GENETIC AND MOLECULAR EPIDEMIOLOGY GROUP

Group Leader
Núria Malats

Research Scientists
M. Evangelina López de Maturana, Bárbara Oldrini (until July)

Gradutate Students
Raquel Benítez, Claudia Coscia (until April), Jiangchuan He (until December), Francisco J Jurado, Alberto Langtry, Victor M. Sobrino, Casper Van Eijk (since August), Nannan Xue

Technicians
Lola Alonso (TS)*, Lidia Estudillo, Yaiza Gutiérrez (since April), Ane Moreno (since November), Vanesa Moreno (TS)*, Sergio Sabroso (TS)*, María Olano (since April)

**Titulado Superior* (Advanced Degree)

Students in Practice
Sol Aletta (since Aug) (Rabdoud University, Nijmegen, The Netherlands), Tania Chadha (March-July) (*ETSIT*, Madrid, Spain), Daniel de San Sebastián (until July) (*UPM*, Madrid, Spain), Soraya Martín (April-June) (*IFP Corredor del Henares*, Spain), Harold Mena (Aug.-Oct.) (Del Rosario Univ., Bogotá, Colombia), Anaëlle Mescam (June-Aug.) (École de Biologie Industrielle, Cergy, France), Nataly Moreu (March-June) (*IES* Mirasierra, Spain)

Visiting Master's Students
Miguel Maquedano (until July) (*UAM*, Madrid, Spain), César Mediavilla (May.-Dec.) and Juana Serrano (since Aug.) (Master's in Bioinformatics, ISCIII-ENS, Madrid, Spain)

Visiting Scientists
Isabel A. Martín (*Univ. CEU San Pablo*, Madrid, Spain), Esther Molina (July-Sep) (*Univ. de Vic*, Barcelona, Spain), Ashwag M. Mukhtar (until May) (Al Neelain University, Sudan, Africa) (Science by Women Programme)

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.

"Oral, faecal, and pancreatic microbiome dysbiosis are associated with pancreatic cancer, with stool microbiota-based classifiers that predict pancreatic cancer with high accuracy and specificity."

RESEARCH HIGHLIGHTS

Research findings

In 2022, GMEG contributed to the pancreatic cancer (PC) field by proposing a faecal metagenomic classifier that identifies PC with an accuracy of 0.84 area under the receiver operating characteristic curve (AUROC) in a Spanish cohort, based on 27 species. The accuracy improves up to 0.94 AUROC when combined with CA19–9 serum marker. The classifier was validated in an independent German PC cohort (0.83 AUROC), and PC disease specificity was confirmed against 25 publicly available metagenomic study populations with various health conditions (n=5792). The presence of marker taxa enriched in faecal samples (*Veillonella*, *Streptococcus*, *Akkermansia*) and also taxa with differential abundance in healthy and tumour pancreatic tissues (*Bacteroides*, *Lactobacillus*, *Bifidobacterium*) was validated by fluorescence *in situ* hybridisation (FIGURE 1). The presented PDAC-specific microbiome signatures, including links between microbial populations across tissues, provide novel microbiome-related hypotheses regarding disease aetiology, prevention, and possible therapeutic intervention. In addition, we also collaborated in elucidating that *GATA4* and *GATA6* cooperate to maintain the classical PC phenotype. We further explored the immune repertoire landscape of 9522 tumour and adjacent non-tumour samples across 28 tumour types in the Cancer Genome Atlas project, and performed diversity and network analysis. We identified differences in diversity and network statistics across tumour types and subtypes and observed a trend towards increased clonality in primary tumours compared to adjacent non-tumour tissues. Regarding **bladder cancer** (BC), GMEG participated in 1 study that delivered suggestive evidence for a multiplicative interaction between the most common class of disinfection by-products, trihalomethanes, and a bladder cancer susceptibility variant (rs907611). Furthermore, we contributed to the validation of BlaDimiR, a urine-based miRNA score for accurate bladder cancer diagnosis and follow-up.

Methodological contributions

We proposed an approach allowing Mendelian randomisation estimation in strata of the population while avoiding collider bias (FIGURE 2). This approach constructs a new variable, the residual collider, as the residual from regression of the collider on the genetic instrument, and then calculates causal estimates in strata defined by quantiles of the residual collider. The new approach generated unbiased estimates in all the simulation settings, and can be used to perform Mendelian randomisation studying heterogeneity among subgroups of the population while avoiding collider bias. Furthermore, GMEG continued exploring the analytic strategies and tools to integrate omics and non-omics data into the cancer risk models, and made progress in the integration of medical image information (radiomics and digital pathology).

