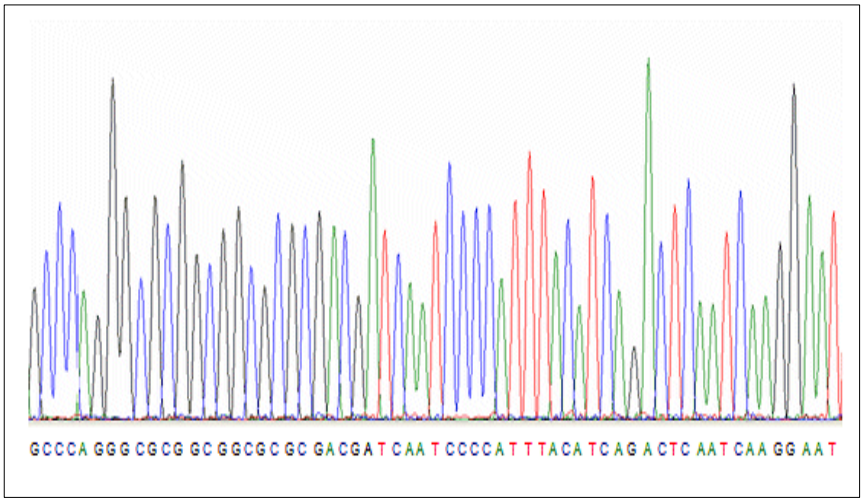

Fig. 13: Example of electropherogram obtained by Genetic Analyzer.

3.7 Next Generation Sequencing

To perform the comprehensive sequence analysis of the entire PAH gene by a NGS approach, we used a Long-PCR strategy to enrich our target gene. Totally, we analyzed 106.270 bp, including 22.000 bp at the 5' end of the transcription start site and 1000 bp at the 3' end of the coding gene. The genomic coordinates of PAH gene were derived from the website "http://genome.ucsc.edu/" (NM 000277.1, Human Genome assembly GRCh37/hg19, Feb.2009, genome coordinates Chr 12: 103,227,104-103,333,381). Sixteen overlapping Long-PCR amplicons, ranging in size between 3-12 kb, were generated in order to cover the promoter and all the coding and non-coding regions of PAH gene. Specific primers were generated using Primer3 software "http://frodo.wi.mit.edu/primer3/". To design an efficient and gene-specific primer set, known SNPs and regions of the gene that contain repeated sequences were masked. Primers inclusion criteria were: 20 base pairs in length; melting temperature (Tm) of 57–60 °C; G/C content of 45–60%. Primers were also checked by BLASTn against the

NCBI genome data bank for specificity (http://www.ncbi.nlm.nich.gov/tools/primer-blast/) and were tested for loop formation by using the DINAMelt web server (http://mfold.rna.albani.edu/?q=DINAMelt/Two-state-folding). Consecutive primer sets were designed to overlap at least 150-600 base pairs to ensure good coverage and continuity at the ends of amplicon sequences. The amplification conditions of each amplicon were optimized using control DNAs. Primer testing was performed using 60-120 ng of human genomic DNA in a PCR core mix containing 200 nM of each forward and reverse primers, 250 mM each dNTP, 50 ng/ml BSA, 1X PFU Ultra II Reaction Buffer (Agilent Technologies. Inc.) and two units of PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies. Inc.). PCR primers and amplicon lengths are listed in Table 4. Thermal cycling reactions were performed on a DNA Engine ® (Bio-Rad Laboratories, Hercules, CA) using the following "touch down" cycling profile: one cycle at 92℃ for 2 min (initial denaturation), followed by five cycles at 92℃ for 10 s (denaturation), at 62℃ for 20 s (annealing; -1℃ every cycle) at 68℃ for 4 min (extension), followed by 25 cycles at 92℃ for 10 s (denaturation), at 58℃ for 20 s (annealing), at 68℃ for 4 min (extension), followed by a final extension step at 68℃ for 5 min. The above mentioned PCR amplification conditions were used for the following amplicons: PKU1, PKU2, PKU3, PKU4a and the ones at the 3' end PKU 10 and PKU 11 (see Table 4).

For amplicons that failed to be amplified, we attempted three additional modified thermal cycling conditions. The first profile reduced the annealing temperature for the first cycle after the initial denaturation to 65℃ with a further -1℃ reduction of the annealing temperature every cycle for the next 4 cycles. The second profile reduced the annealing temperature for the first cycle after the initial denaturation to 60℃ with a further -1℃ reduction of the annealing temperature every cycle for the next 4 cycles. The third profile, regarding 6b and 9a exons (see Table 4), increased the extension times of the last 25 cycles by 8 seconds every cycle. Any primer pair that failed to amplify with all four PCR profiles was discarded and new primers were designed. All the PCR products were visualized on 0.7% agarose gel.

Amplicon	*Region size (bp)*	*Primer*	*Primer sequences*
PKU1	**11706**	**1F**	***5'*-GGGCTGTCAGTGGAAATGTT-*3'***
		1R	***5'*-TGGTGCCTGATGCACATTAT-*3'***
PKU2	**11644**	**2F**	***5'*-CTTCTGCCTGGAAGGTTCTG-*3'***
		2R	***5'*-TATTCGCACAACTGCTCCTG-*3'***
PKU3	**9975**	**3F**	***5'*-TGCCTCACCCTCAGTCTTCT-*3'***
		3R	***5'*-TCTTGAGTGCATTCCATCCA-*3'***
PKU4a	**11724**	**4aF**	***5'*-GTTCATGCTTGCTTTGTCCA-*3'***
		4aR	***5'*-GTGGGCTATCTGGGTTCAAA-*3'***
PKU4b	**4545**	**4bF**	***5'*-CCCTCCCAGAAGAGGAAATC-*3'***
		4bR	***5'*-ACTGAACCCCCAACACAAAG-*3'***
PKU5a	**4620**	**5aF**	***5'*-CCTCCCTTCCAAAACTCTCC-*3'***
		5aR	***5'*-TGCTGTTATTTTATGAAGACAGTGTG-*3'***
PKU5b	**5719**	**5bF**	***5'*-TCTCCATTTTGTTGCGTTAGG-*3'***
		5bR	***5'*-GAAGAAACTCAAGAAGGCAAGG-*3'***
PKU6a	**6983**	**6aF**	***5'*-CGAGTCTGACAGCAGGTTCA-*3'***
		6aR	***5'*-AATGCACAGCCTCATTTTCC-*3'***
PKU6b	**6158**	**6bF**	***5'*-AGGTGGCCCGATAAGAATTT-*3'***
		6bR	***5'*-TCCCAGCCCTCGTGTAAATA-*3'***
PKU7	**5119**	**7bF**	***5'*-GATCCCCACTTCTGATCTCA-*3'***
		7bR	***5'*-GAACTCTGACAGGGCCTCAG-*3'***
PKU8a	**3428**	**8aF**	***5'*-TTCACTTCCAATGTGGGTCA-*3'***
		8aR	***5'*-AGGGGATAGAAGACGGGAGA-*3'***
PKU8b	**7708**	**8bF**	***5'*-TGGTTTAAAAGCCCCTTGTG-*3'***
		8bR	***5'*-CTTGTGGTAGCCAGCAATGA-*3'***
PKU9a	**8092**	**9aF**	***5'*-CTTGGGCAAGAGCTTACCTG-*3'***
		9aR	***5'*-CCCCCATGTTCTCTGTGTCT-*3'***
PKU9b	**3759**	**9bF**	***5'*-GCCCTGCTTGAGACACCTAT-*3'***
		9bR	***5'*-CTCCCTGGGCTCAACTCAT-*3'***
PKU10	**9792**	**10F**	***5'*-CTTAAGACCCTTGGCTGCTG-*3'***
		10R	***5'*-CTGCCTCTAAGCTCCAATGC-*3'***
PKU11	**8762**	**11F**	***5'*-TCTGAGACTGTTGGCCCTCT-*3'***
		11R	***5'*-CAACAAATGTGGGTGCTGTC-*3'***

Table 4: Primer sequences and amplicon lengths of the 16 overlapping regions PAH gene to be sequenced by NGS.

Each amplicon was individually amplified and purified using an equal volume of AMPure Reagent beads (Agencourt, Beverly, MA) according to the manufacturer's protocol. After

appropriate quality assessment (DNA 7500 LabChip, 2100 BioAnalyzer, Agilent) and quantification (Picogreen assay, Invitrogen), the amplification products from the same DNA sample were pooled in equimolar ratio. Each amplicon pool was used to generate a library to be processed using the GS FLX System (454 Life Science and Roche) by shotgun approach, according to the specifications of the manufacturer. Briefly, 5 µg of amplified and pooled DNA from each patient were nebulized into small fragments, purified and assessed for quality. Blunt-ends were generated and specific adaptors with Multiplex Identifiers (MID) were ligated to the ends of each fragment. MID adaptors include a sequence tag which is unique for each MID and allow the simultaneous sequence of different patients. So, we generated a single stranded library with a specific MID for each sample. Then, different libraries were pooled together, immobilized on the surface of microscopic beads and clonally amplified within the droplets of a water-in-oil emulsion. Amplification of the entire fragments collection was done in parallel, resulting for each fragment in a copy number of several million per bead. The clonally amplified fragments were enriched, loaded into wells of a fibre-optic slide (PTP), and bi-directionally sequenced using a pyrosequencing protocol (33 b).

3.8 Bioinformatic Analysis

All the obtained sequencing reads were quality filtered and trimmed by the data analysis software, so that sequencing adaptors were removed and the reads from each of the pooled libraries were identified by their MID tag and correctly assigned to each analyzed sample. Next, the filtered reads were mapped against the human genome reference sequence to generate a list of variants for further investigations.

4. Results

4.1 Molecular diagnosis of PAH mutations

We have analyzed 390 PKU patients and 373 carriers. Patients are mainly from the centre of Screening Annunziata Hospital of Naples and, based on the L-Phe concentrations found in the blood, were divided into three groups: 121 were classified as (15.8%) PKU, 64 (8.4%) as HPAII and 205 (26.9 %) as HPAIII.

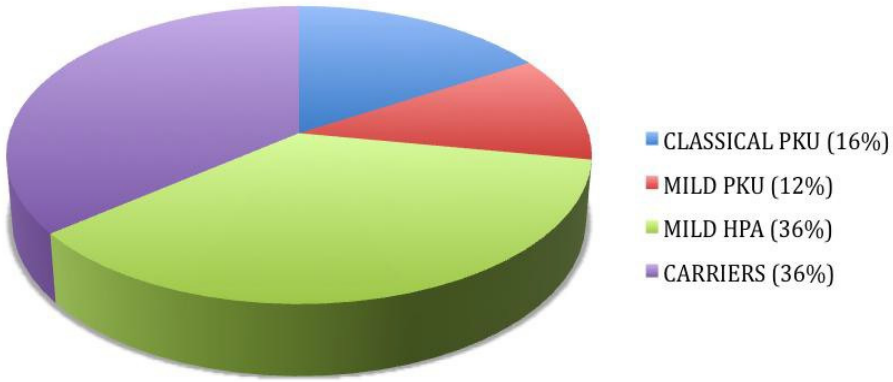

Fig. 14: Graphical representation of studied population. We analyzed both patients and carriers.

Molecular analysis identified mutations in 763/780 alleles with an allele detection rate of 97.8%. Complete sequencing of the 13 exons, the intron–exon boundaries and the promoter region of the PAH gene was carried out. Complete genotyping was carried out in 763/780 alleles. In seventeen patients only one causative mutation was found.

