The Human Cancer Genetics Programme (HCGP) is a translational research programme working on areas related to genetics, genomics, pharmacogenetics, molecular cytogenetics and the environmental bases of human cancer. The HCGP works in close collaboration with the clinical community.

In 2021, the Human Cancer Genetics Programme was composed of 2 Research Groups: Hereditary Endocrine Cancer and Genetic and Molecular Epidemiology; and 3 Units: Human Genotyping-CEGEN, Molecular Cytogenetics and the Familial Cancer Clinical Unit. In addition, the Programme includes a Familial Cancer Consultancy for the evaluation of families with cancer and the provision of genetic counselling, which is located at the Hospital de Fuenlabrada.

The Programme collaborates closely with the clinical community, not only to foster cooperation in genetic diagnosis but also to promote training and education. In 2021, the Familial Cancer Consultancy at the Hospital de Fuenlabrada attended around 387 consultancies, and the HCGP performed 900 genetic diagnoses and carried out 1,106 cytogenetic studies. In addition, the HCGP hosted 3 medical residents from different Spanish hospitals for 3-month training periods. The Programme also offers professionals and students from different national and international research centres the opportunity to join, either as visitors or for training visits consisting of short-term stays of 1-3 months; a total of 18 national visitors and students were hosted in 2021. In terms of education, 16 national PhD students worked on their research projects, 4 of whom already successfully defended their theses.

One of the main objectives of the Programme is to establish research collaborations with national and international groups; this is well demonstrated by its publication record as well as the key roles held by several of the Programme’s members in consortia and international projects. In this regard, in January 2021, several of the HCGP’s Group members, including our former HCGP Director Javier Benítez, Ana Osorio and Anna González-Neira, participated in a major international study on the inheritance of breast cancer, published in the prestigious *New England Journal of Medicine*.

Other major achievements of the Programme in 2021 include: COST Action CA20122 - Harmonizing clinical care and research on adrenal tumours in European countries, awarded to Mercedes Robledo; Ramón Areces Foundation grant to study the microbiome of high-grade non-muscle invasive bladder cancer, awarded to Nuria Malats and the Weizmann Institute in Israel, in collaboration with Francisco X. Real; 1st prize “SETGyC Paper of the Year 2020-2021” in the Gene Therapy category (Molecular Cytogenetics Unit, *Nat. Commun.*); and the 2021 Molecular Cloud “Infinite Possibilities Teams Award” for genome editing (Molecular Cytogenetics Unit).

Upon the retirement of Javier Benítez in 2020, an international search was launched for a new director to lead a new Cancer Genomics Programme. The new programme will incorporate the Groups and Units from the former programme, together with other Groups that use biocomputational tools for the study of cancer, currently under the Structural Biology Programme. This area has the support of Raúl Rabadán, Professor in the Department of Systems Biology and Biomedical Informatics and Director of the Center for Topology of Cancer Evolution and Heterogeneity of Columbia University (New York, USA). In July 2021, Raúl was appointed as Adjunct Professor at the CNIO, an honorary appointment that allows him to maintain an official link with the Centre, reinforcing the presence of the CNIO in the field of computational biology approaches to cancer research.

Maria A. Blasco, Director
Óscar Fernández-Capetillo, Vice-Director
OVERVIEW

Our Group is mainly interested in identifying genetic risk factors involved in endocrine tumour susceptibility. Through a comprehensive analysis of tumour genomic features, we have been able to propose diagnostic and prognostic markers, to identify altered pathways that could be therapeutically targeted, and to identify new major susceptibility genes.

We are also interested in defining markers associated with differences in anticancer drug response and toxicity. We are applying targeted and whole-exome next-generation sequencing to a large series of clinically well-characterised patients. The aim is to identify new therapeutic approaches to personalise cancer treatment. These efforts will collectively improve the diagnosis, prognosis, and treatment of patients.

