HUMAN GENOTYPING-*CEGEN* UNIT

Anna González Neira
Unit Head

Graduate Students
Hugo Tejera, Alejandro Velasco (since October)

Technicians
Charo Alonso, Núria Álvarez, Belén Herráez, Rocío Nuñez (TS)*, Guillermo Pita (TS) *

Student in Practice
María Rodrigo (*Universidad Complutense de Madrid*)

Titulado Superior (Advanced Degree)

OVERVIEW

The most abundant types of genetic variation are single nucleotide variants (SNVs) and copy number variants (CNVs). Association studies involving the large-scale analysis of both SNVs and CNVs in thousands of patients can help to identify genes underlying complex diseases such as cancer and drug responses. In this Unit we implement different high-throughput and cost-effective methods to measure from one to millions of SNVs and CNVs. In addition, epigenetic studies using whole-genome methylation arrays are performed in this Unit. Complementarily, research focused on the identification of predictive biomarkers for precision medicine is also undertaken.

"Matching cancer patients with treatments that are likely to be more effective and cause fewer side effects is what we strive for."

RESEARCH HIGHLIGHTS

Pharmacogenetic variants and response to neoadjuvant single-agent doxorubicin or docetaxel: a study in locally advanced breast cancer patients participating in the NCT00123929 phase 2 randomised trial. Docetaxel and anthracycline are widely used in the treatment of breast cancer despite the benefit being limited to a small proportion of patients, and preoperative biomarkers predictive of clinical outcome remain lacking. We carried out a pharmacogenetic study in 181 patients with locally advanced breast cancer who were previously enrolled in a phase 2 randomised clinical trial (NCT00123929), in which patients were randomly assigned to receive doxorubicin (anthracycline) or docetaxel (taxane) in neoadjuvance. We assessed whether genetic variants in 15 key transport or metabolism genes relevant to doxorubicin and docetaxel drugs could play a role as predictive biomarkers. We identified a genetic variant, located in the promoter of ABCC2, as having the strongest association with tumour response observed in patients treated with doxorubicin (P=0.009). We also identified a significant association for an intronic variant, located in CYP1B1, associated with docetaxel tumour response (P=2.15x10-4). Our integrated pathway-based approach allows revealing promising genetic biomarkers for treatment outcome in breast cancer patients (Ruiz-Pinto S *et al.*, 2018).

Genome-wide association study (GWAS) identifies three new loci associated with Ewing sarcoma susceptibility. Ewing sarcoma (EWS) is a paediatric cancer characterised by the EWSR1-FLI1 fusion. Our previous GWAS identified susceptibility loci at 1p36.22, 10q21 and 15q15. We performed a GWAS of 733 EWS cases and 1346 unaffected individuals of European ancestry. Our study replicates previously reported susceptibility loci at 1p36.22, 10q21.3 and 15q15.1, and identifies new loci at 6p25.1, 20p11.22 and 20p11.23. In the analyses of the new loci, there is evidence of informative eQTLs with nearby biologically plausible candidate genes that could be likely target genes for future functional investigations. It is remarkable that 6 independent susceptibility regions with relatively large effect sizes (estimated OR > 1.7) have been discovered in a sample of 733 EWS cases. In conclusion, our study provides support for a strong inherited genetic component to EWS risk and suggests that interactions between germline variation and somatically acquired EWSR1-FLI1 translocations are important etiologic contributors to EWS risk (Machiela MJ *et al.*, 2018).

New loci associated with risk to develop tobacco-induced lung cancer: genome-wide association study in heavy smokers. We genotyped 2.37 million SNPs across the genome in heavy smokers that either developed NSCLC at an early age (extreme cases), or did not present NSCLC at an advanced age (extreme controls), selected from a discovery set (n = 3631). We validated significant SNPs in 133 additional subjects with extreme phenotypes selected from databases including >39,000 individuals. Two SNPs were validated: rs12660420 (p combined = 5.66 × 10-5; OR combined = 2.80), mapping to a noncoding transcript exon of PDE10A; and rs6835978 (p combined = 1.02 × 10-4; OR combined = 2.57), an intronic variant in ATP10D. We assessed the relevance of both proteins in early-stage NSCLC. PDE10A and ATP10D mRNA expressions correlated with survival in 821 stage I-II NSCLC patients (p = 0.01 and p < 0.0001). PDE10A protein expression correlated with survival in 149 patients with stage I-II NSCLC (p = 0.002). In conclusion, we validated 2 novel variants associated with risk of developing tobacco-induced NSCLC in heavy smokers (Fusco JP *et al.*, 2018). ■

PUBLICATIONS

- Wu L *et al.* (incl. Benitez J, González-Neira A) (2018). A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. *Nat Genet* 50, 968-978.
- Ghoussaini M *et al.* (incl. Benitez J, González-Neira A) (2018). Publisher correction: evidence that breast cancer risk at the 2q35 locus is mediated through *IGFBP5* regulation. *Nat Commun* 9, 16193.
- Machiela MJ *et al.* (incl. González-Neira A) (2018). Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. *Nat Commun* 9, 3184.
- Remacha L *et al.* (2018). Gain-of-function mutations in DNMT3A in patients with paraganglioma. *Genet Med* 20, 1644-1651.
- Mavaddat N *et al.* (incl. González-Neira A) (2018). Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet.* PMID: 30554720.
- Shu X *et al.* (2018). Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis. *Int J Epidemiol.* PMID: 30277539.
- Paumard-Hernández B *et al.* (2018). Whole exome sequencing identifies PLEC, EXO5 and DNAH7 as novel susceptibility genes in testicular cancer. *Int J Cancer* 143, 1954-1962.
- Colombo M *et al.* (incl. Benitez J, González-Neira A) (2018). The BRCA2 c.68-7T > A variant is not pathogenic: a model for clinical calibration of spliceogenicity. *Hum Mutat* 39, 729-741.
- Fusco JP *et al.* (incl. González-Neira A) (2018). Genomic characterization of individuals presenting extreme phenotypes of high and low risk to develop tobacco-induced lung cancer. *Cancer Med.* PMID: 29766673.
- Ruiz-Pinto S *et al.* (incl. Benitez J, González-Neira A) (2018). Pharmacogenetic variants and response to neoadjuvant single-agent doxorubicin or docetaxel: a study in locally advanced breast cancer patients participating in the NCT00123929 phase 2 randomized trial. *Pharmacogenet Genomics* 28, 245-250.