The failure to identify a mutation in one of the 2 alleles may be due to:

1) Alterations present in the promoter region or in the region of the 3' untranslated
2) Molecular defects that affect introns
3) Mutations of an enhancer
4) Large deletions or duplications of the PAH gene

A total of 105 distinct mutations were identified and these were not uniformly distributed along the PAH gene sequence (Table 4).

Of the 105 variants, 76 were missense mutations (74%), 20 alter splicing site (18%), 5 were deletions (4%) and 4 were frame-shift (4%). The data obtained are in agreement with those reported in the literature and

confirm the highest frequency of the missense mutations in PKU (60.68%).

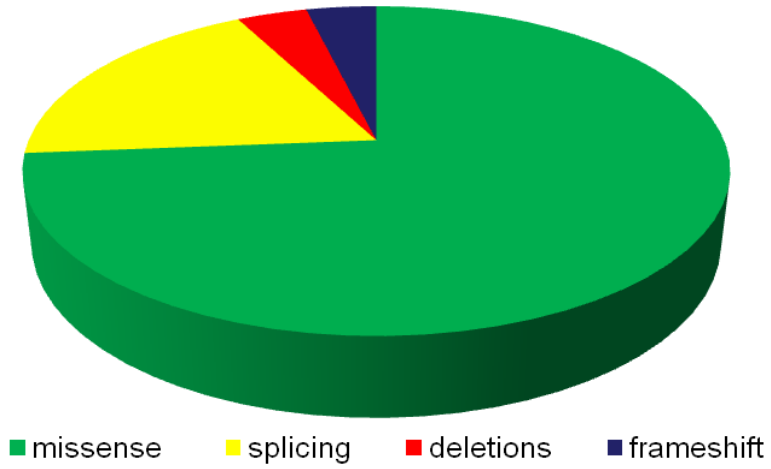

Fig. 15: Graphical representation of mutations of PAH gene in PKU patients

4.2 Prenatal Diagnosis

We performed prenatal diagnosis in three families at high risk for HPA:

1. The prenatal diagnosis has been requested for a foetus in a family with an affected baby. Therefore, we performed the molecular analysis of the PAH gene of the affected son, of the parents and finally of the baby. The molecular analysis of the affected son revealed the genotype R261Q/P281L. Molecular analysis performed on DNA extracted from peripheral blood of the parents showed the presence of the R261Q mutation for the father and mutation P281L for the mother in the heterozygous state, so the genotype is compatible with the healthy carrier state. Finally, the molecular diagnosis performed on DNA extracted from amniocytes has demonstrated the absence of both mutations R261Q and P281L, previously identified in the parents, providing a healthy phenotype for phenylketonuria.
2. The prenatal diagnosis has been requested for a foetus with an affected brother. Molecular analysis performed on DNA extracted from peripheral blood of the parents demonstrated the presence of the mutation in the heterozygous p.P281L in the mother and p.A403V in the father, so the genotypes are compatible with the carrier state. The molecular diagnosis performed on DNA extracted from chorionic villi showed the presence of

mutations p.A403V and p.P281L both in the heterozygous allowing, therefore, to provide a genotype p.A403V/p.P281L compatible with a phenotype affected by hyperphenylalaninemia.

3. The prenatal diagnosis has been requested for a foetus by an affected mother. Molecular analysis performed on DNA extracted from peripheral blood of the mother demonstrated the presence of the mutations p.W187C and p.R261Q, both heterozygous, so the genotype is p.W187C/p.R261Q, and this genotype is compatible with the state of suffering from phenylketonuria. On the contrary, the molecular analysis performed on DNA extracted from peripheral blood of the father showed no causal mutations of disease. It should be emphasized, however, that the analysis of sequencing used includes the regions comprising the 13 exons and exon junctions-intron of the PAH gene, therefore, it is not possible to exclude mutations present in other regions of the PAH gene, not identifiable with current techniques in use. The molecular diagnosis performed on DNA extracted from amniocytes showed the presence of the mutation in heterozygous p.R261Q and the absence of the mutation p.W187C, allowing, therefore, to predict a phenotype of bearer not suffering from phenylketonuria.In this case, we suggested to family to control the Phe mother during the pregnancy through a Phe restricted diet in order to prevent damages from Maternal PKU.

4.3 Molecular epidemiology of PAH mutations

The mutations identified in the 390 patients and their frequencies are shown in Table 5 which shows the data for the 780 alleles studied (column 7), and those related to three classes of PKU patients.

The frequencies of the mutations p.R261Q, p.A403V and c.1066-11G> A were higher than 10%, (cumulative frequency= 24.8%); the mutations p.L48S and p.A300S had a frequency in the range of 6-10%, (cumulative frequency = 13.6%); 51 mutations in the range of 0.2%-3.2% (cumulative frequency = 56.9%); 47 mutations have been found in a single mutant allele with a frequency of 0.1%(cumulative frequency = 4.7%). In 17 patients (4.3%) only one mutation was identified. The major part of the mutations (*n*=81) was located along the catalytic domain (78%), 18 mutations were found in the regulatory domain (17.3%) and 5 belonged to the tetrameric domain (4.8%).

The mutation p.R261Q, in the exon 7 is the most common mutation in our population (13.4%) being present in 89 of these, 35 patients with PKU phenotype (27 heterozygous and 8 homozygous), in 21 patients with HPAII phenotype

(14 heterozygous and 7 homozygous) and in 33 patients with HPAIII phenotype (32 heterozygous and 1 homozygous). The R261Q mutation is due to a transition G>A, in position 782 to the nucleotide sequence, which corresponds to an amino acid substitution (Arg>Gln), resulting in a lack of interaction of the catalytic site with the cofactor BH4 (45).

The second most frequent mutation in Campania is the p.A403V (11.4%) in exon 12. The mutation is present in 76 subjects with HPAIII phenotype (71 heterozygotes and 5 homozygotes) (18.5%), in 7 patients with HPAII (5.4%) and in 1 patient with PKU (0.4%); therefore this mutation appears to be strongly related to the mild phenotype. p.A403V mutation is characterized by an amino acid substitution (Ala>Val) caused by a transition C>T in position 1222. The mutation is not directly involved in alterations of the catalytic activity, in fact falls within the group of mutations that alter the mechanism of binding of the cofactor; in fact, the exon 12 encodes for some elements of the secondary structure of PAH necessary for the interaction with the BH4 and the mutations determine the production of proteins with decreased affinity for the cofactor (45).

The mutation c.1066-11G> A is the third most frequent mutation (11.1%) and appears to be more closely related to the phenotype PKU. In fact, it was found in 44 PKU patients (18.1%) (33 heterozygous and 11 homozygous), in 15 patients with HPAII (11.7%) and in 17 HPAIII patients (4.1%). The mutation, which falls in exon 10, causes a splicing site alteration with the insertion of 9 nucleotides between exon 10 and exon 11 leading to the production of a protein with 3 additional amino acids (Gly-Leu-Gln). The insertion of these 3 amino acids probably alters the structure of the protein, which forms protein aggregates and loses its functionality (46).

Another very frequent mutation in Campania is the p.L48S (7.6%). This mutation was reported in 21 PKU patients (8.7%), in 11 HPAII patients (8.6%) and in 21 HPAIII patients (5.1%) (18 heterozygotes and 3 homozygotes), so it is mainly related to PKU phenotype and HPAII. p.L48S falling in exon 2 is probably involved in alterated mechanism of auto-regulation of the enzyme. p.L48S is a mutation that replaces the amino acid Leu in Ser; it is due to a transition T>C in position 143 of the nucleotide sequence. Finally the last most frequent mutation detected in Campania is the A300S (6%). A300S is due to a G>T at position 898 of the nucleotide sequence of the gene, which causes an amino acid substitution (Ala>Ser) and directly alters the binding site of the enzyme for its cofactor thus determining a reduced affinity for BH4 (42).

It was reported that mutations in exon 8 are related to changes in the catalytic activity as well as to a decrease interaction PAH-BH4.

Exon	Protein Domain and Subdomain	Mutation	PKU (242)	HPAII (128)	HPAIII (410)	Total (780)
1		c.47-48delCT	0,4	0,7	0,2	0,4
2	Regulatory Domain nt 1-426 aa 1-142	p.F39del	0	0	0,5	0,2
		p.L48S	8,6	10,9	5,6	7,6
		p.F55L	0	0	0,5	0,2
		c.168+1G>A	0	1,56	0	0,2
		c.168+5G>C	0,4	0	0,2	0,2
		p.R53H	0	0	0,2	0,1
		p.F55>Lfs	5	0,8	1,2	2,3
3	Auto Regulatory Domain nt 55-99 aa 19-33	p.T921	0	0,8	0,2	0,2
		p.E76X	0,4	0	0,2	0,2
		p.I95_K96delinsK	0	0	0,2	0,1
		p.I95T	0	0	0,2	0,1
		p.I95del	0,4	0	1,5	0,9
		p.L62X	0	0	0,2	0,1
		p.S67P	1,2	0	0,5	0,6
		p.I65T	0	0,8	0	0,1
		p.I65M	0	0	0,2	0,1
		p.I94S	0	0	0,5	0,2
4		c.441+5G>T	2	0,8	1,2	1,4
	Catalytic Domain nt 427-1230 aa 143-410	p.N133_Q134>Rfs	0,4	0	0	0,1
		c.441+1G>A	0,4	0	0,7	0,5
		p.D145V	0	0	0,2	0,1
		p.Q134X	0,4	0	0,2	0,2
5	Cofactor Binding Region # 1 nt 733-798 aa 245-266	p.R158Q	4	3	1,2	2,4
		p.R158W	0,4	1,6	0	0,4
		p.R155H	0	0	0,2	0,1
		p.R169H	0	0	0,2	0,1
		c.442-5C>G	0	1,6	2	1,3
		c.442-2A>C	0,4	0,8	0,2	0,4
		p.G148S	0	0,8	0	0,1
6	Cofactor Binding Region # 2 nt-838-849 aa 280-283	p.D222G	0,8	0	1,4	1
		p.N223Y	0	0	0,5	0,2
		p.S231P	1,2	0,8	1,5	1,3
		p.N207S	0	0,8	0	0,1
		p.E178G	0	0,8	1	0,6
	Cofactor Binding Region # 3 nt 964-978 aa 322-326	p.R176L	0	0	0,1	0,1
		p.R176X	0,8	0,8	0,8	0,7
		p.V230I	0	0,8	1,5	0,8