“We identified NOP10 as a new metastatic risk marker for paraganglioma, and recommend genetic screening in patients with metastatic non-clear cell renal cancer, as they show high prevalence of pathogenic germline mutations in renal cancer-predisposition genes.”
Analysis of telomere maintenance related genes reveals NOP10 as a new metastatic-risk marker in Phaeochromocytoma/Paraganglioma (PPGL)

One of the main problems we face with PPGL is the lack of molecular markers capable of predicting the development of metastatic patients. Telomere-related genes, such as TERF1 and ATRX, have been recently described as being associated with disease progression in PPGL. However, the contribution of other genes involving the telomere preservation machinery has not been previously investigated. We have analyzed the prognostic value of a comprehensive set of genes involved in telomere maintenance in an ongoing series of genetically characterized PPGL samples. In addition to TERF1 and ATRX, NOP10 showed differential expression between metastatic and non-metastatic cases, and alterations in these genes were associated with a shorter time to progression. Analysis of telomere length by Q-FISH in patients, as an in vitro model, showed that NOP10 overexpression is linked to an intermediate-length telomere phenotype without ALT, and in vitro results suggested that NOP10 has a role in telomerase dependent telomere maintenance. Altogether, the results have allowed us to propose NOP10 as a new risk marker for metastatic disease in patients with PPGL, and to propose the use of NOP10 immunohistochemistry as a tool for easy implementation in the clinical setting to stratify patients according to their risk at the time of diagnosis (see FIGURE).

Prevalence of pathogenic germline variants in patients with metastopic potential

Up to 20 genes have been described to be involved in PPGL susceptibility so far, highlighting the remarkable diversity of biological pathways influencing the development of these tumours. Recently our group identified DLST, a gene encoding one of the OGDH complex subunits of the TCA cycle, as a new susceptibility gene for PPGL. Though TCA cycle enzymes are common targets of PPGL predisposing mutations, it seems that the mechanism leading to tumorigenesis in DLST-mutated PPGLs is different from what has been previously described. We used a cell model to examine the potential consequences of mutated DLST in terms of its cellular location, function and affected molecular pathways. We found a significant decrease in the overall protein succinylation levels in the presence of DLST alterations. This post-translational modification (PTM) provides significant chemical and structural changes to proteins, most likely influencing their function. Accordingly, we found that the dysregulation of succinylation seems to have an impact in several essential pathways within cellular metabolism. These results suggest a key role of DLST in protein succinylation and support the relevance of this PTM in the development of different types of cancer.


OVERVIEW

The scope of the research carried out by the Genetic and Molecular Epidemiology Group (GMEG) ranges from the identification of aetiological agents and genetic pathways to the translation of the findings into the clinical and public health domains, focusing on bladder, pancreatic, and breast cancers.

We employ a wide variety of biomarkers, including omics data, to better characterise exposures, genetic susceptibility patterns, and cancer outcomes. While omics data provide a unique opportunity in this regard, their integration with non-omics data poses important challenges, and GMEG explores this methodological field in epidemiologic studies.

The strategic goals of the Group are to:

→ Identify non-genetic and genetic factors, as well as their interactions, associated with cancer development and progression and with its molecular/omics subphenotypes.
→ Develop and apply statistical/informatics tools to model the risk and course of patients with cancer by integrating epidemiological and clinical data with omics information.
→ Assess clinical and public health strategies for cancer control using newly developed biomarkers and algorithms.

“Characterising the tumour immune infiltrating repertoire provides insights to understand its immunogenicity and its association with molecular and clinical factors for better patient stratification.”

Núria Malats
Group Leader

M. Evangelina López De Mauroca
Staff Scientist

Raquel Benítez, Claudia Crocìa, Jiangchuan He, Francisco J Junaido,
Graduate Students

Lola Alonso (TS)*, Lidia Estudillo, Vanessa Maneo (TS), Olga Rico (TS), Sergio Sabinro (TS)
Technicians

Alberto Langtry, Víctor M. Sobrino, Nannan Xue (since November)
Post-Doctoral Fellow

Tânia Silva (Advanced Degree)
Student in Practice

Tania Chadha and Daniel de San Sebastián (Sept.-Nov.) (Bachelor’s Degree Final Project, Univ. Politécnica de Madrid, Spain), Dina Echchat
Visiting Master’s Students

Miguel Maquedano (Aug-Dec.) (Univ. Autónoma de Madrid, Spain), Laura Gutiérrez (until May) (Univ. Complutense de Madrid, Spain)
Visiting Scientists

Helena Fidalgo (until May) (CIBER, Madrid, Spain), Isabel A. Martín (Univ. CEU San Pablo, Madrid, Spain), Ashwag M. Mukhtar (since Dec.) (AI Neelain University, Sudan, Africa)
(Science by Women Programme)
Reserach Highlights