Translational activities

GMEG actively supports several clinical trials of immunotherapy in BC at the methodological level. We continue to sustain the Spanish Familial PC Registry (PanGen-FAM) and the European Registry of PC (PancreOS). We chair the Spanish Alliance for Pancreatic Cancer Research (ALIPANC) to accelerate the translation of research results into the clinical and public health domains. We lead the Research Work Stream of the Pancreatic Cancer Europe (PCE) multistakeholder platform, and we have moved ahead in increasing awareness of PC. We also contributed to the publication of the UEG position paper on pancreatic cancer. Finally, we joined an initiative of the European Alliance for Personalised Medicine to express concerns that disrupting the current balance of the pharmaceutical legislation to meet objectives that are more precisely targeted could have unintended consequences in the EU, reducing rather than increasing the flow of innovative treatments for rare diseases. ■

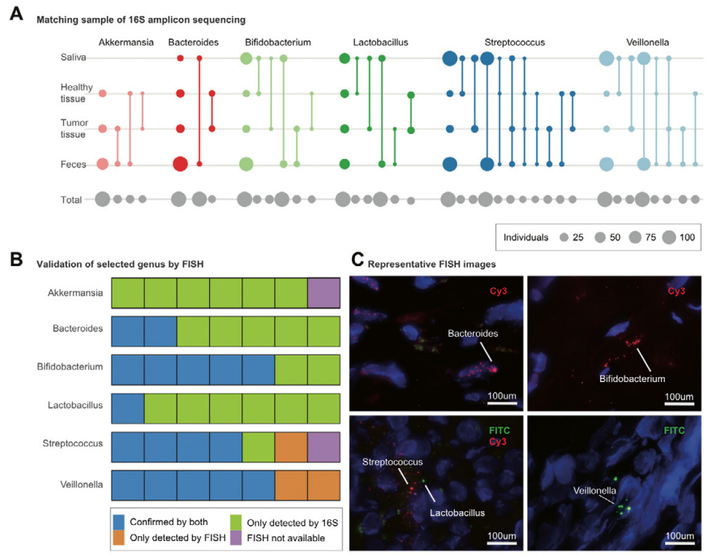

FIGURE 1 Presence of bacteria in 4 different body sites including faecal, saliva, pancreatic tumour and healthy tissue samples (**A**). Bacterial presence/absence with both 16S amplicon and FISH methods in 7 selected pancreatic tissue samples (**B**). FISH microscopy images for *Bacteroides* (intranuclear, tumour tissue), *Bifidobacterium* (extranuclear, tumour tissue), *Lactobacillus* (extranuclear, non-tumour tissue), *Streptococcus* (extranuclear, non-tumour tissue), and *Veillonella* (extranuclear, tumour tissue) (**C**).

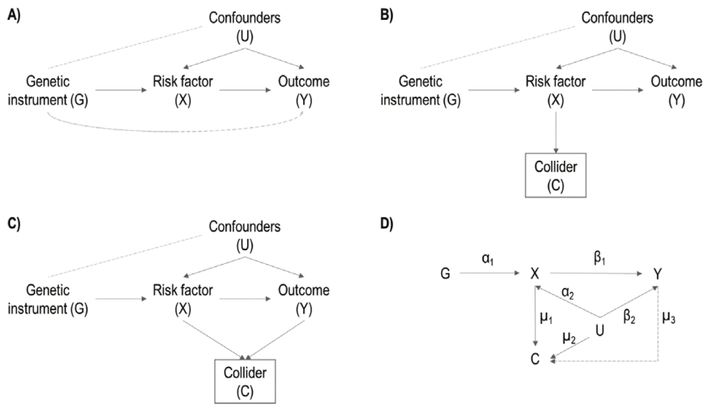

FIGURE 2 Directed Acyclic Graphs (DAGs) illustrating relationships between the variables. (**A**) Mendelian randomisation causal diagram with the instrumental variable assumptions. (**B**) DAG considering a collider variable C, being a common child of risk factor X and outcome Y. (**D**) DAG illustrating the variables and parameters used for the simulation study.