	Cofactor Binding Region # 4 nt 1129-1137 aa 377-379	p.G218V	0	0	0,2	0,1
		p.V177M	0	0	0,8	0,4
		p.W187C	0,4	0	0	0,1
		p.Q232Q	0,4	0	0	0,1
		p.W187X	0,8	0,8	0,8	0,8
		p.Q235X	0,4	0	0	0,1
		p.P211T	0	0	0,2	0,1
		p.L213P	0,8	0	0,8	0,6
		p.V177L	0	0	0,2	0,1
		c.592-613del22	0,8	0	0	0,2
7		p.R261Q	17.07	22	8,3	13,4
		p.R261X	2,4	1,5	0	1
		c.842+3G>C	1,2		1,2	1
		p.R252Q	0	0,8	0,2	0,2
		p.R243Q	0	0	0,2	0,1
		p.P275L	0	0,8	0	0,1
		c.842+1G>A	0,4	0	0	0,1
		p.281L	5,8	6,2	0,8	3,2
		p.L242F	0	0,8	0	0,1
		p.G272X	0,8	0	0	0,2
		p.243Q	0,4	0	0	0,1
		p.R241C	0	0	0,5	0,2
		p.E280K	0,8	0	0,2	0,4
		p.R252W	1,6	0	0,8	0,9
		p.R243X	0,4	0	0,5	0,4
		p.R241H	0	0	0,2	0,1
		p.V245A	0	0,8	1,2	0,8
		c.707-2delA	0,4	0	0	0,1
8		p.A300S	0	4,6	9,7	6
		p.I283N	0	0	0,2	0,1
		p.R297L	0	0	0,2	0,1
		p.Q301P	0,4	0	0	0,1
		p.S303A	0	0	0,2	0,1
9		c.969+43G>T	0	0	0,5	0,2
		p.L321I	0	0	0,2	0,1
		p.A309V	0	0,8	0	0,1
		c.913-7A>G	0	0	0,2	0,1
		p.I306V	0	0	0,2	0,1
10		**p.W326S**	0	0,8	0	0,1
		c.1055delG	1,6	0	0	0,5
		p.G352fs	2,9	0	0,8	1,3
		p.D338Y	0	2,3	1,2	1
11		c.1066-11G>A	22,7	11.07	4	11,10
		p.K363fs	0,4	0	0	0,1
		p.A395G	0	0	0,2	0,1

		p.E390G	0	2,3	2	1,4
		p.P366H	0,8	0	3,6	2,2
		p.F382L	0	0	0,2	0,1
		p.V388M	0	0	0,5	0,2
		p.T380M	0	0	2	1
		c.1066nt4A>C	0	0	0,2	0,1
		p.Y387H	0	0	0,2	0,1
		c.1199nt4A>C	0	0	0,2	0,1
		c.1066-3C>T	0	0,8	0	0,1
		p.K398N	0	0	0,2	0,1
		c.1315nt2T>C	0	0,8	0	0,1
12	Tetramerizazion Domain nt 1231-1356 aa 411-452	c.1315+1G>A	0,8	1,5	0,2	0,6
		p.A403V	0,4	5,4	20	11,4
		p.Y414C	0	1,5	1,5	1
		p.N426H	0	0	0,2	0,1
		p.R408Q	0	0	0,5	0,2
		p.R408W	3	0,8	1,2	1,6
		p.I406M	0	0	0,2	0,1
		p.Q419R	0	0	0,2	0,1
		p.T418I	0	0	0,5	0,2
13			0	0	0	0

Table 5: Alleles and relative frequencies of 103 mutations in PAH gene in 780 alleles of HPAs patients in Campania. Novel mutations are highlighted in red;

Flow chart definition

Based on the epidemiologic data and according to phenotype of patients, we created a comprehensive flow chart for the detection of nucleotide alterations (Fig. 16). This flow chart represents a simple, elegant and less expensive step-by-step procedure for reliable detection of PAH mutations. The method is simple to perform and easy to implement in a routine diagnostic laboratory. Based on the biochemical phenotype of the patient who arrives at our lab, we firstly proceed to the amplification and sequencing of exons more implicated (with high frequency) in that class of patients. Successively, if the mutations responsible for the disease are not identified, we proceed to the analysis of exons gradually less frequently mutated in the specific phenotype. This method can reduce both costs and analysis time of patients starting from their biochemical phenotype.

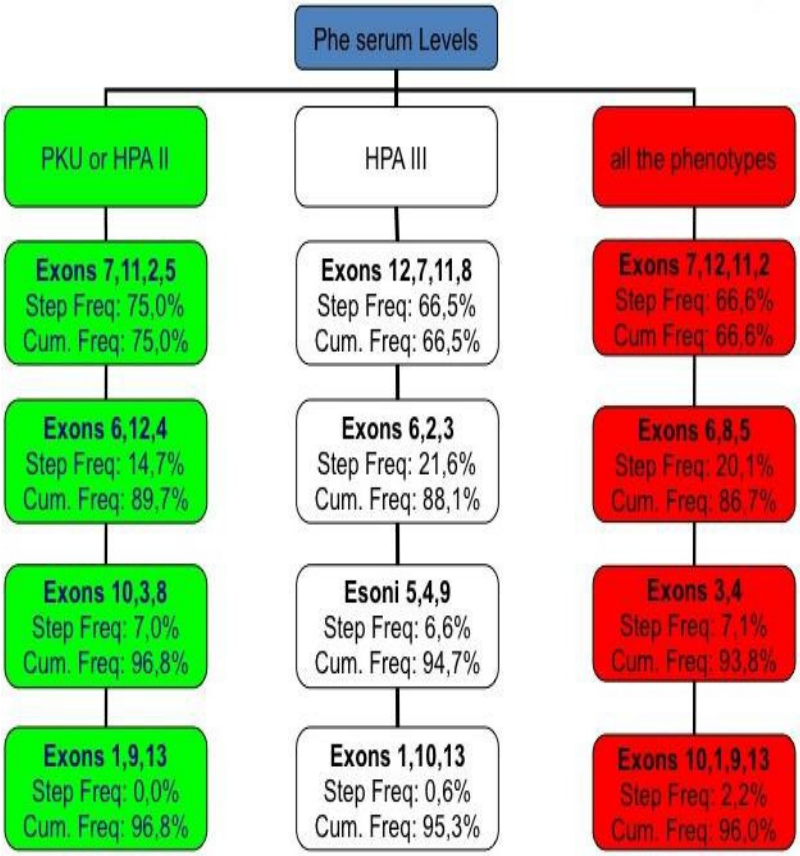

Fig. 16: Analytic flow chart for the molecular analysis of the PAH gene in Campania patients affected by PKU.

The mutations found in patients fall throughout the entire PAH gene without a real mutational hot spot, each causing different effects on the activity of PAH enzyme. In Table 6 are reported the mutations found in our population, each with the relative frequencies as well as the exon in which are present. As shown in Table 5 the exon 7 is the region in which the majority of mutations fall; it contains about the 23.8% of the total mutations, the percentage arises to 33.7% considering only the PKU patients. Moreover, the PAH gene corresponding to the exon 11 together with the exon 12 contains about 34.4 % of total mutations (the percentage arises to 39% considering only the HPAIII patients and 28% considering PKU). Exon 2 contains 11.5% of total mutations. The remaining exons (1, 3, 4, 5, 6, 8, 9 and 10) contain between 0.4 and 8.3% of total mutations. In exon 1 has been identified a single mutation and in exons 9 and 10, five and four respectively. No mutation has been identified in exon 13.

The frequencies of mutations per exons are represented in Fig. 17. Panels A, B, C, D show mutation frequency in PKU, HPAII, HPAIII and total patients respectively.

Exon	Protein Domain and Subdomain	PKU	HPAII	HPAIII	Total
1		0.4%	0.7%	0.2%	0.4%
2	Regulatory Domain nt 1-426 aa 1-142	14.9%	13.4%	8.9%	11.5%
3	Auto Regulatory Domain nt 55-99 aa 19-33	2.19%	1.6%	4.10%	3%
4	Catalytic Domain nt 427-1230 aa 143-410	3%	0.7%	2.5%	2.5%
5	Cofactor Binding Region #1 nt 733-798 aa 245-266	5.26%	7.9%	4.10%	5%
6	Cofactor Binding Region #2 nt 838-849 aa 280-283 Cofactor Binding Region #3 nt 964-978	7%	4.8%	10%	8.32%

	aa 322-326 Cofactor Binding Region #4 nt 1129-1137 aa 377-379				
7		33.7%	34.1%	14.8%	23.8%
8		0.4%	5%	11%	6.71%
9		0%	0.7%	1.3%	0.8%
10		4.8%	3%	2%	3%
11		23.6%	17.4%	15%	18%

12	Tetramerizazion Domain nt 1231-1356 aa 411-452	4.4%	10.3%	25.3%	16.3%
13		0%	0%	0%	0%

Table 6: Distribution of the mutations along the PAH

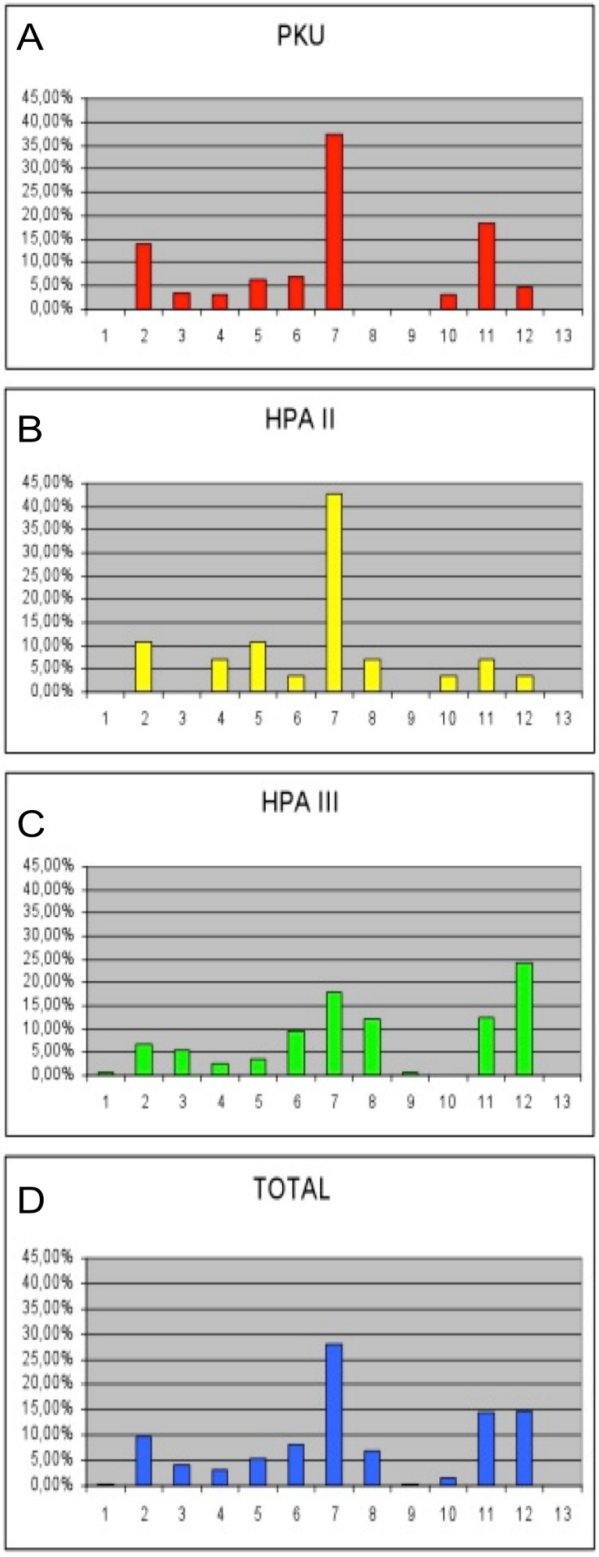

Fig. 17: Graphical representation of mutation frequencies per exon in total PKU patients. Panels A, B, C, D show mutation frequency in PKU, HPAII, HPAIII and total patients respectively.