In 2021, GEMG continued to contribute to the pancreatic and bladder cancer epidemiological fields. Regarding pancreatic cancer (PC), we aimed to explore the immune repertoire landscape of PC and its association with risk factors and overall survival by extracting the immunoglobulins (IGs) and T cell receptors (TCRs) from the RNA-sequencing of 144 PC from The Cancer Genome Atlas (TCGA) and 180 pancreatic normal tissue from the Genotype-Tissue Expression (GTEX) project. RC presented richer and more diverse IG and TCR infiltration than normal pancreatic tissue. Higher IG infiltration was present in heavy smokers and women, and it was associated with better overall survival (FIGURE 1). In detail, specific IG clonotypes classified samples with better prognosis (FIGURE 1). On the other hand, greater TCR infiltration was present in patients with previous history of diabetes and was associated with lower non-metastatic load. We also characterised the risk pattern of diabetes mellitus and BMI associated with PC using causal inference methods. When exploring the association between PC and gallbladder conditions, we concluded that this relationship may be non-causal and/or largely due to diagnostic attention and/or reverse causation. In the genetic susceptibility field, an integrated GWAS approach allowed us to identify further 57 variants associated with PC. We also participated in international consortia to identify loci increasing the risk of PC among smokers and the association of hepcidin-regulating iron metabolism genes and genetic variants affecting microRNAs with PC. In the clinical field, we were involved in 2 studies on liquid biopsy revealing that KRAS negativity and longer overall survival were associated with hereditary/familial PC and identifying a plasma protein biomarker panel for early detection of sporadic PC. Regarding bladder cancer (BC), in 2021 GEMG participated in several European projects in the clinical domain that have assessed the relationship between patient sex and treatment outcomes after intravesical BCG (bacillus Calmette-Guérin) immunotherapy for T1G3 HG bladder cancer and identified the risk factors for residual disease at re-TUR (repeat transurethral resection) in T1G3 patients and novel genes associated with recurrence and progression in non-muscle-invasive bladder cancer. Furthermore, we have been involved in a multi-omics analysis to identify prognostic molecular subtypes of non-muscle-invasive bladder cancer.

Methodological Contributions

GEMG also continued to explore the analytic strategies and tools to integrate omics and non-omics data into the cancer risk models, and made progress integrating medical image information (radiomics and digital pathology).

Translational Activities

GEMG actively supports several clinical trials on immunotherapy in BC at the methodological level. We continue to support the Spanish Familial PC Registry (PanGen-FAM) and the European Registry of PC (PancreoER). We are involved in the PC research platform (PC-CAM) to accelerate the translation of research results into the clinical domain. We lead the Research Work Stream of the Pancreatic Cancer Europe (EPC) multi-stakeholder platform, and we have moved ahead in increasing awareness about PC. We contributed to the UEG position paper on PC published in 2021. In addition, we joined an initiative of the European Alliance for Personalised Medicine to bring onco-innovation to Europe’s healthcare systems by reviewing the potential of biomarker testing in the real world to empower personalised medicine.

*PUBLICATIONS*

- Molina-Montes E, Dos Santos J, de la Pena-Campos M, Márquez M (incl. Real FX, Malats N) on behalf of the PanGenEU Study Inves-
  tigators (2021). Deriving the complex imaging phenotypic landscape of PC and its association with risk factors and overall survival by extracting the immunoglobulins (IGs) and T cell receptors (TCRs) from the RNA-sequencing of 144 PC from The Cancer Genome Atlas (TCGA) and 180 pancreatic normal tissue from the Genotype-Tissue Expression (GTEX) project. RC presented richer and more diverse IG and TCR infiltration than normal pancreatic tissue. Higher IG infiltration was present in heavy smokers and women, and it was associated with better overall survival (FIGURE 1). In detail, specific IG clonotypes classified samples with better prognosis (FIGURE 1). On the other hand, greater TCR infiltration was present in patients with previous history of diabetes and was associated with lower non-metastatic load. We also characterised the risk pattern of diabetes mellitus and BMI associated with PC using causal inference methods. When exploring the association between PC and gallbladder conditions, we concluded that this relationship may be non-causal and/or largely due to diagnostic attention and/or reverse causation. In the genetic susceptibility field, an integrated GWAS approach allowed us to identify further 57 variants associated with PC. We also participated in international consortia to identify loci increasing the risk of PC among smokers and the association of hepcidin-regulating iron metabolism genes and genetic variants affecting microRNAs with PC. In the clinical field, we were involved in 2 studies on liquid biopsy revealing that KRAS negativity and longer overall survival were associated with hereditary/familial PC and identifying a plasma protein biomarker panel for early detection of sporadic PC. Regarding bladder cancer (BC), in 2021 GEMG participated in several European projects in the clinical domain that have assessed the relationship between patient sex and treatment outcomes after intravesical BCG (bacillus Calmette-Guérin) immunotherapy for T1G3 HG bladder cancer and identified the risk factors for residual disease at re-TUR (repeat transurethral resection) in T1G3 patients and novel genes associated with recurrence and progression in non-muscle-invasive bladder cancer. Furthermore, we have been involved in a multi-omics analysis to identify prognostic molecular subtypes of non-muscle-invasive bladder cancer.

