

Arrays di microRNA per l'identificazione di tumori con origine sconosciuta

Ferracin Manuela, PhD

Università di Ferrara, Laboratorio per le Tecnologie delle Terapie Avanzate (LTTA)

I tumori di origine ignota (Cancer with Unknown Primary, CUP) rappresentano un problema clinico frequente e rilevante. Evidenze scientifiche hanno dimostrato che la maggior parte dei CUP sono metastasi di carcinomi il cui sito di origine non viene identificato dalle procedure diagnostiche. Partendo dall'ipotesi che la conoscenza del sito di origine tumorale sia un elemento fondamentale per la corretta terapia del paziente, abbiamo deciso di utilizzare i microRNA microarray per l'identificazione del sito di origine di metastasi e CUP, partendo da campioni FFPE.

Abbiamo identificato una signature di 47 microRNA in grado di discriminare tra 10 diversi possibili siti di tumore primario (accuratezza 100%) e applicato questa signature nella predizione del sito di origine di metastasi con primitivo noto (accuratezza del 86% considerando i siti predetti con probabilità > 90%) e del sito di origine di CUP. La signature è stata validata anche in un dataset indipendente.

Questi risultati hanno stabilito che i campioni FFPE possono essere utilizzati per effettuare esperimenti di microarray e per identificare il sito di origine di tumori metastatici e CUP, la cui diagnosi corretta rimane al momento non definita.

Daniela Furlan

**ACCURACY OF PYROSEQUENCING FOR EGFR MUTATION DETECTION:
DIAGNOSTIC AND CLINICAL IMPLICATIONS**

EGFR activating mutations predict responsiveness to EGFR tyrosine-kinase inhibitors (TKIs) in non-small-cell lung cancer (NSCLC) patients. Mutation screening is crucial to support therapeutic decisions and is commonly conducted using dideoxy sequencing, although its sensitivity is suboptimal in clinical settings. To evaluate the diagnostic performance of pyrosequencing, we examined *EGFR* mutation status and *EGFR* copy number in a retrospective cohort of 53 patients with NSCLCs selected for TKI therapy based on only clinico-pathological criteria.

EGFR mutations were investigated by pyrosequencing and by dideoxy sequencing. Detection rates of both methods were determined by titration assays using NCI-H1975 and HCC-827 cell lines. Increased EGFR copy number was assessed by FISH.

Pyrosequencing showed a higher detection rate than dideoxy sequencing. Tumor control rate of cases with mutant and wild-type *EGFR* was 86% and 29%, respectively ($p=0.002$). *EGFR* amplification was significantly associated with *EGFR* mutation and a positive correlation between high percentages of mutant alleles and clinical response to TKI was observed.

Pyrosequencing is more sensitive than dideoxy sequencing in mutation screening for *EGFR* mutations. Sensitivity of dideoxy sequencing was suboptimal when low frequencies of mutant alleles or low tumor cell contents were observed. Pyrosequencing enables quantification of mutant alleles that correlates well with increased EGFR copy number assessed by FISH.

Pamela Magini

Microdelezioni/microduplicazioni: la refertazione

L'analisi array-CGH viene ormai considerata e indicata come prima indagine diagnostica, al posto del cariotipo convenzionale, per i soggetti affetti principalmente da disabilità intellettive o disordini dello spettro autistico. A suo favore giocano diversi aspetti vantaggiosi, sia a livello tecnico che di resa diagnostica; infatti la necessità di una bassa quantità di DNA iniziale e l'indipendenza dalle cellule in divisione consente di ottenere un risultato in tempi brevi e l'alta risoluzione che contraddistingue l'array-CGH, soprattutto se si utilizzano piattaforme a oligonucleotidi, permette di rilevare riarrangiamenti genomici sbilanciati (copy number variants, CNV) di piccole dimensioni, non visibili con le tradizionali tecniche di bandeggio cromosomico. Quest'ultima caratteristica, oltre ad essere il punto di forza dell'array-CGH rappresenta allo stesso tempo un aspetto critico non tanto dovuto alla tecnica in sé, ma piuttosto legato ad una incompleta conoscenza delle funzioni svolte dalla maggior parte del genoma che spesso impedisce una chiara e non ambigua interpretazione del significato clinico dei riarrangiamenti identificati, rendendo complicata la compilazione di un referto di indagine genetica, nel quale il loro potenziale effetto fenotipico andrebbe indicato per facilitare la comprensione soprattutto a chi non si occupa di genetica. Generalmente, come accade per le sostituzioni di singoli nucleotidi, la migliore conferma che una variante sia patogenetica deriva dal suo riscontro in altri individui, non correlati, con un fenotipo simile e non in controlli sani. A tal fine esistono diversi database pubblici che catalogano CNV identificate in individui apparentemente sani (Database of Genomic Variants, DGV) o riscontrate in soggetti affetti da patologie genetiche, dei quali riportano, anche se non sempre, una descrizione delle principali caratteristiche cliniche (DECIPHER, ISCA).

La situazione è chiara quando le alterazioni coinvolgono regioni genomiche già note essere associate a patologia, come ad esempio le del/dup22q11.21 e le del/dup15q11q13, oppure riportate in DGV perché delete o duplicate in un notevole numero di individui sani. In questi casi i riarrangiamenti possono essere definiti rispettivamente patologici e benigni. Se l'alterazione è ereditata da un genitore sano ed è riportata in DGV anche con una frequenza non molto alta, ma non nei database clinici, è probabile che essa sia una variante polimorfica ed è possibile quindi classificarla come potenzialmente benigna. Al contrario, se un'anomalia cromosomica è di nuova insorgenza oppure segrega con la malattia all'interno di una famiglia, la probabilità che possa essere causativa del fenotipo osservato è elevata, soprattutto se coinvolge numerosi geni ed è stata riscontrata in altri individui con un quadro clinico sovrapponibile; per cui nel referto corrispondente è possibile indicarla come potenzialmente patogenetica. Problemi maggiori di interpretazione si hanno con quelle alterazioni che vengono definite "di significato incerto" perché non sono mai state descritte, non sono riportate in nessun database e sono trasmesse da un genitore sano. Infatti in questi casi non è possibile escludere che ci sia penetranza incompleta o espressività variabile, per cui, a seconda del particolare background genetico in cui si trova, il riarrangiamento comporta o meno l'espressione di un fenotipo clinico.

Diverse sono quindi le informazioni che possono essere utilizzate per dare un significato alle microdelezioni o microduplicazioni identificate dall'array-CGH: origine parentale, contenuto genico, overlap con alterazioni riportate nei diversi database, ma non sempre si arriva ad una conclusione definitiva sul potenziale patogenetico di un'alterazione, soprattutto perché esiste un'ampia variabilità di interpretazione tra i diversi laboratori che eseguono questa analisi. Alla luce di tutto ciò, diventa pertanto fondamentale una standardizzazione nella refertazione e nella comunicazione dell'esito al paziente che si basa anche su una maggiore condivisione dei risultati e

delle caratteristiche cliniche associate al fine di migliorare la definizione delle conseguenze fenotipiche delle alterazioni cromosomiche riscontrate.

RNA-Seq: a revolutionary tool for transcriptome analysis

Ernesto Picardi

University of Bari and IBBE-CNR

Sequencing of RNA has long been recognized as an efficient method for gene discovery and remains the gold standard for annotation of both coding and noncoding genes. Compared with earlier methods, massively parallel sequencing of RNA (RNA-Seq) has vastly increased the throughput of RNA sequencing and allowed global measurement of transcript abundance. In addition, RNA-Seq provides a powerful experimental tool for investigating post-transcriptional modifications as alternative splicing and RNA editing, thus enabling the study of complex eukaryotic transcriptomes and molecular rearrangements occurring herein.