4.4 Genotype-phenotype correlation

We studied correlation between genotype and phenotype. The genotype correlates well with phenotype in about 75% of total patients. These results are in accord with the 79% correlation rate previously reported (14). 59 patients (15%) have a homozygous genotype: 31 PKU phenotype corresponding to 52.5%; 9 HPAII phenotype corresponding to 15.3%; 19 HPAIII phenotype corresponding to 32.2% (Table 7); among them, we found that: the R261Q mutation was present in 16 patients, 8 PKU, 7 HPAII and 1 HPAIII according to the value of residual *in vitro* enzymatic activity; c.1066-11G>A was present in 11 PKU patients; and L48S present in 5 patients, in this case, the L48S homozygosity is associated with different phenotypic classes, i.e. HPAII and HPAIII. In addition, 5 HPAIII patients are homozygous for the mutation A403V according to the residual *in vitro* enzymatic activity; 4 PKU patients are homozygous for the mutation p.F55>Lfs and 4 HPAIII patients are homozygous for P366H. 3 PKU patients are each homozygous for the mutations p.G272X, p.P281L and p.R261X, respectively. From the analysis of the genotype-phenotype data, it results that p.R261Q and c.1066-11G>A mutations are mainly associated with severe PKU phenotype while p.A403V mutation is strongly associated to a mild phenotype while, in contrast, p.L48S is associated with HPA II and HPAIII phenotypes.

In addition, considering the PKU heterozygous patients, the most frequent mutation is the c.1066-11G>A (21%) followed by p.R261Q (17.3%), p.L48S (8.6%), p.281L (5.8%), p.F55>Lfs (5%). In HPAII patients, the most common mutations are p.R261Q (23%), c.1066-11G>A (13.2%), p.L48S (10.9%), p.281L (6.2%) and p.A403V (6.2 %); finally, in HPAIII patients, the most common mutations are p.A403V (20%), p.A300S (9.7%), p.R261Q (8.3%), p.L48S (5.6%) and c.1066-11G>A (5%). From the analysis of these data, we may conclude that c.1066-11G> A and p.R261Q mutations are more related to severe phenotype (PKU and HPAII); p.A403V p.A300S mutations are more related to mild HPAIII phenotype. Therefore, it can be concluded in 75% of analyzed patients, there is a good correlation genotype-phenotype.

Furthermore, there are some patients (about 5%) that, although presenting the same genotype, manifest different phenotypes; for example patients with genotype p.L48S/p.R261Q have a severe PKU or mild HPAIII phenotypes. Then, for few patients there is a discrepancy between genotype and phenotype. In conclusion, the metabolic phenotype of the analyzed patients is not

completely in accord to that expected, conclusion not feasibly justifiable for a monogenic disease such as PKU. Several reasons can be attributed to this discrepancy:

- ❖ Phenotypic classifications incorrect;
- ❖ Combination of different subunits encoded by mutant alleles in the homotetramer protein
- ❖ The action of modifier genes

Note that we must take into account that the classification on the basis of tolerance to L-Phe depends on both the maximum concentration of L-Phe that can be administered according to the age of the patient since it is known that the concentration of L-Phe varies with growth.

With regards to the combination of different monomers to form tetramers, there are 16 different combinations and the result of this interaction is not always predictable.

Finally the phenotypic variability that concerns patients with the same genotype may depend on differences between individuals regarding the presence of molecular chaperons aiming to stabilize the mutant protein.

Mutation	PKU	HPAII	HPAIII	Total
p.R261Q	8	7	1	16
c.1066-11G>A	11	0	0	11
p.L48S	0	2	3	5
p.A403V	0	0	5	5
p.F55>Lfs	4	0	0	4
p.P366H	0	0	4	4
p.D222G	0	0	2	2
c.1055delG	2	0	0	2
p.R252W	2	0	0	2
p.T380M	0	0	1	1
p.I94S	0	0	1	1
c.592_613del22	1	0	0	1
p.G352fs	1	0	0	1
p.G272X	1	0	0	1

p.P281L	1	0	0	1
p.R261X	1	0	0	1
p.V230I	0	0	1	1

Table 7: Classification of mutations in homozygous patients.

4.5 BH_4 Responsiveness

Our analysis showed that about 75% of the analyzed patients are BH4-responsive mutations. The responsiveness to BH4 can be tested both *in vivo* and *in vitro*. The BH4 loading *in vivo* test shows BH4-responsiveness with a decline of L-Phe blood by more than 30%; a positive test indicates that the patient may benefit from the administration of the cofactor in conjunction with a diet therapy. As regards the 18 new mutations, 5 mutations have been identified as *in vivo* BH4 responsive while the other will have to be tested in the future.

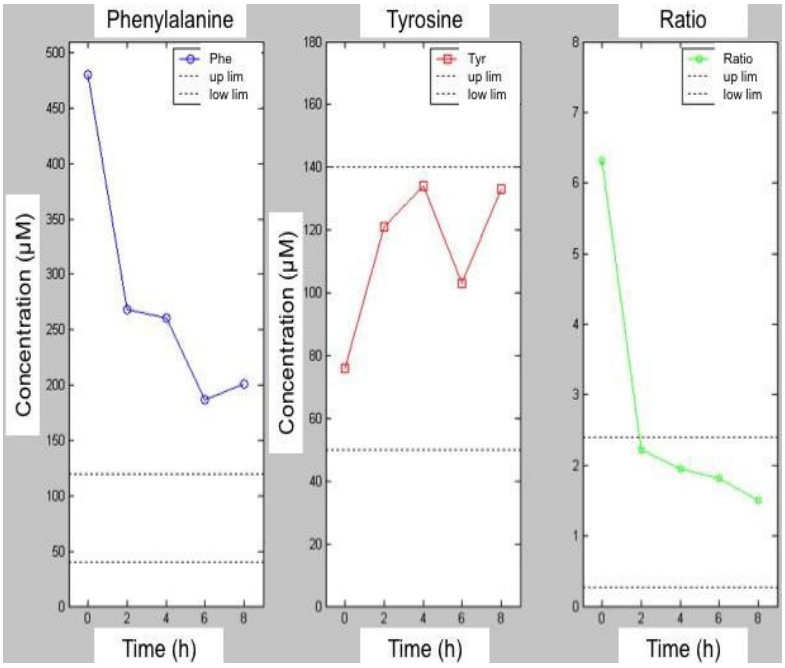

Fig. 18: Representation of the BH4 responsiveness in PKU patients bearing novel mutation.

4.6 Mutations of novel identification

In this thesis work, among the mutations identified in the Southern Italy population, 18 were novel, 14 are missense mutations, 3 are splicing mutations and one deletion. Moreover, all these mutations are located in regulatory domain, catalytic domain and tetramerization domain. The mutations are:

1. p.E76X
2. p.I95T
3. p.I65M
4. p.N223Y
5. p.Q235X
6. c.707-2delA
7. p.R297L
8. p.Q301P
9. p.L321I
10. p.W326S
11. p.F382L
12. c.1199nt4A>C
13. c.1315nt20C>T
14. p.K328N
15. p.N426H
16. p.I406M
17. p.Q419R
18. p.T418I

The missense mutations p.I95T, p.L321I, p.N223Y, p.K398N, p.F382L, p.Q419R, p.N426H, p.I1406M, p.T418I e p.R297L may be related to phenotype HPAIII. The mutation p.I65M may be associated with a severe PKU phenotype according to other variants of codon 65 previously described (p.I65N, p.I65T, p.I65V and p.I65S). The missense mutations p.Q235X and p.Q301P, that are located in the cofactor binding region 4, are associated to severe phenotypes. Only the missense mutation p.W326S is associated to the intermediate phenotype HPAII, while the p.E76X is associated to both PKU and HPAIII.

The c.707-2delA mutation involves the accepting splicing site of exon 7 causing the complete skipping of this exon. As a consequence of the skipping of the whole exon 7, a new open reading frame containing a frameshift generates a premature stop codon after 60 codons that leads to the formation of a truncated protein (p.T236MfsX60). This mutation, in our population, is associated to severe phetotype PKU. Finally the splicing mutations

c.119nt4A<C E c.1315nt20C<T are associated to HPAIII and HPAII mild phenotype, respectively.

4.7 Mutations identified by Next Generation Sequencing

17 patients in which, by using direct sequencing, we did not identified a mutation and their parents have been selected for NGS and analyzed as described under Methods. In particular, one sequencing run was performed to analyze all the study population obtaining more than 397 Mb. This was equivalent to a coverage of 100x for each subject, assessing the high sensitivity of our methodology. The high quality filtered reads were mapped against the human genome reference sequence to identify variants. NGS analysis identified mutation alterations in 9 patients as shown in Table 8. In detail, we found in 4 patients p.Glu178Gly mutation, in 1 patient the c.1066-11 G>A mutation and in 1 patient the g.50448_51402del955 deletion. Each mutation was confirmed in the patient and in the respective parent. In addition we identified also 3 variants in the promoter region confirmed in the patients' corresponding parent.

Region	SNPs and mutations identified by NGS	Phenotype
ex 6	p.Glu178Gly	HPA III
in 10	c.1066-11G>A omo	HPA I
in 4	g.50448_51402del955	HPA III
ex 6	p.Glu178Gly	HPA II
ex 6	p.Glu178Gly	HPA II
ex 6	p.Glu178Gly	HPA II
Promoter	c.-30T>C	HPA I
Promoter	c.-71A>C	HPA III
Promoter	c.-296T>C	HPA II

Table 8: list of additional PAH mutations identified through NGS and the corresponding patients phenotype.

Additional data analysis are still in progress for the identification of other possible causative variations also in the remaining patients. In particular, we are developing a specific bioinformatic pipeline for the analysis of gene deletions based on sequencing coverage in different PAH regions. Finally, we are also deeply investigating possibly interesting variations in the promoter to be selected for further functional evaluations.

5. Discussion and conclusions

The present thesis work enlarges the molecular epidemiology of PAH mutations, in particular, with respect to Southern Italy. By scanning of the PAH gene using direct sequencing, we detected a mutation in 97.8% of chromosomes from patients affected by PKU. Complete genotyping was carried out in 763/780 alleles. In seventeen patients only one causative mutation was found.

The failure to identify a mutation in one of the 2 alleles may be due to:

1) Alterations present in the promoter region or in the region of the 3' untranslated
2) Molecular defects that affect introns
3) Mutations of an enhancer
4) Large deletions or duplications of the PAH gene

Our study confirms the wide heterogeneity of PAH mutations in HPA patients also in our Campania region. Given the wide heterogeneity of PAH mutations, it is necessary to scan the whole coding region of the gene for molecular analysis. Indeed, the detection rate obtained with panels of exons/mutations is unsatisfactory; they cannot be used to plan cascade carrier screening, or for prenatal diagnosis, and with such a procedure some genotype-phenotype correlations may elude detection (40). The sequence analysis set-up in our study is sensitive and rapid, because all exons may be amplified under the same conditions and can be routinely used in reference diagnostic laboratories.