*Methodological Contributions*

GEMG also continued to explore the analytic strategies and tools to integrate omics and non-omics data into the cancer risk models, and made progress integrating medical image information (radiomics and digital pathology).

*Translational Activities*

GEMG actively supports several clinical trials on immunotherapy in BC at the methodological level. We continue to support the Spanish Familial PC Registry (PanGen-FAM) and the European Registry of PC (PancreoER). We are involved in the PC research platform (PC-CAM) to accelerate the translation of research results into the clinical domain. We lead the Research Work Stream of the Pancreatic Cancer Europe (EPC) multi-stakeholder platform, and we have moved ahead in increasing awareness about PC. We contributed to the UEG position paper on PC published in 2021. In addition, we joined an initiative of the European Alliance for Personalised Medicine to bring onco-innovation to Europe’s healthcare systems by reviewing the potential of biomarker testing in the real world to empower personalised medicine.

**FIGURE 1** Linear regression model showing the association between IG entropy and smoking (A). Biplot showing the association between TCR and diabetes and the association between IG expression and sex (B). Pearson correlation showing the association between the TCR and IG

**FIGURE 2** Heatmap showing the IG clonotypes selected to discriminate PDAC (TCGA) vs. Normal Pancreas (SIGs) (A). Bi-plot showing the 2 first principal component analyses using the 24 clonotypes selected and applied to both
discovery and validation datasets (B). Overall survival Kaplan–Meier curves with the 3 clusters obtained and the biopsy with the tumour purity (C). Comparison between the 3 clusters with the cell content deconvoluted using cellx (D).
Mosaic variated aneuploidy (MVA). Using WES we found biallelic loss-of-function MAD1 mutations in a 30-year-old female with multiple cancers, together with a constellation of benign neoplasia derived from the 3 embryonic layers, and developmental defects. The patient showed the highest rate of variated aneuploidy associated with mutations in any gene involved in SAC (FIGURE 1) and MAD1 protein absence, a circumstance that is in principle incompatible with embryonic development. In collaboration with other CNIO groups, we are investigating the relationship between aneuploidy and tumorigenesis, trying to identify a second mechanism — a dysfunction in APC/C or in another biological process involved in ensuring chromosome segregation — that could contribute to cell viability or evasion of embryonic lethality.

Deciphering the role of rare variants in Breast Cancer (BC). Earlier in 2021, the 2 largest BC case-control studies published so far confirmed 11 genes associated with BC susceptibility and that therefore should be included in diagnostic testing. All these genes were already considered clinically relevant, except BARD1, whose role in BC susceptibility had not been clearly established until now. Based on these results, we included BARD1 in our diagnostic panel, and we analysed the gene in a large series of 2000 BRCA1 Spanish families. We found a strikingly high proportion of large deletions in the gene that suggest the existence of regional differences in the spectrum of mutations in the Spanish population.

Identification of new BC susceptibility genes. We previously identified RECQL5, a member of the RECQL helicase family, as a new BC susceptibility candidate. More recently, we identified 13 additional candidate genes that are currently being analysed by NGS to evaluate their role in the disease in a large set of 3000 Spanish BC families.

DNA glycosylase inhibitors as a new therapeutic approach in hereditary BC patients. We are exploring the possible therapeutic use of OGG1 glycosylase inhibitors in BC patients. We found that the inactivation of BER by the OGG1 inhibitor THS487 increases the accumulation of oxidised bases at the telomeres, leading to telomere loss and post-mitotic defects. Moreover, we discovered that THS487 enhances the activity of the PARPi inhibitor olaparib in BRCA1 deficient cells.