We present here an overview of RNA-Seq technology as well as main bioinformatics strategies to handle huge amount of short reads allowing the identification of differently expressed genes and post-transcriptional modifications as RNA editing and alternative splicing.

Approcci metagenomici per la comprensione del microbioma umano

Monica Santamaria

Istituto di Biomembrane e Bioenergetica del CNR

Bari

Ogni individuo umano ospita da 10 a 100 trilioni di microrganismi rappresentati principalmente dai batteri che colonizzano il sistema digerente. Il catalogo di questi microbi e dei loro geni costituisce ciò che viene definito microbioma umano la cui biodiversità sormonta ampiamente quella dell'ospite e cambia considerevolmente tra diversi individui e diverse parti del corpo. Sempre più evidenze sperimentali supportano l'inestricabile legame esistente tra il microbioma e il metabolismo del suo ospite: molti dei microrganismi svolgono, infatti, un effetto benefico ed essenziale sulla salute umana, contribuendo al processamento dei nutrienti, allo sviluppo del sistema immunitario, ai processi di resistenza alla colonizzazione e a molte altre attività dell'ospite. D'altra parte, alterazioni del microbioma possono essere correlate all'eziopatogenesi di una serie di disordini tra i quali la sindrome dell'intestino irritabile, l'obesità, il cancro al colon e le malattie infiammatorie dell'intestino.

Ne deriva che la possibilità di intervenire in maniera controllata sulla composizione del microbioma simbiote potrebbe portare allo sviluppo di nuovi strumenti per migliorare la salute umana contribuendo all'evoluzione della medicina personalizzata. Ovviamente, ciò non può prescindere dalla conoscenza approfondita della complessità tassonomica e funzionale delle comunità microbiche umane e delle sue variazioni su diverse scale spaziali e temporali. La metagenomica rappresenta un approccio di elezione in questo contesto permettendo di ampliare enormemente la biodiversità osservabile non più limitata dalla necessità dell'isolamento e coltura dei singoli organismi, possibile per meno dell'1% delle specie microbiche viventi nella maggior parte degli ambienti. Tale approccio prevede generalmente l'estrazione del materiale genetico totale, costituito dal metagenoma o dal metatrascrittoma, direttamente dal campione ambientale, proveniente, nel caso specifico, da determinate localizzazioni del corpo umano. Segue il suo sequenziamento, shotgun o diretto a selezionati marcatori genetici, attraverso tecnologie di nuova generazione (Next Generation Sequencing). Infine, grazie alla disponibilità crescente di nuovi e potentissimi strumenti di analisi bioinformatica, è oggi possibile estrarre dall'enorme quantità di sequenze prodotte dalle piattaforme NGS informazioni sulla composizione tassonomica e funzionale di qualsivoglia comunità microbica ad una profondità senza precedenti.

Diagnostic Exome Sequencing in Patients with Severe Intellectual Disability

Joep de Ligt, Marjolein Willemsen, Bregje van Bon, Tjitske Kleefstra, Helger Yntema, Thessa Kroes, Anneke Vulto-van Silfhout, David Koolen, Petra de Vries, Christian Gilissen, Marisol del Rosario, Alexander Hoischen, Hans Scheffer, Bert de Vries, Han Brunner, Joris Veltman, Lisenka Vissers

Department of Human Genetics, Nijmegen Centre for Molecular Life Sciences, Institute for Genetic and Metabolic disease, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.

Intellectual disability (ID) is a common condition that carries lifelong medical and social consequences. The causes of ID remain largely unknown due to its extensive clinical and genetic heterogeneity. *De novo* mutations may play an important role in ID as most individuals present as isolated cases without family history, as has recently been demonstrated in a small number of individuals with ID. Here we evaluate the diagnostic potential and role of *de novo* mutations in a cohort of 100 patients with severe ID using family-based exome sequencing.

We sequenced the coding regions of more than 21,000 genes in 100 ID patients with an IQ below 50 and their unaffected parents. All patients were counseled and consented by a clinical geneticist prior to inclusion and studied by SNP microarrays to exclude causal *de novo* CNVs. A data analysis pipeline was developed to identify and classify *de novo*, autosomal recessive, and X-linked maternally-inherited mutations. In addition, a high-throughput re-sequencing strategy was used to confirm new candidate ID genes in a second series of 765 patients with ID. All mutations identified were interpreted in the context of the patients' clinical presentation.

A total of 79 unique coding *de novo* mutations were identified in 53 patients. *De novo* (n=10) and X-linked maternally-inherited (n=3) mutations, predicted to compromise function of known ID genes, yielded diagnoses in 13 patients. In addition, potentially causative *de novo* mutations were found in another 22 patients in genes not previously associated with ID. Screening of additional ID patients for mutations in five of these candidate genes revealed further *de novo* mutations in *DYNC1H1*, *CTNNB1* and *GATAD2B*. More importantly, phenotypic comparison of patients with *de novo* mutations in the same gene showed clear overlapping phenotypes, thereby establishing pathogenicity for these three genes when mutated. Of note, we detected no causative autosomal recessive mutations. Thus, the total diagnostic yield was 16%, mostly comprising *de novo* mutations.

We conclude that *de novo* mutations represent an important cause of ID and that exome sequencing is an effective diagnostic strategy for their detection.

Epigenetic reprogramming of genomic imprinting in pluripotent cells

Robert Feil

Institute of Molecular Genetics (IGMM), CNRS and University of Montpellier, 1919 route de Mende, 34293 Montpellier, France. E-mail: obert.feil@igmm.cnrs.fr

Epigenetic regulation plays essential roles in development and disease, particularly in cancer, and may become perturbed as a consequence of environmental factors (1). One of the best studied epigenetic phenomena in mammals is genomic imprinting, a mechanism through which genes become expressed from one of the two parental alleles only. Epigenetic perturbation of imprinted genes is causally involved in many complex diseases and is frequently observed in cancer (2). Imprinted genes are organised in evolutionarily conserved clusters that are up to 2Mb in size and that are each regulated by an 'imprinting control region' (ICR). These ICRs acquire allelic DNA methylation marks in either sperm or the oocyte (3). In studies on primordial germ cells, we found that this mechanism is linked to transcription and specific chromatin features (4). In the embryo, several ICRs transcribe non-coding RNAs (from the unmethylated allele) that mediate chromatin silencing *in cis*. At the growth-related imprinted *Kcnq1* and *Igf2r* imprinted domains we find that this ncRNA-mediated silencing process involves recruitment of histone methyltransferase complexes (G9A, PRC1/2) (5-7). Our current research explores whether ncRNAs are involved in imprinted gene activation as well. Data will be presented on the *Dlk1-Dio3* locus on mouse chromosome 12, where ncRNA expression in the pre-implantation embryo is linked to the allelic enhancer function of the domain's ICR. Expression and *cis*-accumulation of these nuclear ncRNAs is detected in ES and pluripotent iPS cells, but not in differentiated cells, and is functionally linked to the activation of imprinted expression along the entire domain. Our combined data indicate that ncRNA expression at this ICR constitutes a marker of pluripotency involved in epigenetic maintenance mechanisms, control of DNA replication timing and imprinted gene expression. The significance of these findings will be discussed.

References

1. Feil R and Fraga MF. *Nature Reviews Genetics*, 13, 97-109, 2012
2. Hirasawa R and Feil R *Essays Biochem.*, 48, 187-200, 2010
3. Kota SK, Feil R (2010). *Dev. Cell*, 19, 676-686.
4. Henckel A *et al. EMBO journal*, 31, 606-615, 2011
5. Umlauf, D *et al. Nature Genet.*, 36, 1296-3000, 2004
6. Wagschal A *et al. Mol. Cell. Biol.*, 28, 1104-1113, 2008
7. Nagano T *et al. Science*, 322, 1717-1720, 2008.

