Section of Genetics (GEN)

Section head
Dr Paul Brennan

Genetic Epidemiology Group (GEP)

Group head
Dr Paul Brennan

Scientists
Dr Estelle Chanudet-van den Brink (until July 2020)
Dr Ana Carolina de Carvalho Peters
Dr Aida Ferreiro-Iglesias
Dr Mattias Johansson
Dr Sandra Perdomo Velasquez
Dr Hilary Robbins
Dr Mahdi Sheikh
Dr Shama Virani

Technical assistants
Ms Karine Alcala
Ms Valérie Gaborieau
Ms Sandrine Magat (until January 2020)
Ms Hélène Renard

Laboratory technician
Ms Priscilia Chopard

Project assistant
Ms Laurène Bouvard

Secretariat
Ms Juliette Prazak
Ms Aurélie Rosado (until October 2021)
Ms Andreea Spanu
Ms Charlotte Volatier (until January 2020)

Visiting scientist
Dr Emily Banks (until March 2020)

Postdoctoral fellows
Dr Ricardo Cortez Cardoso Penha
Dr Xiaoshuang Feng
Dr Aida Ferreiro-Iglesias (until July 2020)
Dr Florence Guida (until August 2021)
Dr Jean-Noël Hubert (until December 2020)
Dr Daniela Mariosa
Dr Maja Milojevic
Dr Dariush Nasrollahzadeh Nesheli (until July 2021)
Dr Justina Onwuka
Dr Fadoua Rafii (until September 2021)
Dr Sergey Senkin
Dr Mahdi Sheikh (until June 2021)
Dr Karl Smith-Byrne (until November 2021)

Students
Ms Elmira Ebrahimi
Mr Luis Enrique Romero Cruz
Ms Hana Zahed

Genetic Cancer Susceptibility Group (GCS)

Group head
Dr James McKay

Scientists
Dr Behnoush Abedi-Ardekani
Dr Nicolas Alcala
Dr Lynnette Fernandez-Cuesta
Dr Matthieu Foll
Dr Florence Le Calvez-Kelm

Visiting scientists
Dr Arash Nikmanesh
Dr Dariush Nasrollahzadeh Nesheli

Laboratory technicians
Ms Amélie Chabrier
Ms Nathalie Forey

Bioinformatician
Dr Catherine Voegele

Secretariat
Ms Isabelle Rondy

Postdoctoral fellows
Dr Nicolas Alcala (until March 2020)
Dr Joshua Atkins
Dr Alex Di Genova Bravo
Dr Alexandra Sexton-Oates

Students
Mr Ricardo Blazquez Encinas Rey (until December 2021)
Mr Antoine Boureille (until August 2020)
Ms Virginia Diez-Obrero (until December 2020)
Ms Aurélie Gabriel (until January 2021)
Mr Colin Giacobi (until August 2020)
Mr Boris Lipinski (until August 2020)
Ms Lise Mangiante
Ms Emilie Mathian
Ms Mathilde Persch (until September 2021)

The Section of Genetics (GEN) includes the Genetic Epidemiology Group (GEP) and the Genetic Cancer Susceptibility Group (GCS). The work of the Section combines large population-based studies as well as laboratory and bioinformatics expertise to identify specific genes and genetic profiles that contribute to the development of cancer and to elucidate how they exert their effect along with environmental factors. GEN also tries to identify individuals who are at high enough risk that they are likely to benefit from potential screening strategies.

GEN projects usually involve extensive fieldwork in collaboration with external investigators to develop large-scale epidemiological studies with appropriate clinical and exposure data, as well as biosample collection. This typically occurs within GEP. Germline genetic analysis usually comprises genome-wide genotyping studies, as well as extensive sequencing work. GEP studies also assess non-genetic exposures, partly in recognition of the importance of non-genetic factors in driving cancer incidence, and also to facilitate accurate assessment of gene–environment interactions. In contrast, GCS places more focus on identification of uncommon or rare genetic variants that may have a larger effect than common single-nucleotide polymorphisms but that are not sufficiently frequent to be captured by current genome-wide association genotyping arrays. The approach of GCS has been to use genomics and bioinformatics techniques to complement more traditional approaches for the study of rare genetic variants. GCS also uses genomics to explore how variants may be conferring genetic susceptibility to cancer. Thus, the research programme of GCS complements that of GEP, and also provides a facility for high-throughput genomics techniques and the related bioinformatics to support GEN's molecular epidemiology projects and other IARC genomics projects.