PUBLICATIONS

- de Andrés MP, Jackson RJ, Felipe I, Zagorac S, Pilarsky C, Schlitter AM, Martinez de Villareal J, Jang GH, Costello E, Gallinger S, Ghaneh P, Greenhalf W, Knösel T, Palmer DH, Ruemmele P, Weichert W, Buechler M, Hackert T, Neoptolemos JP, Notta F, Malats N, Martinelli P, Real FX (2022). GATA4 and GATA6 loss-of-expression is associated with extinction of the classical programme and poor outcome in pancreatic ductal adenocarcinoma. *Gut*. PMID: 36109153.
- Kartal E *et al.* (incl. Molina-Montes E, Rodríguez-Perales S, Márquez M, Torres-Ruiz R; MAGIC Study investigators; PanGenEU Study investigators; Real FX, Malats N)P (2022). A faecal microbiota signature with high specificity for pancreatic cancer. *Gut* 71, 1359-1372. **GutTop Paper** 2022.
- Nagata N *et al.* (incl. Molina-Montes E, Estudillo L, Malats N) (2022). Metagenomic identification of microbial signatures predicting pancreatic cancer from a multinational study. *Gastroenterology* 163, 222-238.
- Yuan C, Kim J, Wang QL, Lee AA, Babic A; PanScan/PanC4 I-III Consortium; Amundadottir LT, Klein AP, Li D, McCullough ML, Petersen GM, Risch HA, Stolzenberg-Solomon RZ, Perez K, Ng K, Giovannucci EL, Stampfer MJ, Kraft P, Wolpin BM (2022). The age-dependent association of risk factors with pancreatic cancer. *Ann Oncol* 33, 693-701.
- Suarez-Cabrera C *et al.* (incl. Estudillo L, Real FX, Malats N) (2022). BlaDimiR: a urine-based miRNA score for accurate bladder cancer diagnosis and follow-up. *Eur Urol* 82, 663-667.
- Mouron S, Bueno MJ, Lluch A, Manso L, Calvo I, Cortes J, Garcia-Saenz JA, Gil-Gil M, Martinez-Janez N, Apala JV, Caleiras E, Ximénez-Embún P, Muñoz J, Gonzalez-Cortijo L, Murillo R, Sánchez-Bayona R, Cejalvo JM, Gómez-López G, Fustero-Torre C, Sabroso-Lasa S, Malats N, Martinez M, Moreno A, Megias D, Malumbres M, Colomer R, Quintela-Fandino M (2022). Phosphoproteomic analysis of neoadjuvant breast cancer suggests that increased sensitivity to paclitaxel is driven by CDK4 and filamin A. *Nat Commun* 13, 7529.
- Galesloot TE, Grotenhuis AJ, Kolev D, Aben KK, Bryan RT, Catto JWF, Cheng KK, Conroy S, Dyrskjøt L, Fleshner NE, James ND, Lamy P, Lindskrog SV, Malats N, Mengual L, Verhaegh G, Zeegers MP, Kiemeney LALM, Vermeulen SH (2022). Genome-wide meta-analysis identifies novel genes associated with recurrence and progression in non-muscle-invasive bladder cancer. *Eur Urol Oncol* 5, 70–83.
- Coscia C, Gill D, Benítez R, Pérez T, Malats N, Burgess S (2022). Avoiding collider bias in Mendelian randomization when performing stratified analyses. *Eur J Epidemiol* 37, 671-682.
- Beane Freeman LE *et al.* (incl. Malats N) (2022). Disinfection by-products in drinking water and bladder cancer: evaluation of risk modification by common genetic polymorphisms in two case-control studies. *Environ Health Perspect* 130, 57006.
- Yu K, Ravoor A, Malats N, Pineda S, Sirota M (2022). A pan-cancer analysis of tumor-infiltrating B cell repertoires (2022). *Front Immunol* 12, 790119.
- Daca-Alvarez M, Martí M, Spinelli A, de Miranda NFFC, Palles C, Vivas A, Lachtford A, Monahan K, Szczepkowski M, Tarnowski W, Makkai-Popa ST, Vidal R, López I, Hurtado E, Jiménez F, Jiménez-Toscano M, Álvaro E, Sanz G, Ballestero A, Melone S, Brandáriz L, Prieto I, García-Olmo D, Ocaña T, Moreira R, Moreno L, Carballal S, Moreira L, Pellisé M, González-Sarmiento R, Holowatyj AN, Perea J, Balaguer F; GEOCODE and SECOC consortia (2022). Familial component of early-onset colorectal cancer: opportunity for prevention. *Br J Surg* 109, 1319-1325.
- Serrano MT *et al.* (incl. Sabroso S) (2022). Mortality and causes of death after liver transplantation: analysis of sex differences in a large nationwide cohort. *Transpl Int* 35, 10263.
- Mera-Berriatua L *et al.* (incl. López de Maturana E) (2022). Unravelling the gut microbiota of cow's milk-allergic infants, their mothers, and their grandmothers. *J Investig Allergol Clin Immunol* 32, 395-398.
- Horgan D, Koeva-Balabanova J, Capoluongo E, Jagielska B, Cattaneo I, Kozaric M, Tumiene B, El Ahl JP, Lal JA, Kalra D, Malats N (2022). Making sure that orphan incentives tip the right way in Europe. *Healthcare (Basel)* 10, 1600.

PATENT

- Malats N, Bork P, Kartal E, Molina-Montes E, Rodríguez S, Estudillo L, Real FX, Schmidt S, Zeller G, Wirbel J, Maistrenko OM. Faecal microbiota signature for pancreatic cancer. PCT application (2022). *PCT/EP2022/077087. WO2023052486A1.*