The hereditary gastric cancer

Guglielmina Nadia Ranzani

Dip. di Biologia e Biotecnologie Università degli Studi di Pavia

Gastric cancer (GC) remains a leading cause of cancer-related deaths worldwide, even though a decrease in its incidence and mortality rate has been observed in recent decades. GC is the end result of a multifactorial and multistage process. Various environmental exposures such as *Helicobacter pylori* infection, high salt diet, inadequate vitamin C uptake and smoking, have been identified as risk factors.

The great majority (about 95%) of sporadic GCs are adenocarcinomas that can be classified into intestinal and diffuse histotypes. Intestinal GCs are well differentiated and characterized by cells that resemble intestinal columnar epithelial cells, whereas diffuse GCs are characterized by poorly cohesive cells and show wide and diffuse infiltration into the gastric wall.

About 90% of GCs are apparently sporadic, while 10% of cases can be classified as familial. These cancers occur in families more often than would be expected by chance and often occur at an early age: they can be the sign of shared environmental and lifestyle factors, or, more likely, they can be the hallmark of genetic variants increasing cancer risk.

A small fraction of all GC cases (1-3%) are attributable to the Hereditary Diffuse Gastric Cancer (HDGC) syndrome, an autosomal dominant condition highly predisposing to diffuse-type GC, as well as to lobular-histotype breast cancer (LBC). By linkage analysis in Maori kindreds with early-onset diffuse-type GC, Guilford and co-workers (Guilford *et al.*, Nature 1998) identified for the first time *CDH1* as the susceptibility gene. *CDH1* follows the “two hit model”, with one allele constitutively mutated and the wild-type allele inactivated in tumor cells by somatic events, including gene deletion and promoter methylation (Grady *et al.*, Nat Genet 2000; Oliveira *et al.*, Oncogene 2004). *CDH1* encodes the cell-cell adhesion protein E-cadherin that plays a crucial role in the maintenance of cell polarity and epithelial tissue architecture. Protein loss is associated with poorly differentiated cancers; accordingly, diffuse-type GC, often with signet ring cell morphology, as well as LBC, are the clinical features of the HDGC syndrome.

During the last decade, different criteria have been proposed for the selection of HDGC families; currently, besides diffuse-type GC patients with family history of GC and/or LBC, early-onset diffuse GC cases are also considered for *CDH1* genetic testing (Fitzgerald *et al.*, J Med Genet 2010). Indeed, about 7% of apparently sporadic early-onset diffuse GCs have been found to be associated with germline alterations of *CDH1* (Suriano *et al.*, Hum Mol Genet 2003); in addition, germline mutations have been associated with invasive LBC in the absence of diffuse GC (Masciari *et al.*, J Med Genet 2007). On the whole, the estimated cumulative risk of GC by age 80 years is 67% for men and 83% for women; for women, the cumulative risk of breast cancer is 39% and the combined risk of GC and breast cancer is 90% by age 80 years (Pharoah *et al.*, Gastroenterology 2001).

To date, more than 100 different germline *CDH1* alterations have been reported; these are mainly represented by point mutations, including truncating (about 80%) and missense (about 20%) mutations (Guilford *et al.*, Gastric Cancer 2010). Constitutive large deletions and allelic expression imbalance have also been detected by appropriate molecular approaches (Oliveira *et al.*, Hum Mol Genet 2009; Pinheiro *et al.*, Hum Mol Genet 2010).

However, in spite of complementary methods utilized for *CDH1* genetic testing and of selection criteria refinement, about 50% of HDGC cases remain without a molecular diagnosis; this fraction goes up to 80-90% when probands are selected in geographic areas that are characterized by high incidence of gastric tumors. Without mutations of unequivocal pathogenicity, genetic counseling, surveillance and management of the affected individuals, and of their family members at risk for cancer, remain problematic. In *CDH1* mutation-positive healthy individuals prophylactic total

gastrectomy is recommended; indeed, the systematic histological analysis of prophylactic gastrectomies has revealed intramucosal signet-ring cell carcinoma and pre-invasive lesions (Huntsman *et al.*, N Engl J Med 2001).

To increase mutation detection rate, improving genetic counseling and management of the families at risk, and to shed new light on genotype-phenotype correlations, we started a systematic investigation of HDGC patients recruited in Italy, a geographic area characterized by a moderate-high GC risk. By analyzing both DNA and RNA with a series of complementary approaches, we searched for *CDH1* germline mutations in 32 unrelated probands selected according to consensus criteria; we identified disease-causing mutations in 20% of cases (Molinaro *et al.*, communication-SIGU 2012). Therefore, in addition to *CDH1*, other genes are likely to be associated with HDGC; new approaches, such as GWAS and next generation sequencing technologies, are expected to shed new light on GC susceptibility genes.

L'uso di SNP-ARRAY in diagnosi prenatale

Vanna Pecile

L'analisi del cariotipo di cellule in coltura è considerata il gold standard per la diagnosi prenatale da oltre 30 anni. Questa procedura ha dimostrato di essere altamente affidabile per identificare anomalie cromosomiche di numero (aneuploidie) e grandi riarrangiamenti strutturali nelle cellule fetali ottenute da prelievo di liquido amniotico o villi coriali. Ma l'uso diffuso della tecnologia array per la diagnosi di riarrangiamenti genomici in bambini con ritardo mentale idiopatico, ritardo dello sviluppo ed anomalie congenite multiple ha stimolato l'interesse per l'applicazione di tale tecnologia anche nel campo della diagnosi prenatale. Pur rappresentando il vantaggio di una maggiore risoluzione rispetto al cariotipo convenzionale, il suo utilizzo in diagnosi prenatale è attualmente limitato da diversi fattori, tra cui l'impossibilità di rilevare oltre ai riarrangiamenti cromosomici bilanciati, l'individuazione di variazioni di numero di copie di significato clinico incerto, e mosaicismi con una linea cellulare scarsamente rappresentata (inferiore al 10%). Attualmente tale tecnologia è principalmente utilizzata come approfondimento diagnostico nei casi in cui il feto presenti segni ecografici riconducibili ad una patologia cromosomica ma il cui cariotipo è risultato normale, o in presenza di anomalie cromosomiche (riarrangiamenti sbilanciati, riarrangiamenti apparentemente bilanciati de novo e cromosomi marcatori) individuate attraverso l'analisi citogenetica classica.

Rispetto alla CGH-array l'analisi SNP-array costituisce attualmente uno dei metodi più avanzati di analisi citogenetica, che permette la contemporanea determinazione del copy number con risoluzione dell'ordine delle Kb, e degli eventi di Loss of Heterozygosity (LOH) e Uniparental Disomy (UPD). La molteplicità di informazioni deriva dalla possibilità di interpretare il dato sulle sonde SNP sia a livello qualitativo (genotyping) che quantitativo (numero di copie di DNA in corrispondenza di ogni marker genetico).

Inoltre gli SNP-array, mediante il Log R e la frequenza allelica, offrono il vantaggio di determinare casi di mosaicismo che possono sfuggire alla citogenetica classica. Questo poiché talvolta cellule con anomalie cromosomiche possono avere un ciclo replicativo diverso dalle cellule normali, risultando talvolta in una mancata identificazione dell'anomalia.