With the start of the new IARC Medium-Term Strategy 2021–2025 and the new organizational structure as of 1 January 2021, GEN was renamed as the Genomic Epidemiology Branch, to better capture the broad range of scientific activities under way.

Mutographs: Building up a large cancer biorepository across five continents

GEP, in collaboration with GCS and the Section of Environment and Radiation (ENV), the Laboratory Services and Biobank Group (LSB), and the Section of Support to Research (SSR), has devoted substantial resources to recruiting large series of cancer cases, comprising extensive questionnaire information and biological samples, as part of the Mutographs Grand Challenge project. Delays due to the COVID-19 pandemic in 2020 and 2021 had a major impact on recruitment and timelines. However, GEP has worked closely with all centres participating in the project, has adapted the recruitment and processing protocols to comply with COVID-19 restrictions, and continues to support the progress of the project locally.

Ontario Institute for Cancer Research, Toronto, Canada. Courtesy of Pancreatic Cancer Toronto, Canada.

Digestive Diseases Research Institute, Tehran, Islamic Republic of Iran. Courtesy of Digestive Diseases Research Institute, Tehran University of Medical Sciences, Islamic Republic of Iran.

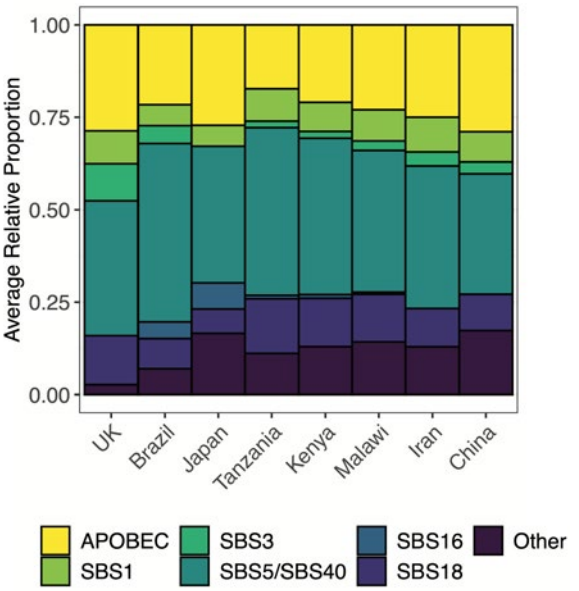
Hospital Italiano de Buenos Aires, Argentina. Courtesy of Hospital Italiano de Buenos Aires, Argentina.

During the 2020–2021 biennium, the Genetic Epidemiology Group (GEP) has continued research efforts on how genomics can be used to understand the causes of cancer, as well as how genomics can contribute to the early detection and outcome prediction of cancer (Ginsburg et al., 2021a). Some prominent examples of the work of the Group over the biennium are described here.

Table 1. Current progress on sample collection and processing by cancer type in the Mutographs project

Sample type	Samples and data at IARC	Samples shipped to Wellcome Sanger Institute	Whole-genome sequencing released	Target	Proportion of target achieved (%)
Oesophageal squamous cell carcinoma	1482	671	552	552	100
Renal cell carcinoma	1261	1141	594	1000	59
Colorectal cancer	1378	386	150	1000	15
Head and neck cancer	622	186	0	300	0
Pancreatic ductal adenocarcinoma	693	297	115	650	18
Oesophageal adenocarcinoma	704	262	110	650	17
Total	6140	2943	1521	3852	39

