STUDY OF HEREDITARY VARIANTS AND ALTERED METHYLATION PATTERNS IN IMPRINTING DISORDERS

Carolina Monzó[1] , Arrate Pereda[2], Javier Errea[2], Azahara Fuentes[1], Verónica Lendínez[1], Vicente Arnau[3], Pablo Marín[4], Ana Bárbara García[5], Guiomar Perez de Nanclares[2], Felipe Javier Chaves.[1,5,6]

1 Genomic and Genetic Diagnosis Unit, INCLIVA Research Institute; 2 Molecular (Epi)Genetics Laboratory, BioAraba National Health Institute, Hospital Universitario Araba-Txagorritxu;
3 Informatics Department, School of Engineering, University of Valencia, I2SysBio; 4 Experimental Sciences and Mathematics Department, Catholic University of Valencia, Mgviz.org;
5 CIBER of Diabetes and Associated Metabolic Diseases (CIBERDEM); 6 Sequencing Multiplex SL

INTRODUCTION

Genomic imprinting refers to specific allelic transcriptional silencing regulated by differentially DNA methylated regions (DMRs). Imprinting marks elude postzygotic reprogramming and are maintained throughout development on all somatic tissues.

Disruption of expression in parentally imprinted genes, results in imprinting disorders, a group of congenital diseases affecting growth, puberty, behavior, development and metabolism. Several molecular changes have been described associated with imprinting disorders: cytogenetic rearrangements, point mutations in imprinted genes, uniparental disomy and epimutations.

There are currently two hypothesis on imprinting methylation disruption; environmental exposures causing primary epimutations, and an initial genetic mutation, in a factor involved in the establishment or maintenance of imprinted methylation, resulting in secondary epimutations.

METHODS AND RESULTS

Previous studies performed at the Molecular (Epi)Genetics Laboratory, Araba-Txagorritxu University Hospital identified 10 patients with pseudohypoparathyroidism, 9 of them carrying a methylation alteration on the GNAS locus, with all four DMRs affected and no deletions in the NESP55 gene, and 1 patient with partially altered methylation of the GNAS locus. 33 other patients were diagnosed with imprinting disorders, carrying methylation alterations in the regions associated to their syndrome.