Il ruolo dell'ereditarietà mendeliana nel processo di filtraggio delle varianti

Tommaso Pippucci

Negli ultimi anni, l'applicazione delle tecnologie di Next Generation Sequencing (NGS) ha permesso un notevole progresso nell'individuazione dei difetti genetici associati a rare malattie monogeniche mendeliane. In particolare, l'approccio del Whole Exome Sequencing (WES), una strategia che mira al sequenziamento simultaneo di tutti gli esoni, ha consentito di coniugare costi limitati con la capacità di individuare numerosi geni associati a patologie inaccessibili allo studio mediante metodiche tradizionali di mappaggio.

Tuttavia, diversi fattori possono limitare l'utilizzo di tale approccio. In primo luogo, la tecnologia alla base del WES è soggetta sia a segnali sia falsi negativi che falsi positivi nella rilevazione delle varianti, che possono implicare una ridotta efficienza nell'individuare la mutazione causale. In secondo luogo, distinguere la variante causale dalle centinaia di varianti rare e private presenti in ogni individuo non è semplice, dal momento che gli strumenti utili a predire l'impatto funzionale delle varianti è ancora molto limitato, specialmente se si tratta di varianti dal significato incerto.

Infine, l'eterogeneità genetica che contraddistingue la maggior parte delle malattie mendeliane e la loro contenuta prevalenza complicano ancora di più questo quadro, rendendo spesso difficile la raccolta di un numero sufficientemente adeguato di pazienti con base geneticamente omogenea.

In tale situazione dunque, lo schema di ereditarietà mendeliana (autosomico dominante, autosomico recessivo o X-linked) può essere considerato *di per sé* uno strumento di filtraggio delle varianti. La selezione di varianti in eterozigosi composta o in omozigosi che segue all'assunzione di un modello autosomico recessivo è in grado, ad esempio di ridurre drasticamente (di circa 50 volte) in ciascun soggetto sequenziato la lista dei geni candidati rispetto a un modello autosomico dominante. Tali proprietà del modello di ereditarietà sotteso alla patologia dovrebbero essere tenute in considerazione nel disegno sperimentale di uno studio di WES.

Using pre-existing knowledge for prioritizing candidate genes

Viviana Caputo, PhD

Recent developments in DNA sequencing technologies (Second Generation Sequencing) provide an opportunity for accelerating the disease gene identification process. The development of methods for coupling targeted capture DNA sequences and massively parallel sequencing has made it possible to determine cost-effectively nearly all of the coding variation present in an individual human genome, a process termed “Whole Exome Sequencing” (WES). WES offers an attractive option for new experimental designs aiming to effectively identify genes that underlie Mendelian disorders in circumstances in which conventional approaches have failed and to date has been successfully used to find the causal variant in several Mendelian diseases.

WES data are processed in 4 major steps, including reads mapping, variant calling, variant annotation and variant interpretation. A key challenge of using exome sequencing approach is how to identify disease-related DNA variants among the background of non-pathogenic variants. Candidate genes can be identified via a filter-based methodology, where variants identified in cases are checked for functionality (e.g., non-synonymous variants), novelty (not identified before), and sharing among affected individuals. Selected variants can then be prioritized using sequence conservation information (e.g PhyloP and GERP scores) and estimating functional consequences of aminoacid changes (e.g Polyphen2, SIFT).

Additionally, candidate alleles can be stratified using existing functional annotation (Gene Ontology, protein functional domains, gene expression data, pathways). The first step is to retrieve and integrate information regarding genes harboring novel variants from different resources. To this aim data mining analyses are performed using integrated biological knowledge based tools (e.g. DAVID) directed at systematically extracting biological meaning from large gene lists. Functional annotation data could represent a valuable tool to stratify candidate variants when used in combination with existing a priori biological knowledge of the particular disease under study. Candidate genes can be ranked using information concerning their predicted role in a biological pathway that is already proven to be involved in a similar phenotype or their interactions with genes that are known to cause a similar phenotype.

Additionally, pre-existing knowledge concerning the phenotype can be used to prioritize candidate genes using text mining algorithms (e.g. Génie) that evaluate gene function based on analysis of the literature. By using combinations of terms that characterize the phenotype those algorithms evaluate available MEDLINE records to rank candidate genes potentially linked to the phenotype.

Several approaches to find causal variants by filtering and prioritizing candidate genes have been developed representing one of the most active fields of research on Mendelian diseases. Strategies

based on the use of cellular networks and ontologies of phenotypic features have proven to enhance the power of WES based approach even when coupled to the sequencing of single individuals with rare diseases.

Bioinformatic tools for annotation, interpretation and prediction of pathogenic effect of variants from NGS studies

Matteo Benelli

Diagnostic Genetic Unit, Careggi University Hospital, Firenze, Italy The emergence of recent sequencing-based platforms allows for the investigation of the entire genome or whole exome with unprecedented resolution. However, the large amount of data produced by these technologies introduces significant challenges in the handling and the analysis of the data generated. This is why the need of computational methods that are able to reduce the vastness of inputs into meaningful and fundamental information is so important. In fact, thousands of genomic alterations are typically identified by SNPs or InDels caller algorithms and a key challenge is to pinpoint a small subset of variants that are potentially clinically and functionally relevant.

In the last years many bioinformatics tools have been developed to add meaning (annotation) and select variants and to determine the effect of genomic variants on transcripts and proteins in the context of NGS analysis.

Here, we review the available open source tools that help researchers to search for candidate variants and we discuss how a good use of biological databases can improve our ability in distinguishing between real candidates and noisy calls.

Pamela Magini

Singole famiglie, malattie ultrarare: successi e problematiche inerenti al follow-up delle varianti candidate

Lo sviluppo di tecniche ad alta risoluzione, in particolare del Next Generation Sequencing (NGS), ha consentito in pochi anni l'identificazione di numerosi geni responsabili di malattie genetiche a trasmissione mendeliana, rivelandosi particolarmente efficace nella risoluzione di condizioni patologiche che si manifestano in casi sporadici o in piccole famiglie e che per questo non sono analizzabili utilizzando strategie tradizionali, quali l'analisi di linkage. Poiché questa tecnologia rileva decine di migliaia di varianti nucleotidiche in ogni singolo individuo, l'attenzione si è subito focalizzata sullo sviluppo di strumenti bioinformatici per filtrarle tenendo conto della qualità dell'esperimento e del loro significato biologico in modo da arrivare ad identificare un'unica alterazione potenzialmente patogenetica. Raggiunto questo obiettivo, la dimostrazione definitiva che difetti di quello specifico gene determinano il fenotipo clinico osservato, è data dal riscontro di ulteriori mutazioni in individui affetti non imparentati e/o da evidenze sperimentali che provino una disfunzione del gene correlabile ai meccanismi patogenetici responsabili della malattia. Con patologie geneticamente eterogenee trovare anche solo un altro soggetto con alterazioni dello stesso gene richiederebbe l'analisi di diverse centinaia di individui con l'impiego di notevoli risorse di cui non tutti i laboratori dispongono e con malattie ultra-rare spesso si lavora su fenotipi unici difficilmente osservabili al di fuori della famiglia in esame. Per questo motivo in questi casi è indispensabile dimostrare il ruolo causativo della mutazione attraverso esperimenti funzionali, che costituiscono la fase meno standardizzata del processo di identificazione dei geni-malattia in quanto dipendente da vari fattori che non sono noti a priori al momento del disegno di studio iniziale poiché sono strettamente legati alle caratteristiche specifiche del gene alterato e al tipo di mutazione.