Our study, one of the largest carried out with a specific ethnic group, confirms that the epidemiology of PAH mutations in Southern Italy differs from that in other Italian and European ethnic groups. For example p.R408W, which is the most frequent mutation in most ethnic groups (Zschocke, 2003), has a frequency <2.0% in Southern Italy. Similarly, the most frequent mutations in Northern Italy (p.R252W and c.1315 + 1G>A) are rare in our ethnic group, whereas p.R261Q, c.1066-11G>A and p.L48S are the most frequent PKU mutations in Campania (25). The epidemiology of PAH mutations in Southern and Northern Italy reflects the pre-Roman colonisation of Italy (39): Celts colonized Northern Italy and most of central European, whereas the Greeks colonised Southern Italy, starting from the year 600 BC. There are also some differences between Southern Italian regions, i.e. Tirrenic (Campania) versus Adriatic (Puglia) regions, confirming results obtained for other genetic diseases (39). Finally, the most frequent mutations of our population coincide with those reported in Sicilian patients (25). Unfortunately, the Sicilian study does not distinguish the mutation

epidemiology within the three HPA phenotypes. In Croatia, Iceland, Wales and Taiwan a single mutation had a higher incidence (50; 10). The PAH mutations are differently distributed in our PKU groups of patients, confirming that several mutations are more frequently associated with the PKU phenotype and others (mainly missense). In particular, we confirm that p.R261Q, p.L48S and c.1066-11G>A are more frequently associated with the PKU and mild PKU phenotypes, whereas p.A403V and p.E390G are peculiar to the milder phenotypes. Furthermore, we show that c.842+3G>C, c.284-286delTCA, p.R158Q, p.E390G and p.V230I are peculiar to the mild HPA phenotype. Only five mutations had a frequency >3.0%, and the incidence of the most frequent mutation, i.e., p.R261Q, was 15.7%. Therefore, for few patients there is a discrepancy between genotype and phenotype. Several reasons can be attributed to this discrepancy: phenotypic classifications incorrect; combination of different subunits encoded by mutant alleles in the homotetramer protein; the action of modifier genes. Moreover, it must be taken into account that the classification on the basis of tolerance to L-Phe depends on both the maximum concentration of L-Phe that can be administered according to the age of the patient since it is known that the concentration of L-Phe varies with growth. Finally the phenotypic variability that concerns patients with the same genotype may depend on differences between individuals regarding the presence of molecular chaperons that could stabilize the mutant protein.

In this thesis work, we identified also 18 were novel, 14 are missense mutations, 3 are splicing mutations and 1 deletion. Moreover, all these mutations are located in regulatory domain, catalytic domain and tetramerization domain. 12 of these are missense mutations (p.I95T, p.L321I, p.N223Y, p.K398N, p.F382L, p.Q419R, p.N426H, p.I1406M, p.T418I e p.R297L) that could be considered “mild”, because they are found in patients bearing a “severe” mutation on the second allele (http://www.pahdb.mcgill.ca) and showing a mild biochemical phenotype. Mutation p.R297L is in cis with p.A403V, a “mild” mutation, and in trans with p.R243X, a “severe” mutation, but the patient shows a mild biochemical phenotype. Hence, p.R297L may also be considered a “mild” mutation. On the contrary, mutation p.I65M can be considered a “severe” mutation because the four previously described variants of codon 65 (p.I65T, p.I65N, p.I65V, p.I65S) are associated with a “severe” phenotype. Moreover, the patient bears a “mild” mutation on the second allele (i.e. p.A403V) and shows a mild phenotype.

The p.Q301P mutation was found in a compound heterozygous patient affected by an HPAII phenotype and bearing the p.L48S mutation on the other allele. The change leads to a protein with 4.4% residual enzyme activity and 8–10% residual expression, both tested in vitro. Two mechanisms appear to occur with this mutant protein: a lower stability that diminishes the protein level in the cell environment and a misfolding/destabilization of the tetrameric/dimeric structure, which impairs the catalytic function of the molecule. In this regard, it is noteworthy that Q301 is a phylogenetically highly conserved residue and that no mutation has been reported so far at this codon in the human PAH gene. Gln301 is located in the middle of an a-helix; hence, its replacement by Pro, an a-helix breaker residue, results in a drastic structural re-arrangement. Such a distortion might affect the structure and orientation of the Ca8 helix, which contains residues (i.e. R297 and Q304) anchoring a neighboring subunit, thereby stabilizing the dimer interface. The altered expression and function of the p.Q301P mutant protein may be attributed to destabilization of the monomer and/or to an altered oligomeric assembly. At the molecular level, the PAH tetramer may be formed from various combinations of mutated alleles. Homo- and heterotetramers can be formed at different ratios depending on the effects produced by mutations (i.e. folding defects, reduced stability or low levels of expression) (19). Being embodied in homo- or heterotetrameric proteins, the resulting enzyme may influence the overall in vivo activity (19). *In vivo*, the patient bearing mutation p.Q301P presents a HPAII phenotype and is BH4 responsive. This phenotype may be attributable either to the L48S allele or to the stabilizing effect of BH4 on the p.Q301P monomer. Mutation p.L48S was shown to produce a protein in vitro that underwent accelerated proteolytic action, as revealed by pulse-chase studies (49). Interestingly, the p.R158Q and p.P281L mutations increase the proportion of aggregates and produce less PAH tetramer (24), whereas the p.R261Q mutation produces a well known folding defect. Residue R261 plays a structural role (16) in that it contributes to the stabilization of the tertiary structure of the catalytic domain through a connection of different secondary structure elements.

We evaluated the biochemical phenotype in several patients homozygous for HPA mutations: p.P281L and c.1066-11G>A were associated with the PKU phenotype according to Bardelli's classification (1). p.R261X and p.G272X were associated with the PKU phenotype, and p.A403V and p.P366H caused mild HPA (the latter

mutation was defined “un- classified” by Bardelli et al. 2002). Of the five patients carrying p.R261Q in homozygosity, three were affected by PKU and two by mild PKU (although the L-Phe level was borderline between the clinical types). p.R261Q is known to be associated with different phenotypes within the same family (Kleiman et al. 1993). L-Phe levels differed widely in two patients carrying L48S in homozygosity, one being classified as “mild PKU” and the other as “mild HPA.” We speculate that other genes inherited independently of PAH may mitigate or exacerbate the biochemical genotype, as is the case for some Mendelian disorders (15; 41; 42). Furthermore, mutations in genes involved in the BH4 pathway may cause some cases of PKU and HPA (4), but we ruled out such mutations in most (7 out of 10) of our patients with unidentified PAH mutations (data not shown). Large deletions of the PAH gene in heterozygosity, undetectable by current scanning procedures, could account for unknown mutations (7; 51). Furthermore, in other genetic diseases, sequence variations that do not change amino acids within the protein sequence have been associated with impaired mRNA splicing and then as disease-causing mutations (37).

Moreover**,** about 75% of HPA patients from our population, mainly affected by mild PKU or HPA, had mutations previously classified as “BH4-sensitive” (44), and we identified five putative new BH4-sensitive mutations, which confirms that many patients may be treated with BH4 supplementation. BH4 sensitivity is usually diagnosed by the BH4 loading test, but identification of the BH4-sensitive mutation may decide the diagnosis in borderline cases and lead to a patient-rational treatment.

In addition, we performed prenatal diagnosis of three foetus in high-risk families; the molecular analysis of the PAH gene has been conducted in both parents and in chorionic villi or amniocytes taken at the sixteenth week of gestation. In one of the analyzed families, the mother of the foetus is affected by the HPA while the father carried no mutation; the foetus was a carrier of a single mutation and therefore not affected. In the other two families analyzed, instead, there is the presence of another affected child and both parents are carriers of a single mutation. One of the foetus was affected by HPA as carrier of both mutations in the heterozygous. The other foetus resulted not affected by the absence of both mutations of the parents.

In conclusion, our study, has allowed us to:

- Define correlations between some mutations and phenotypes of the disease.

- Define the responsiveness of BH4 in PKU patients. These data will allow patients to take a most appropriate therapy and personalized.
- Define the epidemiology of mutations in PAH gene of patients from Southern Italy, especially in Campania. These data may be useful for developing kits for molecular analysis of large-scale HPA patients and their families
- Calculate the residual risk to be carriers, and in cases where it is required, to carry out prenatal diagnosis in cases of family history of PKU.
- Define the differences in the epidemiology of these mutations, between different geographical areas.

As additional improvement of all the other molecular strategies successfully applied in the present thesis, we developed also a NGS strategy to fully analyze the PAH gene including the promoter, the introns and the 3' UTR, for a total target of about 200 kb.

This approach allowed the identification in one patient, and in his parent, of a 955 bp deletion causing the skipping of the exon 5. In other 4 patients we identified, as second causative mutation, the p.Glu178Gly mutation. Finally, other interesting variants in PAH regulatory regions are still under evaluation. In particular, 3 interesting variants in the promoter region have been identified and, given their possible pathogenetic role, selected for further functional evaluations.

Totally, by using this novel approach, we were able to additionally increase our diagnostic sensitivity ameliorating the mutation detection rate to a value of 98.8%.

These data may help to reconstruct the historical and anthropological point of view, the various settlements that have occurred in our country.

Bibliography

1. Bardelli, T., Donati, M. A., Gasperini, S., Ciani, F., Belli, F., Blau, N., Morrone, A. & Zammarchi, E. Two novel genetic lesions and a common BH4-responsive mutation of the PAH gene in Italian patients with hyperphenylalanine- mia. Mol Genet Metab. 2002;77, 260–266.
2. Bélanger-Quintana A, Burlina A, Harding CO, Muntau AC. Up to date knowledge on different treatment strategies for phenylketonuria. Mol Genet Metab. 2011;104 Suppl:S19-25.
3. Bernegger C, Blau N. High frequency of tetrahydrobiopterin-responsiveness among hyperphenylalaninemias: a study of 1,919 patients observed from 1988 to 2002. Mol Genet Metab. 2002;77(4):304-13.
4. Blau N, Erlandsen H. The metabolic and molecular bases of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. Mol Genet Metab. 2004;82(2):101-11.
5. Blau N, Van Spronsen F, Levy H. Phenylketonuria. Lancet. 2010; 376: 1417
6. Blau N, Hennermann JB, Langenbeck U, Lichter-Konecki U. Diagnosis, classification, and genetics of phenylketonuria and tetrahydrobiopterin (BH4) deficiencies. Mol Genet Metab. 2011;104 Suppl:S2-9.
7. Bosco, P., Ceratto, N., Cali, F., Goltsov, A. A., Eisensmith, R. C., Novelli, G., Dalla Piccola, B. & Romano, V. RFLP discordance in a PKU family due to a deletion in the PAH gene. Turk J Pediatr. 1996;38, 497–504.
8. Cerreto M, Mehdawy B, Ombrone D, Nisticò R, Ruoppolo M, Usiello A, Daniele A, Pastore L, Salvatore F. Reversal of metabolic and neurological symptoms of phenylketonuric mice treated with a PAH containing helper-dependent adenoviral vector. Curr Gene Ther. 2012;12(1):48-56.
9. Cerreto M, Cavaliere P, Carluccio C, Amato F, Zagari A, Daniele A, Salvatore F. Natural phenylalanine hydroxylase variants that confer a mild phenotype affect the enzyme's conformational stability and oligomerization equilibrium. Biochim Biophys Acta. 2011 Nov;1812(11):1435-45.
10. Chien, Y. H., Chiang, S. C., Huang, A., Chou, S. P., Tseng, S. S., Huang, Y. T. & Hwu, W. L. Mutation spec-trum in Taiwanese patients with phenylalanine hydroxylase deficiency and a founder effect for the R241C mutation. Hum Mutat. 2004;23, 206.
11. Christensen R, Güttler F, Jensen TG. Comparison of epidermal keratinocytes and dermal fibroblasts as potential target cells for somatic gene therapy of phenylketonuria. Mol Genet Metab. 2002;76(4):313-8.