Clinical and diagnostic activity. During 2021, our consultancy at the Fuenlabrada Hospital was visited by 397 patients and 900 genetic studies were performed in the laboratory.
**MOLECULAR CYTOGENETICS UNIT**

Mariana Barea (since May), Pilar Puig | (TS) | Alejandro Alfonso (since September) | (M. Carmen Martín, Marta Puig, Maria Cruz Casado (since May), Pilar Puig) | 1st prize “Paper of the Year 2020–2021” contest in the gene therapy category for the article: “In vivo analysis of fusion oncoproteins for selective elimination of cancer cells” (Nature Communications, 2021). Bestowed by The Spanish Society for Gene and Cell Therapies (ASEG).
**Overview**

In the Unit we implement high-throughput and cost-effective methods to measure from one to millions of genetic variants, mainly single nucleotide variants (SNVs) and copy number variants (CNVs). Epigenetic analyses are also performed. Complementarily, our research focuses on identifying genes associated with cancer risk and response to therapy to understand the underlying molecular mechanisms of cancer susceptibility and drug efficacy/toxicity and to improve individual risk assessment for the identification of high-risk populations.

“Our goal is to identify predictive markers in cancer that allow individual risk assessment, thus integrating personalised medicine into healthcare practice.”

**Research Highlights**

Breast cancer risk genes – association analysis in more than 113,000 women. This study is the result of the European project BRIDGES (Breast Cancer Risk after Diagnostic Gene Sequencing), in which the Unit participates. The study was based on the analysis of 34 known or suspected breast cancer susceptibility genes in DNA samples from 60,400 women who had developed breast cancer and 53,400 healthy women. The results pinpoint 9 genes for which there is solid evidence of their involvement in the disease: ATM, BRCA1, BRCA2, CHK2, PALB2, BARD1, RAD51C, RAD51D and TP53. This was already known for some of these genes, but for others, such as RAD51C and D and BARD1, their involvement was not so well established. For all genes, more precise risk estimates can now be calculated, and these estimates are tailored for each tumour subtype. By contrast, the study shows that about 15 of the genes that have been used so far in some tests are not indicative of an increased risk for breast cancer and should therefore not be taken into account in risk estimates, at least at this time (Dorling L et al., 2021).

Characterisation of pharmacogenetic variability in the Spanish population. Pharmacogenomics (PGx) allows for patients to be managed in an individualised manner, leading to the better safety and efficacy of treatments. Nevertheless, genetic differences between populations need to be considered before implementing PGx. We used data from 30,066 Spanish individuals to analyse pharmacogenetic variation in 1055 important pharmacogenes for population characterisation. For 21 clinically relevant pharmacogenes, our analysis revealed that 98% of the Spanish individuals harbour at least 1 allele that leads to a change in the therapy. We also identified 7775 putative pathogenic SNVs and 33 CNVs. This study provided useful information to facilitate PGx implementation in our country by (i) confirming that almost all Spanish individuals could benefit from pharmacogenetics diagnostics; (ii) identifying additional pathogenic pharmacovariants with a possible functional role, data freely available to the public by accessing the Collaborative Spanish Variant Server; and (iii) demonstrating that PGx microarrays can be a cost-effective solution for testing.

**PORMT** as a novel susceptibility gene for cardiotoxicity in epirubicin treatment of breast cancer patients. Anthracyclines are among the most used chemotherapeutic agents in breast cancer (BC). However, their use is hampered by anthracycline-induced cardiotoxicity (AIC). To identify novel predictive genes, we conducted a 2-stage genome-wide association study in epirubicin-treated BC patients. The most interesting and replicated finding was rs63342666, located 4kb upstream of PORMT (OR = 5.76, P = 2.23 × 10⁻⁴). This variant regulates the expression of PORMT, a gene that encodes a mitochondrial DNA-directed RNA polymerase, responsible for mitochondrial gene expression. Individuals harbouring the risk allele had decreased expression of PORMT in heart tissue, which may cause impaired capacity to maintain a healthy mitochondrial population in cardiomyocytes under stress conditions, such as is the case with epirubicin treatment. This finding suggests a novel molecular mechanism involved in the development of AIC and may improve our ability to predict which patients are at risk (Velasco-Buitz A et al., 2021).

**Publications**