A parte la possibilità di predire in silico l'effetto della mutazione sullo splicing o sulla struttura tridimensionale della proteina attraverso l'impiego di software o tools bioinformatici dedicati, gli esperimenti più immediati e i cui risultati rispecchiano meglio la reale fisiologia dell'organismo, sono quelli che impiegano materiale biologico prelevato dal paziente, quando il tessuto utile allo scopo è facilmente ottenibile. In caso contrario occorre studiare in vitro la specifica funzione che si pensa possa essere alterata. La scelta del saggio da eseguire dipende dalle informazioni disponibili relativamente al gene in esame, tra cui ovviamente l'attività biologica della proteina codificata, il pattern di espressione tissutale e i pathway molecolari in cui esso interviene. Se questi dati non sono noti per il gene umano, ma sono stati ottenuti e pubblicati per un ortologo, è possibile utilizzare le informazioni ricavate dal modello animale per pianificare gli studi funzionali.

Infine, la migliore conferma che una mutazione sia causativa si ottiene riproducendo il fenotipo clinico osservato nel paziente in un modello animale tramite esperimenti in vivo.

In conclusione, una volta identificata la variante potenzialmente causativa, occorre approfondire la letteratura relativa al gene di interesse e scegliere gli esperimenti funzionali che possano fornire un quadro chiaro e completo dell'effetto che essa ha sulla funzionalità della proteina e sulla fisiologia cellulare. L'esecuzione di questo tipo di studi non si rivela mai semplice, soprattutto perché richiede una certa esperienza nella pratica dei vari saggi, ognuno dei quali è caratterizzato da numerose variabili, e la possibilità di accedere alle attrezzature necessarie di cui non tutti i laboratori dispongono. Un approccio multidisciplinare con gruppi di ricerca che abbiano già sviluppato tra le

loro principali linee di ricerca una parte di analisi funzionale sul gene in esame può risolvere questi problemi e accelerare la dimostrazione della causalità della variante identificata.

L'uso del salbutamolo nell'atrofia muscolare spinale: dagli studi in vitro alla sperimentazione clinica controllata

Francesco D. Tiziano

Ist. Gen. Med., Univ. Cattolica, Roma

L'atrofia muscolare spinale (SMA) è una condizione autosomica recessiva, caratterizzata dalla degenerazione dei motoneuroni delle corna anteriori del midollo spinale, causata dalla perdita in omozigosi del gene *SMN1*. Sia *SMN1* che un gene quasi identico (*SMN2*) codificano per la proteina SMN ma, a causa dello *splicing* alternativo dell'esone 7, i geni *SMN2* producono bassi livelli di proteina SMN full length. I livelli di SMN sono marcatamente ridotti nei pazienti SMA e correlano inversamente con la gravità fenotipica. Sebbene non vi sia ad oggi una terapia per la SMA, diverse evidenze sperimentali hanno messo in evidenza come l'espressione dei geni *SMN2* possa essere incrementata mediante trattamento farmacologico con diversi composti, tra cui il salbutamolo, un agonista del recettore beta2-adrenergico (B2AR). Una prima segnalazione in letteratura, riportata da Kinali et al. (2001), suggeriva in uno studio aperto come il trattamento con salbutamolo incrementasse la forza muscolare in un gruppo di pazienti SMA II-III. Questi dati preliminari ci hanno spinto a valutarne l'effetto in vitro su fibroblasti di pazienti in coltura ed abbiamo dimostrato come il salbutamolo induca un rapido e persistente incremento dei livelli di trascritti e proteina SMN, agendo in maniera prevalente sullo *splicing* alternativo dell'esone 7 (Angelozzi et al., 2008). Successivamente, sono stati effettuati due studi open label in vivo: uno molecolare, che ha dimostrato come il trattamento induca un incremento dei livelli di trascritti *SMN2* full length in leucociti di sangue periferico di pazienti (Tiziano et al. 2010), ed uno clinico che ha dimostrato un incremento della performance motoria di pazienti SMA II (Pane et al., 2008).

Siamo stati infine coinvolti in uno studio in doppio cieco, controllato con placebo il cui obiettivo era di valutare tollerabilità ed efficacia del composto in 45 pazienti adulti affetti da SMA III. I risultati del nostro studio hanno dimostrato che il salbutamolo è ben tollerato, che induce un incremento statisticamente significativo dei trascritti dei geni *SMN2*, ed un miglioramento significativo della performance motoria nella maggioranza dei pazienti (v. anche poster P261).

Tuttavia, sebbene diversi dati suggeriscano l'utilità potenziale del salbutamolo nel trattamento della SMA, i dati oggi disponibili non sono ancora sufficienti ad ottenere la registrazione del farmaco per la condizione. Sono necessari ulteriori studi in doppio cieco controllati con placebo, di fase III e IV, che ne dimostrino l'efficacia clinica.

Prospettive per una Farmacoterapia in Fibrosi Cistica

Mauro Mazzei

Dipartimento di Farmacia - Università di Genova - Viale Benedetto XV, 3 -16132 Genova
mauro.mazzei@unige.it

La Fibrosi Cistica (FC) è una malattia genetica causata da mutazioni di una proteina di membrana (CFTR) che ha la funzione di canale per il cloruro. Sono note ca. 1800 mutazioni del gene che codifica per la CFTR, ma l'alterazione più grave e più diffusa è data dalla delezione della fenilalanina nella posizione 508 (DF508-CFTR). Circa il 90% dei malati FC possiede almeno un allele con tale mutazione. Anche altre mutazioni numericamente meno rappresentate (ad es. G551D) hanno effetti pesantemente negativi.

La mancanza o il malfunzionamento della proteina CFTR impedisce allo ione cloruro e all'acqua ad esso associata di fuoriuscire dalle cellule epiteliali e ciò porta alla formazione di un muco denso e disidratato che ostruisce tutti i dotti interni (soprattutto nei polmoni e nel pancreas) ed è il luogo ideale per lo sviluppo di infiammazioni ed infezioni che nel tempo divengono insensibili alle cure. La diagnosi per la FC può essere fatta con il test genetico o, più facilmente, con il test del sudore (positivo se lo ione cloruro è >60 mEq/L).

Attualmente la vita media della persona recante una mutazione grave è di ca. 40 anni. A raggiungere questo valore hanno contribuito le cure con antibiotici, enzimi pancreatici, anti-infiammatori, ecc., ma la qualità di vita del paziente FC rimane ancora di basso livello in quanto il paziente, oltre che assumere farmaci, deve dedicare giornalmente molto tempo a faticose sedute di fisioterapia per mobilizzare il muco e liberare le vie aeree. Quando la malattia respiratoria si aggrava l'ultima possibilità rimane il trapianto di polmone.

In attesa di una efficace terapia genica, si sta tentando da qualche tempo un approccio farmacologico il cui razionale prende l'avvio da una più approfondita conoscenza della fisiologia della CFTR. La CFTR è stata clonata nel 1989, però non è ancora possibile avere tutta la proteina cristallizzata. Poiché si hanno solo dati cristallografici incompleti, le ricerche non si possono avvalere di modelli computazionali che faciliterebbero il lavoro nella sintesi di farmaci. Per questa ragione il ritrovamento di sostanze attive sulla CFTR è soprattutto avvenuto mediante "screening" di librerie composte da migliaia di molecole usando metodi ad alte prestazioni (HTS) che permettono di esaminare un gran numero di potenziali farmaci in breve tempo. Si sono così individuati molti composti capaci di attivare la CFTR mutata: queste sostanze sono definite "correttori" se riescono a far giungere la DF508-CFTR sulla membrana plasmatica senza essere degradata a livello del reticolo endoplasmatico (RE) o "potenziatori" se riescono ad attivare la proteina mutata quando è in membrana ma ha il canale non funzionante (caso della G551D). Il meccanismo di azione di questi sostanze non è ancora ben chiarito. Esistono poi delle sostanze

(acido 4-fenilbutirrico, matrina) che agiscono a livello dei chaperoni che normalmente aiutano il folding delle proteine. E' stato trovato che l'inibizione del chaperone HSC70 permette alla DF508-CFTR di uscire indenne dal RE e di collocarsi sulla membrana.