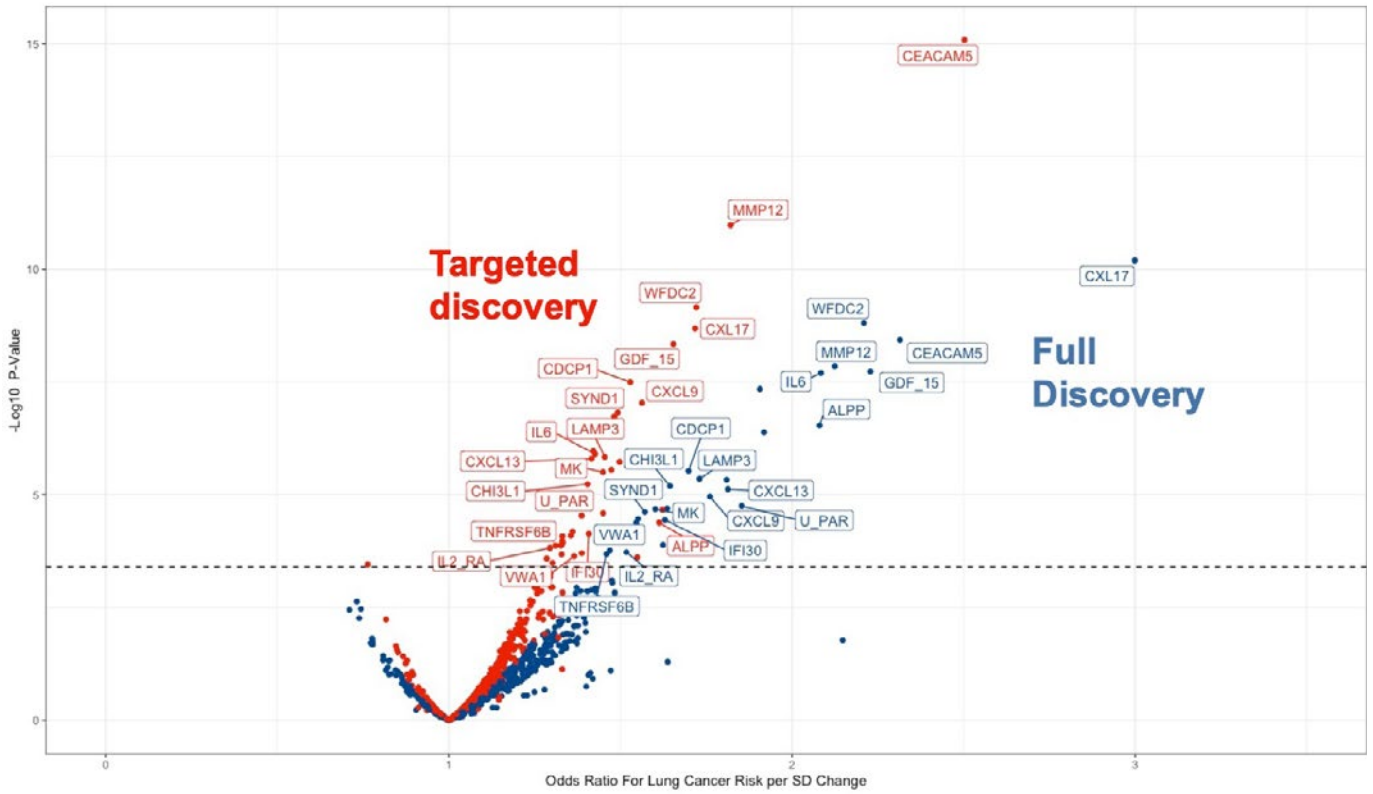
Figure 1. Average relative attributions of single-base substitution (SBS) COSMIC signatures are broadly similar between all countries. Signatures accounting for less than 5% on average (with the exception of SBS3 and SBS16) are grouped together into the "other" category. Reproduced with permission from Moody et al. (2021). © 2021, The Author, under exclusive licence to Springer Nature America, Inc.



Identifying novel causes of multiple cancer types through analysis of mutation signatures: the Mutographs study