,

12. Daniele A, Scala I, Cardillo G, Pennino C, Ungaro C, Sibilio M, Parenti G, Esposito L, Zagari A, Andria G, Salvatore F. Functional and structural characterization of novel mutations and genotype-phenotype correlation in 51 phenylalanine hydroxylase deficient families from Southern Italy. FEBS J. 2009;276(7):2048-59.
13. Daniele A, Cardillo G, Pennino C, Carbone MT, Scognamiglio D, Esposito L, Correra A, Castaldo G, Zagari A, Salvatore F. Five human phenylalanine hydroxylase proteins identified in mild hyperphenylalaninemia patients are disease-causing variants. Biochim Biophys Acta. 2008;1782(6):378-84.
14. Daniele A, Cardillo G, Pennino C, Carbone MT, Scognamiglio D, Correra A, Pignero A, Castaldo G, Salvatore F. Molecular epidemiology of phenylalanine hydroxylase deficiency in Southern Italy: a 96% detection rate with ten novel mutations. Ann Hum Genet. 2007;71(Pt 2):185-93.
15. Dipple, K. M. & McCabe, E. R. B. Modifier genes convert “simple” mendelian disorders to complex traits. Mol Genet Metab. 2000;71, 43–50.
16. Erlandsen H & Stevens RC. The structural basis of phenylketonuria. Mol Genet Metab. 1999;68, 103–125.
17. Erlandsen H, Stevens RC. A structural hypothesis for BH4 responsiveness in patients with mild forms of hyperphenylalaninaemia and phenylketonuria. J Inherit Metab Dis. 2001;24(2):213-30.
18. Erlandsen H, Patch MG, Gamez A, Straub M, Stevens RC. Structural studies on phenylalanine hydroxylase and implications toward understanding and treating phenylketonuria. Pediatrics. 2003;112(6 Pt 2):1557-65.
19. Fincham JRS & Pateman JA. Formation of an enzyme through complementary action of mutant ‘alleles’ in separate nuclei in a heterocaryon. Nature. 1957;179, 741–742.
20. Fitzpatrick PF. The aromatic amino acid hydroxylases. Adv Enzymol Relat Areas Mol Biol. 2000;74:235-94.
20b. Feliubadaló L, Lopez-Doriga A, Castellsagué E, Del Valle J, Menéndez M, Tornero E, Montes E, Cuesta R, Gómez C, Campos O, Pineda M, González S, Moreno V, Brunet J, Blanco I, Serra E, Capellá G, Lázaro C. Next-generation sequencing meets genetic diagnostics: development of a comprehensive workflow for the analysis of BRCA1 and BRCA2 genes. Eur J Hum Genet. 2012 Dec 19.
21. Gersting SW, Lagler FB, Eichinger A, Kemter KF, Danecka MK, Messing DD, Staudigl M, Domdey KA, Zsifkovits C, Fingerhut R, Glossmann H, Roscher AA, Muntau AC. Pahenu1 is a mouse model for tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency and promotes analysis of the pharmacological

chaperone mechanism in vivo. Hum Mol Genet. 2010 May 15;19(10):2039-49.
22. Gersting SW, Staudigl M, Truger MS, Messing DD, Danecka MK, Sommerhoff CP, Kemter KF, Muntau AC. Activation of phenylalanine hydroxylase induces positive cooperativity toward the natural cofactor. J Biol Chem. 2010;285(40):30686-97.
23. Gibbs BS, Wojchowski D, Benkovic SJ. Expression of rat liver phenylalanine hydroxylase in insect cells and site-directed mutagenesis of putative non-heme iron-binding sites. J Biol Chem. 1993;268(11):8046-52.
24. Gjetting T, Petersen M, Guldberg P & Guttler F. In vitro expression of 34 naturally occurring mutant variants of phenylalanine hydroxylase: correlation with metabolic phenotypes and susceptibility toward protein aggregation. Mol Genet Metab. 2001;72, 132–143.
25. Giannattasio, S., Dianzani, I., Lattanzio, P., Spada, M., Ro- mano, V., Cali, F., Andria, G., Ponzone, A., Marra, E. & Piazza, A. Genetic heterogeneity in five Italian re- gions: Analysis of PAH mutations and minihaplotypes. Hum Hered. 2001;52, 154–159.
25b. Goossens D et al. (Simultaneous mutation and copy number variation (CNV) detection by multiplex PCR-based GS-FLX sequencing. Hum Mutat. 2009;30, 472-6.
26. Guldberg P, Rey F, Zschocke J, Romano V, François B, Michiels L, Ullrich K, Hoffmann GF, Burgard P, Schmidt H, Meli C, Riva E, Dianzani I, Ponzone A, Rey J, Güttler F. A European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. Am J Hum Genet. 1998;63(1):71-9.
27. Hennermann JB, Bührer C, Blau N, Vetter B, Mönch E. Long-term treatment with tetrahydrobiopterin increases phenylalanine tolerance in children with severe phenotype of phenylketonuria. Mol Genet Metab. 2005;86 Suppl 1:S86-90.
28. Hufton SE, Jennings IG, Cotton RG. Structure and function of the aromatic amino acid hydroxylases. Biochem J. 1995;311 (Pt 2):353-66.
29. Jennings IG, Kemp BE, Cotton RG. Localization of cofactor binding sites with monoclonal anti-idiotype antibodies: phenylalanine hydroxylase. Proc Natl Acad Sci U S A. 1991 Jul 1;88(13):5734-8.
30. Kanai Y, Endou H. Functional properties of multispecific amino acid transporters and their implications to transporter-mediated toxicity. J Toxicol Sci. 2003;28(1):1-17.
31. Kleiman, S., Vanagaite, L., Bernstein, J., Schwartz, G., Brand, N., Elitzur, A., Woo, S. L. & Shiloh, Y. Phenylke- tonuria: variable phenotypic outcomes of the R261Q mu-

tation and maternal PKU in the offspring of a healthy homozygote. J Med Genet. 1993;30, 284–288
32. Lindner M, Haas D, Mayatepek E, Zschocke J, Burgard P. Tetrahydrobiopterin responsiveness in phenylketonuria differs between patients with the same genotype. Mol Genet Metab. 2001;73(1):104-6.
33. MacDonald A, Rocha JC, van Rijn M, Feillet F. Nutrition in phenylketonuria. Mol Genet Metab. 2011;104 Suppl:S10-8.
33b. Margulies M, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;437, 376-80.
34. Martínez A, Andersson KK, Haavik J, Flatmark T. EPR and 1H-NMR spectroscopic studies on the paramagnetic iron at the active site of phenylalanine hydroxylase and its interaction with substrates and inhibitors. Eur J Biochem. 1991;198(3):675-82.
35. Matalon R, Surendran S, Matalon KM, Tyring S, Quast M, Jinga W, Ezell E, Szucs S. Future role of large neutral amino acids in transport of phenylalanine into the brain. Pediatrics. 2003;112(6 Pt 2):1570-4.
36. Michals-Matalon K, Bhatia G, Guttler F, Tyring SK, Matalon R. Response of phenylketonuria to tetrahydrobiopterin. J Nutr. 2007;137(6 Suppl1):1564S-1567S; discussion 1573S-1575S.
37. Pagani, F., Stuani, C., Tzetis, M., Kanavakis, E., Efthymiadou, A., Doudounakis, S., Casals, T. & Baralle, F. E. New type of disease causing mutations: the example of the com- posite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet 2003;12, 1111–1120.
38. Pérez-Dueñas B, Vilaseca MA, Mas A, Lambruschini N, Artuch R, Gómez L, Pineda J, Gutiérrez A, Mila M, Campistol J. Tetrahydrobiopterin responsiveness in patients with phenylketonuria. Clin Biochem. 2004;37(12):1083-90.
39. Rendine, S., Calafell, F., Cappello, N., Gagliardini, R., Caramia, G., Rigillo, N., Silvetti, M., Zanda, M., Miano, A., Battistini, F., Marianelli, L., Taccetti, G., Diana, M. C., Romano, L., Romano, C., Giunta, A., Padoan, R., Pianaroli, A., Raia, V., De Ritis, G., Battistini, A., Grzincich, G., Japichino, L., Pardo, F., Piazza, A. et al. Genetic history of cystic fibrosis mutations in Italy. I. Regional distribution. Ann Hum Genet. 1997;61, 411–424.
40. Romano, V., Lio, D., Cali, F., Scola, L., Leggio, L., D'Anna, C., De Leo, G. & Salermo, A. A methodologi- cal strategy for PAH genotyping in populations with a marked molecular heterogeneity of hyperphenylalaninemia. Mol Cell Probes. 2001;15, 13–19.

41. Salvatore, F., Scudiero, O., Castaldo, G. Genotype-phenotype correlation in cystic fibrosis: the role of modifier genes. Am J Med Genet. 2002;111, 88–95.
42. Scriver, C. R. & Waters, P. J. Monogenic traits are not simple: Lessons from phenylketonuria. Trends Genet. 1999;15, 267–272.
43. Scriver CR, Hurtubise M, Konecki D, Phommarinh M, Prevost L, Erlandsen H, Stevens R, Waters PJ, Ryan S, McDonald D, Sarkissian C. PAHdb 2003: what a locus-specific knowledgebase can do. Hum Mutat. 2003;21(4):333-44.
44. Spaapen LJ, Rubio-Gozalbo ME. Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency, state of the art. Mol Genet Metab. 2003;78(2):93-9.
45. Steinfeld R, Kohlschütter A, Ullrich K, Lukacs Z. A hypothesis on the biochemical mechanism of BH(4)-responsiveness in phenylalanine hydroxylase deficiency. Amino Acids. 2003;25(1):63-8.
46. Thöny B, Ding Z, Martínez A. Tetrahydrobiopterin protects phenylalanine hydroxylase activity in vivo: implications for tetrahydrobiopterin-responsive hyperphenylalaninemia. FEBS Lett. 2004;577(3):507-11.
47. Trefz FK, Blau N. Potential role of tetrahydrobiopterin in the treatment of maternal phenylketonuria. Pediatrics. 2003;112(6 Pt 2):1566-9.
48. Van Spronsen FJ, Huijbregts SC, Bosch AM, Leuzzi V. Cognitive, neurophysiological, neurological and psychosocial outcomes in early-treated PKU-patients: a start toward standardized outcome measurement across development. Mol Genet Metab. 2011;104 Suppl:S45-51.
49. Waters PJ, Parniak MA, Akerman BR, Jones AO & Scriver CR. Missense mutations in the phenylala- nine hydroxylase gene (PAH) can cause accelerated proteolytic turnover of PAH enzyme: a mechanism underlying phenylketonuria. J Inher Metab Dis. 1999;22, 208–212.
50. Zschocke, J. Phenylketonuria mutations in Europe. Hum Mutat. 2003;21, 345–356.
51. Zschocke, J. and Hoffmann, G. F. Phenylketonuria mu-tations in Germany. Hum Genet. 1999;104, 390–398.

ELENCO PUBBLICAZIONI

1. Long-range PCR and next generation sequencing for the identification of PAH mutation status in hyperphenylalaninemia Italian patients. V. D'Argenio, G. Guerri, A. Telese, A. Palmieri, A. Daniele, F. Salvatore. 20° IFCC-EFLM European Congress of Clinical Chemistry and Laboratory Medicine and 45° Congress of the Italian Society of Clinical Biochemistry and Clinical Molecular Biology (SIBioC), Milano, 19-23 maggio 2013.

2. Analisi molecolare del gene PAH in pazienti italiani affetti da Fenilchetonuria mediante sequenziamento high throughput del DNA. Palmieri A., D Argenio V., Guerri G., Cacciapuoti G., Di Lorenzo L., Daniele A., Salvatore F. Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2012, 10-11 luglio 2012.

3. Molecular analysis of PAH gene in Italian patients affected by Phenylketonuria by DNA high throughput sequencing. A. Palmieri, V. D'Argenio, G. Guerri, A. De Rosa, C. Canero, A. Daniele, F. Salvatore. 48° Congresso dell'Associazione Italiana di Neuropatologia e Neurobiologia Clinica AINPeNC e 38° Congresso dell'Associazione Italiana di Ricerca sull'Invecchiamento Cerebrale AIRIC, Napoli, 24-26 maggio 2012.