I composti che *in vitro* attivano la CFTR mutata sono oggi un buon numero: uno solo di essi, il VX770 (KalydecoTM), ha finora raggiunto il mercato. Essendo un potenziatore, agisce su pazienti con mutazione G551D ma non su quelli che hanno la DF508-CFTR. Per immettere sul mercato anche farmaci attivi sulla DF508-CFTR la ricerca in FC sta alacremente lavorando.

Nuove prospettive terapeutiche per la sindrome X fragile

Elisabetta Tabolacci

La sindrome X fragile (FXS, MIM #300624) rappresenta la principale causa di disabilità intellettiva ereditaria. E' causata da espansione e successiva metilazione della sequenza CGG localizzata al 5' UTR del gene FMR1, ciò determina il silenziamento del gene e la conseguente assenza della proteina FMRP. Tale assenza causa la mancata repressione del pathway mediato dai recettori metabotropici del glutamato tipo 5 (mGluR5), determinando in parte il fenotipo cognitivo-comportamentale tipico della FXS [Bear et al., 2004]. In un recente studio clinico multicentrico in doppio-cieco, randomizzato e incrociato è stato dimostrato l'effetto benefico (miglioramento dell'attenzione, diminuzione dell'ansia e dell'iperattività) del composto AFQ056, inibitore selettivo dei mGluR5, su pazienti FXS con mutazione del gene completamente metilata rispetto ai pazienti FXS con mutazione parzialmente metilata [Jacquemont et al., 2012].

Per determinare se l'AFQ056 potesse avere un effetto epigenetico secondario sulla metilazione e trascrizione di FMR1, abbiamo trattato con questo composto tre linee cellulari di linfoblasti provenienti da pazienti FXS e un controllo maschile normale. Attraverso RT-PCR quantitativa abbiamo osservato che l'AFQ056 non è in grado di riattivare il gene FMR1 silente. Inoltre mediante sequenziamento previa trasformazione del DNA con bisolfito di sodio abbiamo osservato che lo stato di metilazione del promotore del gene non si modifica rispetto al controllo normale [Tabolacci et al., 2012].

Questi dati dimostrano che l'effetto dell'AFQ056 sui pazienti con mutazione del gene completamente metilata non è dovuto ad un effetto epigenetico di questo composto, ma piuttosto ad un'azione di modulazione sui recettori metabotropici del glutamato. Questi trattamenti aprono la strada all'uso di molecole target che bypassano l'assenza di FMRP, andando a modulare il fenotipo della FXS.

Title:

A mutation threshold distinguishes the antitumorigenic effects of the mitochondrial gene MTND1, an oncojanus function

Authors & affiliations:

Anna Maria Porcelli¹, Claudia Calabrese², Luisa Iommarini¹, Ivana Kurelac², Mariantonietta¹ Capristo, Maria Antonietta Calvaruso², Apollonia Tullo³, Mariano Caratozzolo³, Marcella Attimonelli⁴, Claudio Ceccarelli⁵, Giordano Nicoletti⁶, Patrizia Nanni⁶, Carla De Giovanni⁶, Michela Rugolo¹, Pier Luigi Lollini⁶, Giovanni Romeo², Giuseppe Gasparre².

¹Dipartimento di Biologia Ev. Sp., Università di Bologna, Bologna, Italy

²Dipartimento di Scienze Ginecologiche, Università di Bologna, Bologna, Italy

³Istituto di Tecnologie Biomediche, CNR, Bari, Italy

⁴Dipartimento di Biochimica e Biologia Molecolare, Bari, Italy

⁵Dipartimento di Scienze Istopatologiche, Università di Bologna, Bologna, Italy

⁶Dipartimento di Cancerologia, Università di Bologna, Bologna, Italy

The recently defined function of *oncojanus* concerns the ability of high loads of mutations in NADH-ubiquinone-oxidoreductase (complex I) mitochondrial DNA-encoded genes to arrest tumor growth by impeding hypoxia inducible factor-1 α (HIF1 α) stabilization and, hence, adaptation to hypoxia. Through allotopic expression of wild-type MTND1 oncojanus in MTND1-null osteosarcoma cells, we managed to recover complex I assembly and NADH-oxidase activity, shifting the balance of Krebs cycle metabolites α -ketoglutarate/succinate, essential to recuperate HIF1 α stabilization. Allotopic cells injected in mice generated larger and more aggressive tumor masses compared to MTND1-null cells. Next-generation-sequencing whole-transcriptome profiling and metabolites measurement of allotopic versus MTND1-null xenografts showed that, despite remaining prevalently glycolytic, NADH-oxidase activity was fundamental to activate survival pathways needed for tumors progression and adaptation to hypoxia. Our study defines a specific role in tumorigenesis of a fundamental metabolic enzyme such as the multiprotein respiratory complex I and identifies its NADH-oxidase function as a potential anti-cancer target.

Implicazioni delle mutazioni a carico del DNA mitocondriale nel neuroblastoma

Mario Capasso

Il neuroblastoma (NBL) è un tumore neuroendocrino che origina dalla cresta neurale del sistema nervoso simpatico. Nonostante i notevoli miglioramenti nella percentuale di guarigione di molti tumori infantili, il neuroblastoma rimane un importante problema clinico, pari al 15% dei decessi attribuibili alle condizioni maligne nei bambini (Maris et al., Lancet, 2007). Il tasso di sopravvivenza, dal 1999 al 2005, è migliorato solo per i pazienti con la forma più benigna della malattia (www.seer.cancer.gov), mentre tra i bambini ad alto rischio di sviluppare NBL (40%) ha mostrato solo un modesto miglioramento. Negli ultimi anni, diversi studi hanno evidenziato alterazioni genetiche associate al NBL che possono avere un ruolo cruciale in una migliore stratificazione del rischio e nella guida verso trattamenti efficaci. Sia le mutazioni germinali che quelle somatiche nel gene ALK sono state identificate nei pazienti NBL (Chen et al, Nature, 2008; Janoueix-Lerosey et al, Nature, 2008; Mosse et al, Nature, 2008). Queste mutazioni sono responsabili della maggior parte dei rari casi di NBL familiare, e di una frazione più piccola dei tumori sporadici (6-9%). Di recente, il nostro gruppo di ricerca in collaborazione con un gruppo americano, ha dimostrato, mediante uno studio di associazione “genome-wide” (GWAS), che comuni sostituzioni nucleotidiche polimorfiche (SNPs) nei geni LINC003407 (Maris et al, New England Journal of Medicine, 2008), BARD1 (Capasso et al, Nature Genetics, 2009), LMO1 (Wang et al, Nature, 2009) DUSP12, HSD17B1210, DDX4/IL31RA10 (Nguyen Le et al, Plos Genetics, 2011), HACE1 e LIN28B (Diskin et al, Nature Genetics, 2012) sono altamente associate al neuroblastoma. Questi risultati dimostrano, inequivocabilmente, che multiple comuni variazioni del DNA collaborano ad aumentare il rischio di trasformazione neuroblastica maligna. Nonostante questi importanti risultati, la complessa architettura genetica del NBL rimane in gran parte sconosciuta e sono ancora molti i meccanismi molecolari, responsabili dell’insorgenza del tumore, che devono essere svelati.