The Mutographs project aims to understand the causes of five different cancer types across five continents by generating mutational signature profiles. The initial recruitment of approximately 6000 cases has been completed, and 3000 cases have been successfully processed at IARC and sent to the Wellcome Sanger Institute for whole-genome sequencing (Table 1). GEP completed the first group of 1552 cancers sequenced in 2021, and genomic, exposure, and clinical data will be publicly available through the International Cancer Genome Consortium Accelerating Research in Genomic Oncology (ICGC ARGO) platform. The analysis of 552 cases of oesophageal cancer from eight countries with varying incidence rates showed the high prevalence of APOBEC signatures in all cases, as well as specific mutation signatures linked to opium consumption and alcohol consumption, as well as to homologous DNA repair deficiency (Figure 1) (Moody et al., 2021). Analysis of approximately 1000 kidney cancers is continuing, and preliminary results are shedding light on the contribution of environmental causes to the high risk of kidney cancer in central Europe. Most of the processing and sequencing efforts are now focused on cases of head and neck cancer, colorectal cancer, gastro-oesophageal adenocarcinoma, and pancreatic adenocarcinoma. Based on the same design and methodology as the Mutographs project, three additional side studies investigating specific geographical exposures of interest have been initiated. For example, GEP is studying the plausible sources of aristolochic acid exposure that could cause renal and urinary tract cancers in the Balkan region. The Group also aims to elucidate the role of opium consumption in the development of bladder cancer in the Kerman province of the Islamic Republic of Iran, and to

Figure 2. Volcano plot depicting proteins associated with lung cancer risk after accounting for multiple comparisons in both the initial discovery phase and the replication phase. SD, standard deviation. © IARC.



investigate the mutational profile differences in gallbladder cancer cases from regions of high and low incidence in India.

Evaluating strategies to improve early detection and outcome in lung cancer

The overall goal of GEP is to identify individuals at sufficiently high risk of developing lung cancer to justify screening and early detection. The initial approach of the Group was to develop and validate lung cancer risk prediction models using the vast information harmonized in the Lung Cancer Cohort Consortium (LC3), including lung cancer risk factors and outcomes from more than 20 cohorts around the world, representing more than 2.5 million individuals in 15 countries (Robbins et al., 2021). In addition, using the framework of the Integrative Analysis of Lung Cancer Etiology and Risk (INTEGRAL) study, GEP is actively investigating the potential of a wide range of biomarker types for lung cancer risk prediction; initial studies indicate that circulating protein biomarkers have the most promising potential to improve the identification of individuals most likely to benefit from screening. The Group completed the initial discovery scan of 1200 proteins on 252 case–control pairs from two LC3 cohorts. This enabled the identification of the five most informative protein panels, which were assayed in 477 case–control pairs from four additional cohorts (Figure 2).

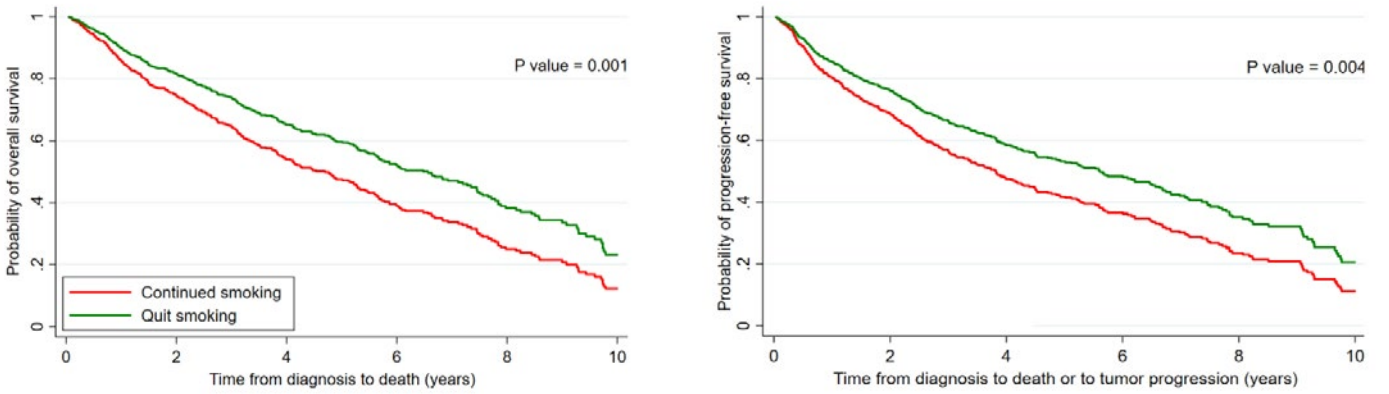
Quitting smoking after diagnosis of lung cancer improves survival

The effect of smoking cessation in lung cancer survival was evaluated using information collected as part of the 15-year collaborative study with the N.N. Blokhin National Medical Research Centre of Oncology of the Russian Academy of Medical Sciences. GEP showed that smoking cessation after lung cancer diagnosis substantially improved overall and progression-free survival among current smokers with early-stage lung cancer; similar effects were observed among mild to moderate smokers and heavy smokers, and among patients with earlier-stage and later-stage tumours (Figure 3) (Sheikh et al., 2021).