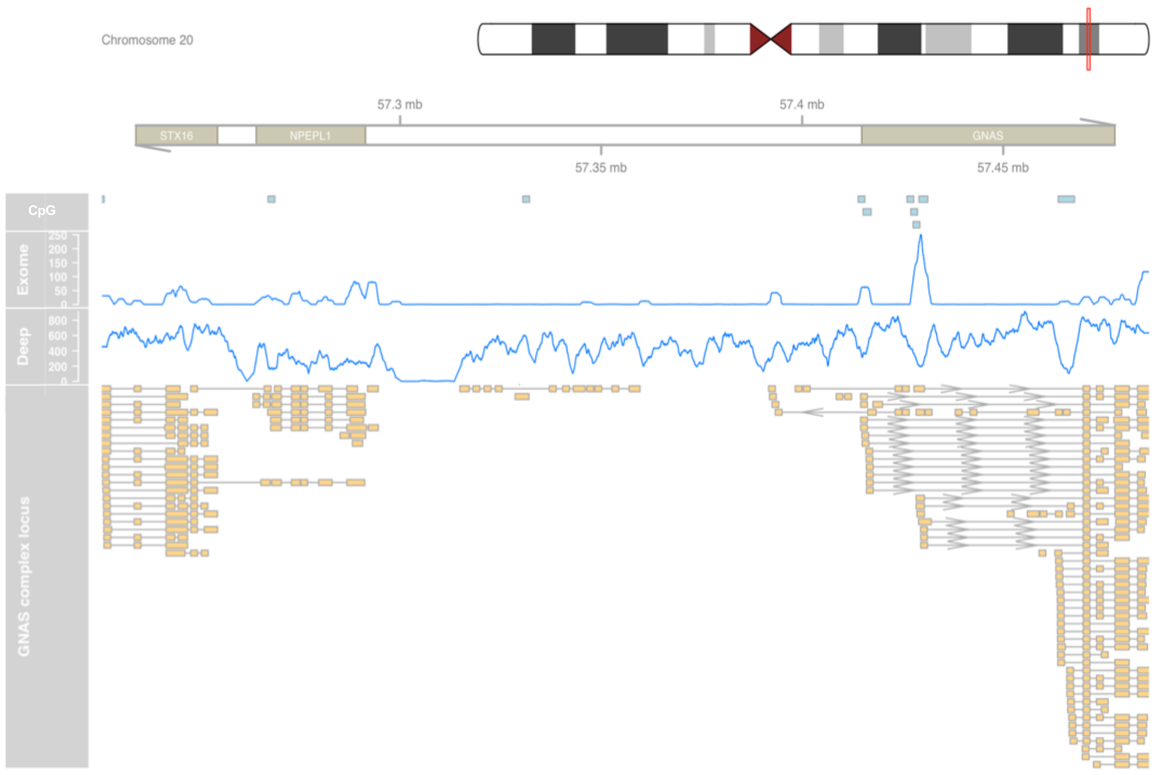

Figure 1. Representation of the GNAS locus; imprinted region affected by aberrant methylation patterns, showing *STX16*, *NESP55* and *GNAS* genes, and coverage differences between Exome and Deep-Sequencing.

DEEP-SEQUENCING STUDY

Deep-Sequencing was performed on 33 patients using custom enrichment capture of imprinting genomic regions “SureSelect DNA” (Agilent©). On a HiSeq2000 system (Illumina Inc.)

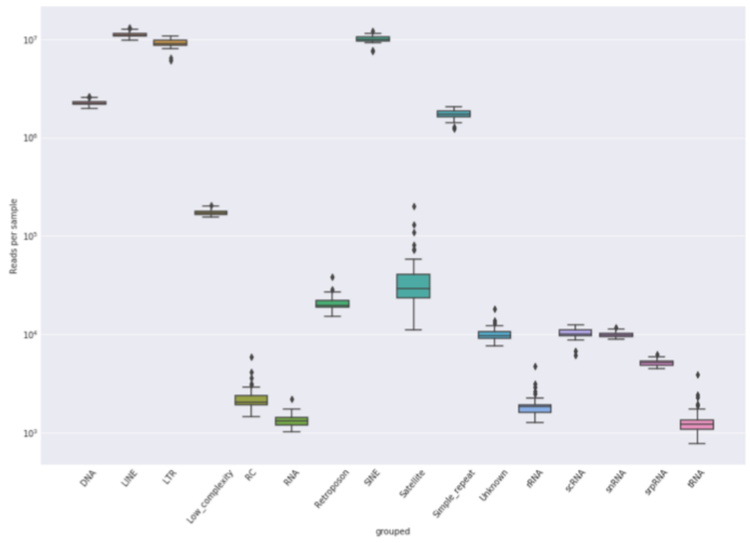

Figure 4. Number of reads mapping to repetitive elements interspersed throughout the genome.

OBJECTIVES

So far mutations in four genes have been associated with imprinting disorders caused by multilocus imprinting disturbances: *NLRP7* (MIM #609661), *KHDC3L* (MIM #611687), *NLRP2* (MIM #609364) and *ZFP57* (MIM #612192). Whereas microdeletions and point mutations have been detected at the ICRs regulating locus-specific DMRs.

- Short term objectives: developing specific bioinformatic pipelines for pedigree analysis and causal variant identification. Developing a tool for Deep-Sequencing data processing and identification of possible methylation modulators.
- Long-term objective: identifying unknown genes associated with disruption of imprinting patterns (secondary epimutations).

EXOME STUDY

Exome sequencing for pedigree analysis was performed on 41 individuals (6 trios, 6 single, 3 quartets and 1 quintet) on a HiSeq2000 system (Illumina Inc.) using exome enrichment capture ”SureSelect Target Enrichment System Human all exon V5 + UTRs” kit (Agilent©).