Il NBL è un tumore molto eterogeneo sia sotto il profilo clinico che genetico. Di recente, il Children Oncology Group ha stabilito fino a quattro categorie di rischio (*very low-risk*, *low-risk*, *intermediate risk* e *high-risk*) e diversi interventi terapeutici basati su tale classificazione. Essa si basa principalmente su fattori clinico-patologici come età alla diagnosi, stadio del tumore (INSS stage), istologia del tumore e fattori genetici quali amplificazione di MYCN e delezione del cromosoma 11q. Tale classificazione di rischio insieme con gli avanzamenti scientifici nel campo della ricerca di base e clinica hanno reso possibile il fatto che più del 90% dei bambini affetti da NBL di tipo *low-risk* guariscono. Purtroppo tra i pazienti con NBL di tipo *high-risk* la percentuale di mortalità rimane elevata, di circa il 60%. Quindi una più accurata stratificazione dei pazienti e lo sviluppo di terapie mirate sono gli elementi cardini per migliorare il decorso clinico e aumentare la sopravvivenza dei bambini ammalati di NBL. Come è stato già ampiamente dimostrato, l’identificazione di nuovi *marker* genetici può dare un significativo contributo sia nel migliorare la stratificazione di rischio del NBL sia nel guidare lo sviluppo di nuove terapie.

Il ruolo dei mitocondri nei complessi eventi molecolari che inducono la progressione tumorale verso la malignità è ancora ampiamente dibattuto, dato che è stato mostrato che le disfunzioni mitocondriali possono avere effetti sia pro- che anti-tumorigenici. Infatti, le alterazioni del DNA mitocondriale (mtDNA) sono generalmente considerate come pro-tumorigeniche ma, in alcuni casi, l’accumulo di mutazioni distruttive nei geni del *complesso I* (CI) sono associate ad un fenotipo benigno, come nel caso dell’oncocitoma. Dunque, i geni del mtDNA possono agire sia come oncosoppressori sia come oncogeni geni di letalità.

Ancora oggi non si sa con esattezza se esiste un diretto legame tra alterazioni genetiche e funzionali del mitocondrio con sviluppo e progressione del NBL. Molto di recente uno studio di proteomica ha

dimostrato che, in linee cellulari di NBL, il Sorafenib (farmaco antitumorale), inibisce la funzione del CI nei mitocondri e questo comporta una mancata funzione del potenziale di membrana mitocondriale e di conseguenza apoptosi cellulare (Bull et al, Journal of Proteome Research, 2012). Un altro recente studio ha dimostrato una massiva riduzione della succinico deidrogenasi (SDH) nei tessuti di neuroblastoma. Gli stessi tessuti non presentavano mutazioni nei gene SDH, SDHAF1, SDH5, nfs-1, LYRM4, frataxin (Feichtinger et al., BMC Cancer, 2010). In accordo con questo lavoro, altri due gruppi di ricerca non hanno identificato mutazioni nei geni SDHD e SDHB in 113 pazienti (De Preter K et al., BMC Cancer et al, 2004 e Astuti et al, British Journal of Cancer, 2004). Comunque, fino ad oggi, nessuna ricerca è stata focalizzata sull'analisi genomica globale del mtDNA nel NBL. Di conseguenza, abbiamo, di recente, iniziato un progetto che prevede il sequenziamento dell'intero genoma del mtDNA sia in pazienti *high-risk* che *low-risk*. Il nostro scopo è di stabilire se elevate frequenze di mutazioni a carico del mtDNA sono associate al neuroblastoma *high-risk* o *low risk*, e se possibile include tali marcatori nella corrente classificazione di rischio del NBL. In questa prima fase della nostra ricerca abbiamo già sequenziato il mtDNA somatico di 16 pazienti *low-risk* e abbiamo quasi concluso il sequenziamento di 19 mtDNA somatico di pazienti *high-risk*.

Risultati preliminari in 16 pazienti con NBL low-risk

Abbiamo identificato 52 mutazioni geniche con una variabilità totale inferiore a 0.2 e tra queste solo 7 mutazioni (13%) nei geni rRNA 16S, tRNA^{cys} and COII avevano una variabilità uguale a zero quando confrontate con i dati depositati in HmtDB (Human Mitochondrial DataBase). Due di queste sette mutazioni erano localizzate nel gene rRNA 16S ed erano comuni a due pazienti. Inoltre, 34 mutazioni erano localizzate nelle regioni D-loop del mtDNA.

Abbiamo frazionato le mutazioni secondo quattro categorie funzionali e abbiamo osservato che i geni appartenenti alla categoria Complex I risultavano avere il numero più elevato di mutazioni (vedi tabella).

Tipo di mutazione	N (%)
<i>Complex I mutations</i>	
ND1, ND2, ND3, ND4, ND4L, ND5, ND6	20 (38.5)
<i>Complex III, IV e V mutations</i>	
Cytb, COI, COII, COIII, ATPase6, ATPase8	19 (36.5)
<i>tRNA e rRNA mutations</i>	
tRNA, 12S rRNA, 16S rRNA	13 (25.0)
<i>D-loop mutations</i>	
D-loop	34

Mentre presi singolarmente, i geni ND5 e rRNA 16S risultavano essere i più mutati (8/52 per ognuno).

Il numero totale delle mutazioni sinonime era pari a 28 mentre quelle non-sinonime era pari a 10. Tra queste ultime vi era una localizzata nel gene ND4 che mostrava un'elevata conservazione tra i

mammiferi e il valore di predizione relativo al danno della proteina più alto rispetto alle altre mutazioni analizzate.

Lo studio qui presentato dimostra che c'è un'alta frequenza di mutazioni che avvengono nel mtDNA in soggetti con NBL di tipo *low-risk*. Questo è il primo lavoro in cui è stato eseguito il sequenziamento di tutto il genoma mtDNA nel NBL e che ha fornito un quadro generale di quelle che potrebbero essere le mutazioni di un sottotipo di NBL con caratteristiche benigne. Il passo successivo sarà quello di completare il sequenziamento e l'analisi dei dati ottenuti dal mtDNA di pazienti *high-risk* e confrontare i risultati con quelli derivanti dall'analisi dei *low-risk*. Sarà poi cruciale verificare quali di queste mutazioni insorgono unicamente al livello del mtDNA somatico e valutare una loro specifica funzione molecolare nel tumore. Sarebbe poi interessante valutare se particolari profili di mutazioni sono associati a variabili clinico-patologiche e valutare infine se essi possono contribuire al miglioramento dell'attuale schema di classificazione del rischio del NBL.

Bioinformatic approaches to study mtDNA variability in cancer

Francesco Maria Calabrese, Mariangela Santorsola, Sofia Angione, Domenico Simone, Roberta Piredda and Marcella Attimonelli

Department of Biosciences, Biotechnology and Pharmacological Sciences, University of Bari "Aldo Moro", Bari, Italy

The debate on the newly mitochondrial Reconstructed Sapiens Reference Sequence (RSRS) [1], which has replaced the forefather revised Cambridge reference sequence (rCRS) [2], rekindled interest in mitochondrial DNA molecule and its related features. HmtDB (Human mitochondrial DataBase) contains, until now, about 10.000 mitochondrial genomes [3], 1/5 of which from pathological specimens. Thanks to today's NGS techniques spread, the same allowing more correct heteroplasmy estimation, this number is growing. HmtDB represents one of the reference resources where specific tools help clinicians to study mitochondriopathies. Onward the information retrieval of site specific nucleotide and aminoacidic statistical data, included in each genome linked schedule (genome card), among the latest added changes, the possibility to detect differences versus different major allele reference genomes (MARGe-s). Sequences subset derived from the same continent or belonging to the same mitochondrial macro-haplogroups (world Phylotree haplogroup classification) have been grouped and the resulted estimated major allele per position sequences has been used to build MARGe-s. The sequence is generated on the basis of the data produced by applying the SiteVar [3] software available in HmtDB. To support the usage of the site_var algorithm in the generation of MARGe, the variability profile used has been compared with the Entropy profile generated starting from the multialignment of the same set of genome. The "Classify" tool for haplogroup predictions through the comparison of each genome with RSRS allows differences detection with respect to the aforementioned MARGe profiles. This leads to a new variability checking method contributing to detect sequences quality, "personal mutations" and their phenotypical effects. This method has been applied to a set of human neuroblastome mitochondrial genomes (Implicazioni delle mutazioni a carico del DNA mitocondriale nel neuroblastoma, Capasso et al, XV CONGRESSO NAZIONALE SIGU Sorrento 21-24 November 2012).