Progress investigating etiological and prognostic factors in head and neck cancers

Two main collaborative projects have showed important evidence of the role of human papillomavirus (HPV) infection in the development of oropharyngeal cancer. Recent evidence from the work in the HPV Cancer Cohort Consortium (HPVC3) has shown that HPV16 E6 is easily detectable in blood and is highly sensitive (> 90%) and specific (> 99%), and that seropositivity can occur decades before cancer is detectable. As part of the VOYAGER study, GEP conducted the largest genome-wide association study analysis with a focus on oropharyngeal cancer. After stratifying by HPV status, a protective effect was shown for one human leukocyte antigen

Figure 3. Association between quitting smoking after diagnosis of lung cancer and the probability of overall survival (left) and progression-free survival (right). © IARC.



locus (rs4713462) against HPV-positive oropharyngeal cancer, which also correlated with low antibody levels against HPV16 E6 (Ferreiro-Iglesias et al., 2021). Furthermore, the first survival analysis in 1463 patients with head and neck cancer in South America after completion of a 3-year follow-up confirmed that survival was better for HPV-related than for HPV-unrelated oropharyngeal cancer, as well as the negative prognostic effect of advanced clinical stage and alcohol consumption (Abrahão et al., 2020).

Genetic Cancer Susceptibility Group (GCS)

The Genetic Cancer Susceptibility Group (GCS) is a multidisciplinary scientific group, covering genetics, genomics, bioinformatics, and pathology. These combined skills are used to undertake genetics and genomics research to identify cancer-related genes, explore their mechanisms of action, and contribute to how tumours are classified and detected. GCS works within international consortia to assemble the large sample sizes needed for informative genetics and genomics studies. GCS's multifaceted genomic analysis and multidisciplinary team provide additional depth to these consortia-based studies.

In the context of early biomarkers, GCS has explored highly recurrent telomerase reverse transcriptase (*TERT*) gene promoter mutations as biomarkers for the early detection of urothelial cancer, and developed assays for the detection of low-abundance *TERT* promoter mutations (Zvereva et al., 2020a). This assay was used in a nested case–control study within a longitudinal population-based prospective cohort of 50 000 individuals in the Islamic Republic of Iran, and demonstrated that *TERT* mutations could be detected in urine samples obtained up to 10 years before the primary diagnosis of bladder cancer, and were not detected in matched controls (100% specificity and 46.6% sensitivity) (Figure 4) (Hosen et al., 2020a). The study demonstrated the presence of these mutations in urine in asymptomatic individuals who later developed bladder cancer, highlighting the potential of urinary *TERT* promoter mutations to be used as a simple, inexpensive, and non-invasive early detection biomarker. These results received broad coverage by international media, including the United Nations News, France Info, the La Chaîne Info news channel, and the *Daily Mail* (in the United Kingdom), as well as science magazines and specific urology websites.

The Rare Cancers Genomics initiative aims at the molecular characterization of rare cancers (http://rarecancersgenomics.com/), including malignant pleural mesothelioma (MESOMICS) and lung neuroendocrine neoplasms (lungNENomics). In the MESOMICS project, GCS has contributed to the comprehensive molecular and pathological evaluation of transitional mesothelioma assisted by a deep learning approach (Galateau Salle et al., 2020) and provided an overview of molecular advances in the classification of pleural mesotheliomas (Fernandez-Cuesta et al., 2021). In the lungNENomics project, GCS generated

Figure 4. Association between *TERT* promoter mutation mutant allelic fractions (MAFs, %) and the time interval from urine collection to clinical diagnosis of bladder cancer. The MAFs of the mutations detected by the UroMuTERT assay (next-generation sequencing-based assay) in the 14 urine samples of the asymptomatic subjects from the Golestan Cohort Study are plotted against the time in years from urine collection to diagnosis of bladder cancer. Reproduced from Hosen et al. (2020a). © 2020 Published by Elsevier B.V.