4. Iperfenilalaninemie: identificazione delle mutazioni attraverso HIGHTHROUGHPUT SEQUENCING. M.T. Carbone, A. Palmieri, V. D'Argenio, G. Guerri, A. Telese, A. Correra, A. Daniele, F. Salvatore. Congresso Nazionale Congiunto SIMMESN e SIMGePeD Malattie genetico-metaboliche tra tecnologia e assistenza, Bologna, 27-29 ottobre 2011. Abstract pag.45.

5. Un caso di Fenilchetonuria Materna: analisi molecolare del gene large neutral amino acid transporter 1. Palmieri A., Cerreto M., Nigro E., Carbone M.T., Correra A., Daniele A., Salvatore F. Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2011, 28 giugno 2011.

6. Rapid and sensitive assessment of PAH mutation status in Hyperphenylalaninemia Italian patients based on DNA high throughput sequencing. A. Palmieri, V. D'Argenio, G. Guerri, V. Sanna, A. Daniele, F. Salvatore. 42° Congresso Nazionale SIBioC, Roma, ottobre 2010. Abstract pag. 428.

7. Determinazione delle mutazioni nel gene PAH in pazienti affetti da HPA attraverso HIGH THROUGHPUT sequencing. Palmieri A., D Argenio V., Guerri G., Daniele A., Salvatore F. Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2009, 7-10 luglio 2010.

20° IFCC-EFLM European Congress of Clinical Chemistry and Laboratory Medicine and 45° Congress of the Italian Society of Clinical Biochemistry and Clinical Molecular Biology (SIBioC), Milano, 19-23 maggio 2013

Long-range PCR and next generation sequencing for the identification of PAH mutation status in hyperphenylalaninemia Italian patients

V. D'Argenio[1,2], G. Guerri[1,2], A. Telese[1,2], A. Palmieri[1], A. Daniele[1,3], F. Salvatore[1,2,4]

[1]CEINGE-Biotecnologie Avanzate Scarl, Napoli; [2]Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Napoli; [3]Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli; [4]IRCCS-Fondazione SDN, Napoli

BACKGROUND. Phenylketonuria (PKU) is the most important disorder of amino acid metabolism, resulting mainly from a deficiency of phenylalanine hydroxylase (PAH) that metabolises phenylalanine to tyrosine. Usually, defective PAH activity is caused by mutations in the PAH gene; more than 500 mutations, scattered over the entire gene length, having been reported to date. Mutation analysis is important to obtain information on expected phenotype, for family counseling, to identify PKU carriers and for prenatal diagnosis. The scanning of PAH gene by direct sequencing detected a mutation in 90% of patients. Therefore, this strategy failed to reveal a point mutation or deletion/insertion on several alleles.

In this study, we developed and tested a method for the comprehensive and sensitive detection of PAH mutations using long-range (LR) PCR and next generation sequencing.

METHODS. The study was carried out on 19 PKU patients for which the conventional Sanger exons screening was able to detect only one causative mutation. DNA samples of patient's parents were also available for molecular testing and results confirmation. Sixteen LR-PCR fragments, between 3,400 and 11,700 bp, containing the promoter, all coding exons, all introns and the 3' UTR of PAH were individually obtained for each study subject. The purified amplicons of the same patient were then pooled in equimolar ratio to obtain a library/sample. The libraries were sequenced with the GS FLX System. Sequence and data analysis was performed using the Roche/454 gsMapper software.

RESULTS. All LR-PCR fragments were completely sequenced. We detected and confirmed 100% of the mutations and SNP previously identified in the same samples by Sanger analysis. In addition, we identified some interesting possibly pathogenetic variations that segregated with the family inheritance (40% of the studied subjects).

CONCLUSION. Our results demonstrates that genomic LR-PCR and next generation sequencing, increasing diagnostic sensitivity, can provide comprehensive genetic information more quickly and accurately than conventional approaches, being an effective method for patient sample analysis of PAH gene.

Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2012, 10-11 luglio 2012.

Analisi molecolare del gene PAH in pazienti italiani affetti da Fenilchetonuria mediante sequenziamento high throughput del DNA

Palmieri A.[(1)], D'Argenio V.[(1)], Guerri G.[(1)], Cacciapuoti G.[(1)], Di Lorenzo L.[(1)], Daniele A.[(2)], Salvatore F.[(1)]

[(1)]CEINGE- Biotecnologie Avanzate Scarl, Napoli [(2)]CEINGE- Biotecnologie Avanzate Scarl, Napoli; Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Facoltà di Scienze del Farmaco per l'Ambiente e la Salute

La Fenilchetonuria (PKU) e le sue varianti, le Iperfenilalaninemie (HPA) derivano dalla deficienza della fenilalanina idrossilasi (PAH: EC 1.14.16.1), enzima epatico che utilizza la tetraidrobiopterina (BH4) per metabolizzare la Fenilalanina (Phe) in Tirosina (Tyr). Gli alti livelli di Phe e dei suoi metaboliti sono neurotossici e determinano gravi danni cerebrali. Recentemente è stato osservato che un trattamento aggiuntivo con BH4 rappresenta in alcuni pazienti un valido trattamento. Il gene PAH mappa sul cromosoma 12q23.2 codifica per una proteina di 50kDa costituita da un dominio regolatorio, uno catalitico e uno di tetramerizzazione. Nel 97% delle forme di HPA, l'attività PAH difettiva è causata da mutazioni nel gene PAH; sono state identificate più di 500 mutazioni la maggior parte delle quali distribuite sull'intera lunghezza del gene (http://www.pahdb.mcgill.ca). L'identificazione delle mutazioni è utile per ottenere informazioni sia sul fenotipo del paziente sia sulla proteina e per poter svolgere una più completa consulenza genetica nelle famiglie a rischio; inoltre la diagnosi molecolare permette l'identificazione dei portatori sani di HPA e l'esecuzione di diagnosi prenatali. L'analisi molecolare permette inoltre una migliore classificazione dei pazienti responsivi al BH4 permettendo un approccio terapeutico più personalizzato e quindi più appropriato. La procedura attualmente utilizzata si basa sul sequenziamento diretto, mediante metodo Sanger del promotore e dei 13 esoni del gene PAH. Con questa strategia di analisi abbiamo identificato mutazioni in più 90% dei pazienti per cui sarebbe auspicabile mettere a punto una procedura di analisi più sensibile che permetta di identificare anche delezioni che coprono uno o più esoni e che risultano difficilmente riconoscibili negli eterozigoti compositi a causa dell'effetto mascherante dell'allele non deleto. L'obiettivo di questo studio è mettere a punto una nuova e più sensibile strategia di analisi delle mutazioni mediante approccio di high-throughput sequencing GS FLX. Materiali e metodi: Abbiamo raccolto campioni di sangue (3 ml) in EDTA ed estratto il DNA di soggetti sani e di pazienti affetti da PKU identificati dallo screening neonatale. Utilizzando il software Primer3, abbiamo disegnato primers specifici per amplificare e analizzare regioni di 100 kb che comprendessero 20 Kb nella regione del promotore, tutti gli esoni ed introni, 5 kb della regione 3 'UTR del gene PAH. Risultati: Utilizzando un protocollo di PCR touch-down, abbiamo generato lunghi ampliconi parzialmente sovrapposti, le cui dimensioni variano tra i 8-12 Kb. Ciascun amplicone è stato poi amplificato e purificato. Dopo valutazione della qualità (2100 Bioanalyzer, Agilent) e quantificazione (Picogreen test, Invitrogen), i prodotti di amplificazione da uno stesso campione di DNA sono stati sequenziati. Conclusioni: l'approccio di amplificazione tramite PCR e sequenziamento highthroughput è in grado di rilevare sia alterazioni di singoli nucleotidi sia grosse delezioni o duplicazioni nei geni target aumentando così lo spettro delle alterazioni genetiche rilevabili e di conseguenza la detection rate.

48° Congresso dell'Associazione Italiana di Neuropatologia e Neurobiologia Clinica AINPeNC e 38° Congresso dell'Associazione Italiana di Ricerca sull'Invecchiamento Cerebrale AIRIC, Napoli, 24-26 maggio 2012.

Molecular analysis of PAH gene in Italian patients affected by Phenylketonuria by DNA high throughput sequencing

A. Palmieri[1], V. D'Argenio[1,2], G. Guerri[1], A. De Rosa[1], C. Canero[1], A. Daniele[1,3], F. Salvatore[1,2,4]

[1]CEINGE-Biotecnologie Avanzate Scarl, Napoli, Italia; [2]Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Napoli, Italia; [3]Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Italia; [4]IRCCS-Fondazione SDN, Napoli, Italia.

Introduction: Phenylketonuria (OMIM 261600) is the most common disorder of amino acid metabolism caused by phenylalanine hydroxylase (PAH: EC 1.14.16.1) deficiency. Accumulation of Phe causes severe brain damage with mental retardation but an early Phe-restricted diet prevents neurocognitive and developmental damage. In PKU, defective PAH activity results from mutations in the PAH gene. To date, more than 600 mutations have been identified (http://www.pahdb.mcgill.ca). Mutation analysis is important to obtain information both about the expected phenotype and for genetic counseling. In our lab, we perform the molecular diagnosis of HPA with a detection rate of ~90% (Daniele, 2008; Daniele 2011). Our aim is to set-up a new and more sensitive strategy using high-throughput sequencing approach.

Materials and Methods: From controls and PKU patients identified from neonatal screening, we collected a blood sample (3 mL) in EDTA and extracted DNA using a standard protocol. Using the Primer3 software (http://frodo.wi.mit.edu/primer3/), we designed specific primers encompassing about 100 kb including 20 kb of the promoter region, all exons and introns, and 5 kb of the 3' UTR region of PAH gene.

Results: Overlapping long-PCR amplicons, ranging in size between 8-12 kb, were generated using a "touch-down" PCR protocol. Each amplicon was individually amplified and purified. After appropriate quality assessment (2100 BioAnalyzer, Agilent) and quantification (Picogreen assay, Invitrogen), the amplification products from the same DNA sample are sequenced.

Conclusions: The PCR amplification/high throughput sequencing approach detects both single-nucleotide polymorphisms and CNVs in target genes thereby increasing the spectrum of detected variations within the analyzed gene.

Congresso Nazionale Congiunto SIMMESN e SIMGePeD Malattie genetico-metaboliche tra tecnologia e assistenza, Bologna, 27-29 ottobre 2011. Abs. pag. 45

IPERFENILALANINEMIE: IDENTIFICAZIONE DELLE MUTAZIONI ATTRAVERSO HIGH-THROUGHPUT SEQUENCING

Maria Teresa Carbone [(1)], Alessia Palmieri [(2)], Valeria D'Argenio [(2)], Giulia Guerri [(2)], Antonella Telese [(2)], A. Correra [(1)], Aurora Daniele [(2,3)], Francesco Salvatore [(2,4)].