References:

1. A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Behar DM et al, Am J Hum Genet. 2012 Apr 6;90(4):675-84.
2. Sequence and organization of the human mitochondrial genome, Anderson et al, Nature. 1981 Apr 9;290(5806):457-65.
3. HmtDB, a genomic resource for mitochondrion-based human variability studies. Rubino F et al, Nucleic Acids Res. 2012 Jan;40(Database issue):D1150-9. Epub 2011 Dec 1.

Basi cliniche e genetiche delle malformazioni congenite del cervelletto e del troncoencefalo

Enza Maria Valente, MD, PhD

Le malformazioni congenite del cervelletto e del troncoencefalo (Cerebellar and Brainstem Congenital Defects - CBCD) interessano circa 1 su 3000-4000 nati con conseguenze rilevanti in termini di mortalità e morbidità, e rappresentano una causa frequente di interruzione volontaria di gravidanza quando diagnosticate in epoca prenatale. Numerose classificazioni anatomiche ed eziologiche sono state sinora proposte, nessuna delle quali è stata accettata su larga scala. Negli ultimi anni, la ricerca ha fatto notevoli progressi nella identificazione delle basi genetiche delle CBCD su base mendeliana, come la sindrome di Joubert e disordini correlati (JSRD), le ipoplasie pontocerebellari (PCH), ed alcune sindromi X-linked caratterizzate da ipoplasia cerebellare (OPHN1, CASK). In particolare, l'identificazione di numerosi geni causativi delle JSRD e le conseguenti correlazioni genotipo-fenotipo hanno consentito una preliminare riclassificazione di queste sindromi su base molecolare, che tuttavia rimane incompleta in quanto mutazioni nei geni ad oggi noti sono identificate in non più della metà dei casi.

A fronte dei risultati ottenuti nelle CBCD mendeliane, poco ancora è noto sulla eziopatogenesi delle numerose forme sporadiche, verosimilmente causate da mutazioni eterozigoti de novo o riarrangiamenti genomici a penetranza variabile. Diverse anomalie citogenetiche sono state riportate in numerosi pazienti affetti da sindrome di Dandy-Walker, e l'identificazione di riarrangiamenti ricorrenti di specifiche regioni genomiche ha portato all'identificazione di alcuni geni candidati del fenotipo malformativo. Tuttavia, studi mutazionali di tali geni non sono mai stati riportati ed il reale impatto di mutazioni in tali geni nella patogenesi della sindrome resta sconosciuto. Molte altre CBCD sporadiche mancano di una chiara definizione genetica, e nuove sindromi malformative (come ad esempio la "pontine tegmental cap dysplasia") sono state solo recentemente descritte, anche grazie al contributo di sofisticate indagini neuroradiologiche quali la trattografia.

Ad oggi, numerose questioni restano aperte riguardo a problematiche come la nosologia delle CBCD, la disponibilità di test genetici per diagnosi pre- e post-natale, la relativa consulenza genetica, e l'identificazione di indici prognostici adeguati. Tuttavia, le nuove tecnologie ad oggi disponibili per l'analisi del nostro genoma (come ad esempio il sequenziamento dell'intero esoma e l'analisi di microriarrangiamenti genomici su larga scala), verosimilmente porteranno nel prossimo futuro alla caratterizzazione patogenetica di un numero sempre più ampio di CBCD, rivoluzionando le nostre conoscenze in questo campo e di conseguenza l'approccio clinico-diagnostico e la gestione dei pazienti.

Inquadramento clinico e genetico delle malformazioni congenite sovratentoriali

Elena Parrini PhD

I progressi effettuati nello studio delle malformazioni dello sviluppo corticale hanno permesso, in particolar modo negli ultimi anni, di evidenziare l'elevata frequenza di queste patologie in soggetti con ritardo mentale, ritardo dello sviluppo ed epilessia, specialmente in età pediatrica. Le malformazioni corticali sono spesso dovute ad un disturbo della migrazione neuronale, un processo che avviene durante le fasi iniziali dello sviluppo embrionale. La diagnostica per immagini costituisce il principale mezzo per l'identificazione e la classificazione delle diverse sindromi malformative e ne permette un rapido inquadramento clinico-diagnostico. Numerosi geni sono stati recentemente identificati come responsabili di alterazioni del processo di migrazione neuronale, attraverso cui si formano i sei strati definitivi della corteccia cerebrale. Lo studio di questi geni è di fondamentale importanza per la comprensione delle anomalie dello sviluppo corticale.

La lissencefalia-pachigiria e l'eterotopia a banda sottocorticale (SBH) costituiscono uno spettro malformativo associato a mutazioni dei geni *LISI* o *DCX*. Mutazioni nel gene *LISI* determinano un pattern malformativo più grave nelle regioni posteriori dell'encefalo e sono generalmente associate ad un grave ritardo dello sviluppo ed epilessia anche se raramente possono causare fenotipi meno severi tra cui l'SBH in sede posteriore dovuta a mutazioni sotto forma di mosaico somatico. Mutazioni nel gene *DCX* sono generalmente associate a lissencefalia nei maschi, predominante nella parte anteriore dell'encefalo, mentre causano SBH nelle femmine. Altri geni implicati in forme più rare di lissencefalia sono: *reelin (RELN)*, nel quale sono state osservate mutazioni in pazienti con una forma autosomica recessiva di lissencefalia associata a grave ipoplasia cerebellare, anomalie morfologiche facciali, ritardo dello sviluppo ed epilessia; *TUBA1A*, implicato in quadri di lissencefalia di grado variabile associata a ritardo mentale; *ARX*, che è alterato in individui di sesso maschile con lissencefalia X-linked associata a agenesia del corpo calloso e genitali ambigui.

L'eterotopia nodulare periventricolare (PNH) X-linked è presente prevalentemente in pazienti di sesso femminile ed è spesso associata ad epilessia. Mutazioni nel gene *FLNA* sono presenti in tutti i casi familiari e nel 26% dei pazienti sporadici. È stata inoltre descritta una rara forma recessiva di PNH in pazienti con mutazioni nel gene *ARFGEF2*.

La polimicrogiria (PMG) è una malformazione della corteccia cerebrale in cui è stata evidenziata, almeno in alcune forme, una causa genetica. La polimicrogiria bilaterale perisilviana (BPP) è geneticamente eterogenea in quanto sono stati descritti modelli di trasmissione autosomica dominante, autosomica recessiva e X-linked. Sono stati descritti inoltre pazienti con BPP e mutazioni di *SRPX2* e pazienti con BPP e delezione della regione cromosomica 22q11.2. La polimicrogiria bilaterale frontoparietale (BFPP) è una forma di polimicrogiria autosomica recessiva associata a mutazioni nel gene *GPR56*. Infine, sono state identificate recentemente identificate mutazioni nel gene *TUBB2B* in pazienti con polimicrogiria asimmetrica.

La genetica molecolare, l'analisi array-CGH e il sequenziamento massivo parallelo rappresentano gli strumenti più importanti per identificare alterazioni in geni noti e per indirizzare lo studio su nuovi loci, al fine di effettuare una migliore diagnosi molecolare nei pazienti con disturbi della migrazione neuronale e offrire una migliore consulenza genetica alle famiglie di questi pazienti.