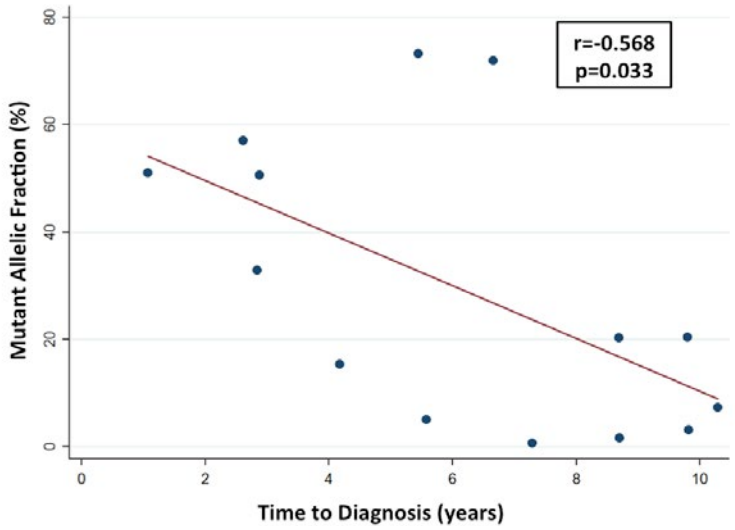
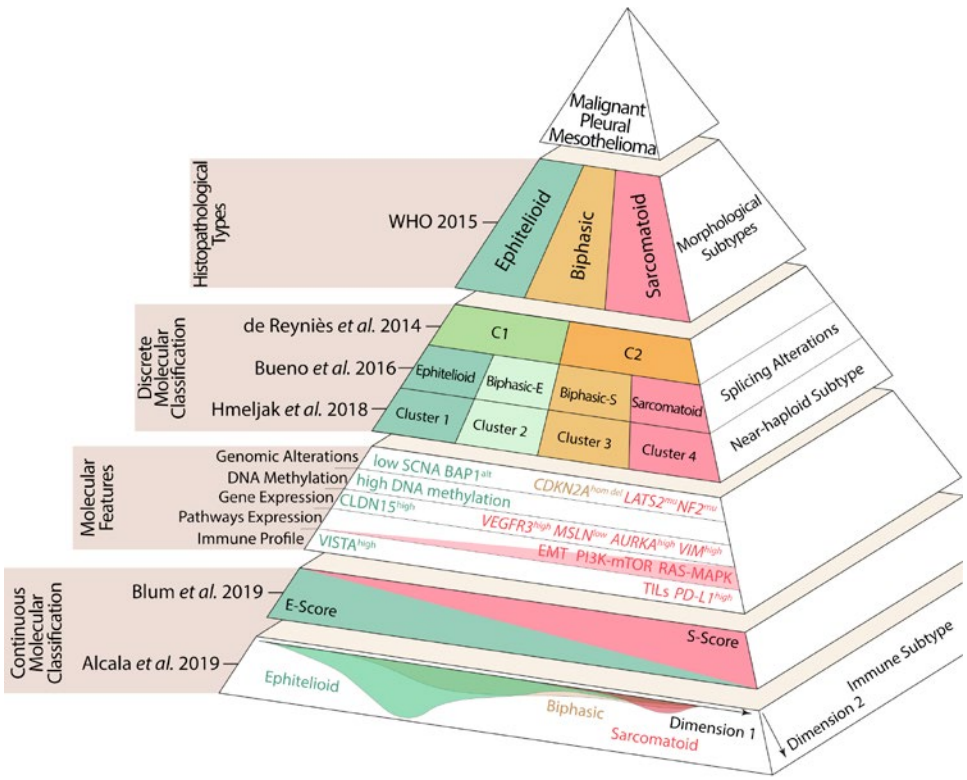


