HUMAN GENOTYPING-*CEGEN* UNIT

Anna González-Neira
Unit Head

Gradutate Students
Hugo Tejera, Alejandro Velasco

Bioinformatician
Guillermo Pita (TS)*

Technicians
Charo Alonso, Núria Álvarez, Belén Herráez, Rocío Núñez (TS)*

Student in Practice
Javier Pérez (May.-Dec.) (Master's Programme in Bioinformatics, *ENS-ISCIII*, Madrid, Spain)

**Titulado Superior* (Advanced Degree)

OVERVIEW

In the Unit, we offer researchers access to state-of-the-art methods for high throughput genotyping and sequencing for a wide range of applications. We currently have available different genotyping and sequencing platforms to be used according to the scale of analysis required, and we are continuously developing new techniques to cover all research project needs. The research carried out in the Unit is based on identifying genetic risk factors of breast cancer susceptibility and treatment response. Our main goals are to: i) improve individual breast cancer risk assessment, ii) develop novel strategies for breast cancer early detection, and iii) provide cancer patients more accurate and safe treatment.

"Our research on breast cancer will improve breast cancer risk prediction and guide risk-stratified breast screening strategies."

RESEARCH HIGHLIGHTS

Breast cancer risks associated with missense variants in breast cancer susceptibility genes. This study is the result of the European project BRIDGES (Breast Cancer Risk after Diagnostic Gene Sequencing), in which the Unit participates. Protein truncating variants in ATM, BRCA1, BRCA2, CHEK2, and PALB2 are associated with increased breast cancer risk, but risks associated with missense variants in these genes are uncertain. We analysed 59,639 breast cancer cases and 53,165 controls for missense variants in these 5 breast cancer genes, evaluating the risk according to *in silico* prediction-of-deleteriousness algorithms, functional protein domain, and frequency. For ATM, BRCA1, and BRCA2, data were compatible with small subsets (7%, 2%, and 0.6%, respectively) of rare missense variants giving similar risk to those of protein truncating variants in the same gene. For CHEK2, data were more consistent with a large fraction (approximately 60%) of rare missense variants giving a lower risk [OR 1.75, 95% CI (1.47-2.08)] than CHEK2 protein truncating variants. Our results could contribute to the clinical reporting of gene panel testing for breast cancer susceptibility (Dorling L *et al.* 2022).

Pathology of tumours associated with pathogenic germline variants in 9 breast cancer susceptibility genes. The main objective of this study was to determine the distribution of intrinsic subtypes in the 9 confirmed breast cancer genes — *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *CHEK2*, *PALB2*, *RAD51C*, *RAD51D*, and *TP53* — harbouring rare truncating variants and likely pathogenic missense variants associated with increased breast cancer risk. For this purpose, we used data from the BRIDGES project, including 42,680 patients and 46,387 control participants. The results suggested that variants in the 9 breast cancer risk genes are generally associated with triple-negative and/or high-grade disease. Together, the 9 genes were associated with 27.3% of all triple-negative tumours in women 40 years or younger. (Breast Cancer Association Consortium *et al.* 2022).

Novel genes and sex differences in Covid-19 severity. The study is the result of the Spanish COalition to Unlock Research on host GEnetics on COVID-19 (SCOURGE) consortium, in which the Unit participates. The consortium was launched in May 2020 to find biomarkers of evolution and prognosis that can have an immediate impact on the clinical management and therapeutic decisions in SARS-CoV-2 infections. We conducted a genome-wide study of COVID-19 with patients recruited in Spain from 34 centres in 25 cities. The discovery stage of the study comprised up to 9,371 COVID-19 positive cases and 5,943 population controls. Replication was pursued in an additional 1,598 COVID-19 cases and 1,068 population controls, and in other studies from the Host Genetics Initiative. When we performed sex-disaggregated genome-wide association studies for COVID-19 hospitalisation, genome-wide significance ($P < 5 \times 10^{-8}$) was crossed for variants in 3p21.31 and 21q22.11 loci only among males ($P = 1.3 \times 10^{-22}$ and $P = 8.1 \times 10^{-12}$, respectively), and for variants in 9q21.32 near TLE1 only among females ($P = 4.4 \times 10^{-8}$). The results in the overall analysis revealed 2 novel risk loci in 9p13.3 and 19q13.12, associated with *AQP3* ($P = 2.7 \times 10^{-8}$) and *ARHGAP33* ($P = 1.3 \times 10^{-8}$), respectively. In summary, new candidate variants for COVID-19 severity and evidence supporting genetic disparities among sexes are provided (Cruz R *et al.* 2022). ■

PUBLICATIONS

- Breast Cancer Association Consortium *et al.* (incl. González-Neira A) (2022). Pathology of tumors associated with pathogenic germline variants in 9 breast cancer susceptibility genes. *JAMA Oncol* 8, e216744.
- Dorling L *et al.* (incl. González-Neira A) (2022). Breast cancer risks associated with missense variants in breast cancer susceptibility genes. *Genome Med* 14, 51.
- Grootes I *et al.* (incl. González-Neira A) (2022). Incorporating progesterone receptor expression into the PREDICT breast prognostic model. *Eur J Cancer* 173, 178-193.
- Ahearn TU *et al.* (incl. González-Neira A, Pita G) (2022). Common variants in breast cancer risk loci predispose to distinct tumor subtypes. *Breast Cancer Res* 24, 2.
- Marchena-Perea EM *et al.* (incl. Tejera-Pérez H, Pita G, Núñez-Torres R) (2022). A large case-control study performed in Spanish population suggests that *RECQL5* is the only RECQ helicase involved in breast cancer susceptibility. *Cancers (Basel)* 14, 4738.
- Dennis J *et al.* (incl. González-Neira A) (2022). Rare germline copy number variants (CNVs) and breast cancer risk. *Commun Biol* 5, 65.
- Cruz R *et al.* (incl. Pita G, Alvarez N, González-Neira A) (2022). Novel genes and sex differences in COVID-19 severity. *Hum Mol Genet* 31, 3789-3806.
- Hurkmans EGE *et al.* (incl. González-Neira A) (2022). SLC7A8 coding for LAT2 is associated with early disease progression in osteosarcoma and transports doxorubicin. *Front Pharmacol* 13, 1042989.
- Benito-Sánchez B, *et al.* (incl. Núñez-Torres R, Pita G, González-Neira A) (2022). Apparent regional differences in the spectrum of BARD1 pathogenic variants in Spanish population and importance of copy number variants. *Sci Rep* 12, 8547.
- Santonja Á *et al.* (incl. González-Neira A) (2022). Role of germline variants in the metastasis of breast carcinomas. *Oncotarget* 13, 843-862.
- Vega T *et al.* (incl. González-Neira A) (2022). Genetic, environmental and lifestyle factors associated with longevity. Protocol and response of the LONGECYL Study. *Gac Sanit* 36, 260-264.