[(1)] Centro Screening Fenilchetonuria, Ospedale SS. Annunziata, ASL Na1, Napoli; [(2)] CEINGE- Biotecnologie Avanzate Scarl, Napoli; [(3)] Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli; [(4)] IRCCS-Fondazione SDN, Napoli

Introduzione: Le Iperfenilalaninemie derivano dalla deficienza della fenilalanina idrossilasi (PAH) che converte la Fenilalanina in Tirosina. Alti livelli di fenilalanina sono neurotossici e determinano gravi danni cerebrali. Il gene PAH codifica una proteina di 50kDa costituita da un dominio regolatorio, uno catalitico, uno di tetramerizzazione. Nel 97% dei pazienti, l'attività PAH difettiva è causata da mutazioni nel gene PAH (http://www.pahdb.mcgill.ca). L'identificazione delle mutazioni è utile per ottenere informazioni sia sul fenotipo del paziente, sia sulla proteina, e per una più completa consulenza genetica; inoltre, la diagnosi molecolare permette l'identificazione dei portatori sani e la diagnosi prenatale. Scopo: mettere a punto una procedura di analisi più sensibile che permetta di identificare mutazioni nel 99% dei pazienti.
Metodi: Il gene PAH è stato amplificato mediante Long-PCR. Ciascun amplicone è stato purificato ed analizzato per qualità (2100 BioAnalyzer, Agilent) e quantità (Picogreen assay, Invitrogen). Gli ampliconi dello stesso paziente sono stati assemblati in rapporto equimolare ed utilizzati per generare librerie da sequenziare mediante sequenziamento high throughput (GS FLX System).
Risultati: Specifici primer sono stati disegnati con il software Primer3 (http://frodo.wi.mit.edu/primer3/) e le condizioni di amplificazione sono state ottimizzate. Il DNA di tutti i pazienti selezionati è stato amplificato in modo da ottenere un pool di ampliconi per ciascun paziente. E' in corso la preparazione delle librerie per il sequenziamento.
Conclusioni: questo nuovo approccio permette l'identificazione delle mutazioni nel gene PAH, compreso il promotore, le regioni codificanti e non codificanti. In questo modo, saranno identificati variazioni nucleotidiche, delezioni/inserzioni e polimorfismi ad oggi non descritti. Questo approccio permetterà di analizzare simultaneamente diversi pazienti riducendo in maniera significativa i tempi dell'analisi.

Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2011, 28 giugno 2011

Un caso di Fenilchetonuria Materna: analisi molecolare del gene large neutral amino acid transporter 1

Palmieri A. [(1)], Cerreto M. [(1)], Nigro E. [(1)], Carbone M.T. [(2)], Correra A. [(2)], Daniele A. [(3)], Salvatore F. [(4)]

[(1)]CEINGE- Biotecnologie Avanzate Scarl, Napoli [(2)]Centro Screening Fenilchetonuria-Ospedale SS. Annunziata-ASL Na1 Napoli Italia [(3)]CEINGE-Biotecnologie Avanzate Scarl, Napoli;Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Facoltà di Scienze del Farmaco per l'Ambiente e la Salute [(4)]CEINGE-Biotecnologie Avanzate S.C.aR.L., Napoli - Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli

La Fenilchetonuria (PKU) è una patologia autosomica recessiva legata nel 98% dei casi a mutazioni del gene Fenilalanina idrossilasi (PAH). Il gene PAH codifica per un enzima ad espressione epatica che catalizza la trasformazione della Fenilalanina (Phe) in Tirosina in presenza di O2 e Tetraidrobiopterina (BH4). La ridotta attività dell enzima comporta un incremento delle concentrazioni plasmatiche di Phe che determinano gravi danni cerebrali se la PKU non è tempestivamente diagnosticata e trattata. Mutazioni nel gene PAH causano mancata o ridotta attività dell'enzima; attraverso l'analisi molecolare del gene PAH si effettua la diagnosi genetica di PKU. Elevati livelli plasmatici di Phe durante la gravidanza esercitano effetti teratogeni sul feto e creano una condizione di Iperfenilalaninemia embrio-fetale, lesiva della morfogenesi, nonch dello sviluppo fisico e cerebrale del prodotto del concepimento che va sotto il nome di sindrome da PKU materna (MSPKU). Il 95% dei neonati affetti da MSPKU presentano ritardo mentale, il 73% microcefalia, il 40% ritardo di crescita intrauterino e il 12% difetti congeniti cardiaci. Sono state ipotizzate correlazioni tra il genotipo PKU, il fenotipo biochimico e quello clinico. Nella patogenesi della MSPKU è stata suggerita un'implicazione della barriera emato-encefalica, in particolare del sistema LNAA (large neutral amino acid transporter 1, LAT1), responsabile del trasporto della Phe. Alterazioni nucleotidiche nel gene codificante LAT1, quindi, potrebbero giocare un ruolo nel determinare una diversa suscettibilità ai danni cerebrali in pazienti affetti da MSPKU. In questo studio riportiamo un caso di MSPKU; si tratta di un neonato partorito da una paziente affetta dalla forma grave di PKU; essa infatti presenta livelli di Phe superiori ai 1200 umol/l. La paziente PKU è arrivata al Centro di Screening per la PKU dell Ospedale SS. Annunziata di Napoli in stato di gravidanza. La paziente, nonostante il counselling genetico, non ha effettuato dietoterapia e il neonato è risultato affetto da MSPKU classica (grave microcefalia, difetti cardiaci). Abbiamo effettuato in questa paziente l'analisi molecolare del gene PAH evidenziando la presenza di due mutazioni riportate nel database delle mutazioni causative di PKU: c.1066-11G A e p.L48S. La prima mutazione si trova nel sito attivo dell enzima, in particolare nel dominio di legame al suo cofattore BH4 e si associa sempre a fenotipo grave. La seconda mutazione, che cade nel dominio di autoregolazione della proteina, è associata a tutti i fenotipi di Fenilchetonuria. Abbiamo effettuato inoltre l'analisi molecolare del gene PAH nel neonato PKU evidenziando la presenza della mutazione p.L48S in eterozigosi. Inoltre abbiamo messo a punto la procedura per effettuare il sequenziamento del gene LAT1.

42° Congresso Nazionale SIBioC, Roma, 2010. Abs pag. 428.

Rapid and sensitive assessment of PAH mutation status in Hyperphenylalaninemia Italian patients based on DNA high throughput sequencing

A. Palmieri[1], V. D'Argenio[1,2], G. Guerri[1], V. Sanna[1], A. Daniele[1,3], F. Salvatore[1,2,4]

[1]CEINGE-Biotecnologie Avanzate Scarl, Napoli; [2]Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Napoli; [3]Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli; [4]IRCCS-Fondazione SDN, Napoli

Phenylketonuria (OMIM database: 261600) and its Hyperphenylalaninemia (HPA) variants are the most important disorders of amino acid metabolism caused by a deficiency of phenylalanine hydroxylase (PAH: EC 1.14.16.1), which metabolizes phenylalanine (Phe) to tyrosine. Phe causes severe brain damage but an early Phe-restricted diet prevents neurocognitive and developmental damage. In most HPA forms, defective PAH activity results from mutations in the PAH gene. To date, more than 500 mutations, scattered throughout the gene, have been identified (http://www.pahdb.mcgill.ca). Mutation analysis is important to obtain information both about the expected phenotype and for counseling in the affected families. In our lab, we perform the molecular diagnosis of HPA with a detection rate of 90% [1, 2]. We have set-up a competitive strategy based on the high-throughput sequencing approach. From a control population and HPA patients identified from neonatal screening, we collected a blood sample (3 mL) in EDTA and extracted DNA using a standard protocol. Using the Primer3 software (http://frodo.wi.mit.edu/primer3/), we designed specific primers encompassing about 100 kb including 20 kb of the promoter region, all exons and introns, and 5 kb of the 3' UTR region of PAH gene. Subsequently, overlapping long-PCR amplicons, ranging in size between 8-12 kb, were generated using a "touch-down" protocol. Each amplicon was individually amplified and purified. After appropriate quality assessment (2100 BioAnalyzer, Agilent) and quantification (Picogreen assay, Invitrogen), the amplification products from the same DNA sample are sequenced. The PCR amplification/high throughput sequencing approach detects both single-nucleotide polymorphisms and CNVs in target genes thereby increasing the spectrum of detected variations within the analyzed gene.

1. Daniele A et al. (2008) BBA, 1782, 378
2. Daniele A et al. (2009) FEBS Journal 276, 2048

Seconda Università degli Studi di Napoli Giornate Scientifiche di Ateneo 2009, 7-10 luglio 2010

Determinazione delle mutazioni nel gene PAH in pazienti affetti da HPA attraverso HIGH-THROUGHPUT sequencing.

Palmieri A.[(1)], D Argenio V.[(2)], Guerri G.[(1)], Daniele A.[(3)], Salvatore F.[(4)]

[(1)]CEINGE- Biotecnologie Avanzate Scarl, Napoli [(2)]CEINGE-Biotecnologie Avanzate S.C.aR.L., Napoli - Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli [(3)]CEINGE-Biotecnologie Avanzate Scarl, Napoli;Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Facoltà di Scienze del Farmaco per l'Ambiente e la Salute [(4)]CEINGE- Biotecnologie Avanzate Scarl, Napoli; Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli Federico II , Napoli; Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Facoltà di Scienze

La Fenilchetonuria (PKU) e le sue varianti, le Iperfenilalaninemie (HPA) derivano dalla deficienza della fenilalanina idrossilasi (PAH: EC 1.14.16.1), enzima epatico che utilizza la tetraidrobiopterina (BH4) per metabolizzare la Fenilalanina (Phe) in Tirosina (Tyr). Gli alti livelli di Phe e dei suoi metaboliti sono neurotossici e determinano gravi danni cerebrali. Recentemente è stato osservato che un trattamento aggiuntivo con BH4 rappresenta in alcuni pazienti un valido trattamento. Il gene PAH mappa sul cromosoma 12q23.2 codifica per una proteina di 50kDa costituita da un dominio regolatorio, uno catalitico e uno di tetramerizzazione. Nel 97% delle forme di HPA, l attività PAH difettiva è causata da mutazioni nel gene PAH; sono state identificate pi di 500 mutazioni la maggior parte delle quali distribuite sull'intera lunghezza del gene (http://www.pahdb.mcgill.ca) (1). L identificazione delle mutazioni è utile per ottenere informazioni sia sul fenotipo del paziente sia sulla proteina e per poter svolgere una pi completa consulenza genetica nelle famiglie a rischio; inoltre la diagnosi molecolare permette l'identificazione dei portatori sani di HPA e l esecuzione di diagnosi prenatali. L analisi molecolare permette inoltre una migliore classificazione dei pazienti responsivi al BH4 permettendo un approccio terapeutico pi personalizzato e quindi pi appropriato. La procedura attualmente utilizzata si basa sul sequenziamento diretto, mediante metodo Sanger del promotore e dei 13 esoni del gene PAH. Con questa strategia di analisi abbiamo identificato le mutazioni in pi 90% dei pazienti (2) per cui sarebbe auspicabile mettere a punto una procedura di analisi pi sensibile che permetta di identificare anche delezioni che coprono uno o pi esoni e che risultano difficilmente riconoscibili negli eterozigoti compositi a causa dell'effetto mascherante dell'allele non deleto. In questo lavoro, abbiamo intenzione di sequenziare mediante il sistema high-throughput GS FLX l'intero gene PAH e di identificare in modo altamente accurato e a costi sostenibili le mutazioni presenti sia nelle regioni esoniche, sia in quelle introniche e sia nelle regioni del promotore del gene PAH; in pi utilizzando questa strategia di analisi potranno essere identificate in modo accurato anche delezioni e/o inserzioni geniche. In questo modo, saremo in grado di evidenziare variazioni nucleotidiche e/o genomiche presenti negli alleli in cui non è stato possibile identificare mutazioni mediante lo screening tradizionale aumentando la detection rate dell'analisi molecolare di HPA.

1. Scriver CR & Kaufman S (2001) McGraw-Hill, New York, NY
2. Daniele A et al. (2009) FEBS Journal 276, 2048