Figure 5. Schematic representation of the different classifications of malignant pleural mesothelioma and their key molecular features. The front face of the pyramid represents the current WHO classifications (top) and the different discrete (middle) and continuous (bottom) molecular classifications that have been proposed. For discrete molecular classifications, the proposed molecular clusters are reported. For Blum et al., a gradient between the epithelioid score (E-score) and the sarcomatoid score (S-score) is depicted. For Alcala et al., the association between the first dimension of the molecular classification and the WHO histological type is shown. The side face of the pyramid represents features and subtypes mentioned in each study that have not been reported to be significantly correlated with the histological types; for Alcala et al., the second molecular dimension is represented, which was shown to summarize features independent of the WHO classification. Colours represent the association between features and the different types: red, sarcomatoid or sarcomatoid-like profiles; orange, biphasic or biphasic-like profiles; green, epithelioid or epithelioid-like profiles; and grey, no proven association with the WHO classification. Reprinted by permission from Fernandez-Cuesta et al. (2021), © 2021.



the first molecular map of lung neuroendocrine neoplasms (Gabriel et al., 2020), led (Foll and Fernandez-Cuesta, 2020) and contributed (Lantuejoul et al., 2020) to reviews in the field of lung neuroendocrine neoplasms, and contributed (Dr Fernandez-Cuesta) to the 2021 fifth edition of the WHO Classification of Tumours of the lung, pleura, thymus, and heart. The Rare Cancers Genomics initiative has a strong computational biology component, particularly for the analysis and integration of –omics data (including whole-genome and/or transcriptome sequencing and methylation arrays), interpretation of histopathological images with deep learning algorithms, and modelling of evolutionary processes associated with tumour progression. GCS actively shares these tools as open-source packages (https://github.com/IARCbioinfo) and maximizes their reuse potential by providing reproducible analyses and online training, ultimately building capacity for cancer genomics (Figure 5).

In the context of how germline variation influences cancer susceptibility, GCS has worked within the International Lung Cancer Consortium (ILCCO) to identify a lung cancer susceptibility variant in the DNA repair gene *ATM*. This variant is a missense variant in *ATM* and has an important genetic effect, with allele carriers having an up to 3–4-fold increase in the risk of lung cancer relative to non-carriers. It also appears to be most relevant to lung cancer in women and lung adenocarcinomas in never-smokers; although it is very rare in most parts of the world, it approaches frequencies of 3% in Ashkenazi Jewish populations (Ji et al., 2020a).

GCS leads the pathology workflow for the Mutographs of Cancers project, a continuing large-scale international study that aims to unveil the carcinogenic role of environmental exposures by analysing the mutational signatures through whole-genome sequencing (https://www.mutographs.org/). GCS also investigated the morphological features of 1000 non-tumoral renal tissues from patients with renal cell carcinoma, and found that the frequency of chronic renal parenchymal changes with the predominance of chronic interstitial nephritis pattern in patients with renal cell carcinoma varies by country; these changes are more frequent in Romania and Serbia (Table 2) (Abedi-Ardekani et al., 2021).

Table 2. Odds ratios and 95% confidence intervals (CIs) for association between country and observation of moderate to severe chronic renal parenchymal changes in non-neoplastic kidney tissues of patients with renal cell carcinoma

Recruiting country	Odds ratio (95% CI)			
	Unadjusted	Model 1[a]	Model 2[b]	Model 3[c]
Russian Federation	Reference	Reference	Reference	Reference
United Kingdom	0.84 (0.25–2.76)	0.64 (0.19–2.16)	0.59 (0.17–2.01)	0.39 (0.10–1.49)
Czechia	1.79 (0.82–3.92)	1.34 (0.59–3.00)	1.54 (0.67–3.57)	1.28 (0.56–2.91)
Romania	3.40 (1.41–8.18)	3.12 (1.26–7.74)	3.16 (1.24–8.03)	2.67 (1.07–6.67)
Serbia	4.63 (1.33–16.08)	5.06 (1.39–18.44)	6.27 (1.40–28.02)	4.37 (1.20–15.96)
Romania and Serbia	3.62 (1.57–8.32)	3.42 (1.44–8.12)	3.53 (1.45–8.58)	2.96 (1.24–7.03)

[a] Adjusted for age, sex, and percentage of medulla.
[b] Adjusted for age, sex, percentage of medulla, stage, and tumour size.
[c] Adjusted for age, sex, diabetes, hypertension, and use of non-steroidal anti-inflammatory drugs (NSAIDs).
Source: Reproduced from Abedi-Ardekani et al. (2021). Copyright © 2021, Abedi-Ardekani et al.