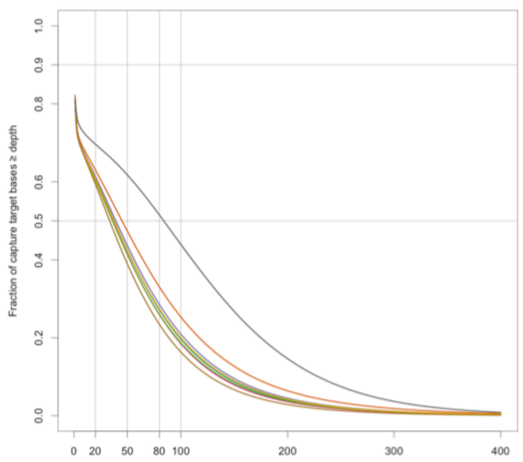

Figure 2. Coverage along the exome target regions; cumulative coverage distribution by fraction of targeted bases covered at different read depths.

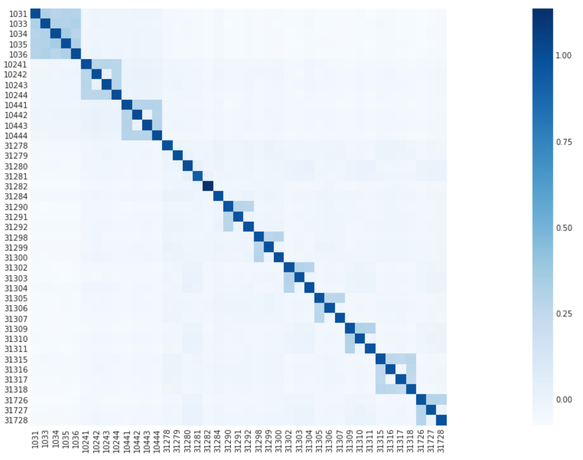

Figure 3. Relatedness between samples; pairwise genetic relationship calculated using autosomal markers.

HEREDITARY VARIANTS ANALYSIS

Variant prioritization on exome samples was performed using GEMINI, a framework for integrating variant, genotype, phenotype and annotation information into a SQL database.

- Impact severity of variants == HIGH
- Minor allele frequency of the variant in EUR population of 1000g project < 0.1
- CADD score for variant pathogenicity > 20

Table 1. Candidate variants obtained through exome analysis by means of inheritance pattern.

Impact	Compound Het	Recessive	De Novo
Splice_acceptor_variant	5	1	198
Splice_donnor_variant	2	0	159
Start_lost	0	0	33
Stop_gained	17	4	510
Stop_lost	0	0	4
Total	24	5	904

CONCLUSIONS

- We are developing a tool that performs reproducible and automated genetic inheritance analysis in relation to hereditary diseases and imprinting disorders. We have tested its performance using previous familiar studies (hypobetalipoproteinemia) and have successfully identified the causal variant.
- Joint analysis of all individuals in the study, permits identification of parent-of-origin variants and segregation analysis, as well as subgroup comparison using known phenotypic or genomic characteristics.
- Deep-Sequencing approaches facilitate variant or copy number alteration identification, in heterozygosis or mosaic that may cause losses in methylation. Development of specific analysis tools is indispensable for in depth study of intron and intergenic sequences.

[1] Bystry,B. et al. (2015) ARResT/AssignSubsets: a novel application for robust subclassification of chronic lymphocytic leukemia based on B cell receptor IG stereotypy. Bioinformatics, 31, 3844-6. [2] Kim,J. et al. (2016) Genomic profile of Chronic Lymphocytic Leukemia in Korea Identified by Targeted Sequencing. Plos One, 11, e0167641. [3] Nadeu,F. et al. (2018) Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. Leukemia, 32, 645-653. [4] Jay,JJ. et al. (2016) Lollyplots in the Clinic: Information Dense Mutation Plots for Precision Medicine. PLoS ONE, 11, e0160519