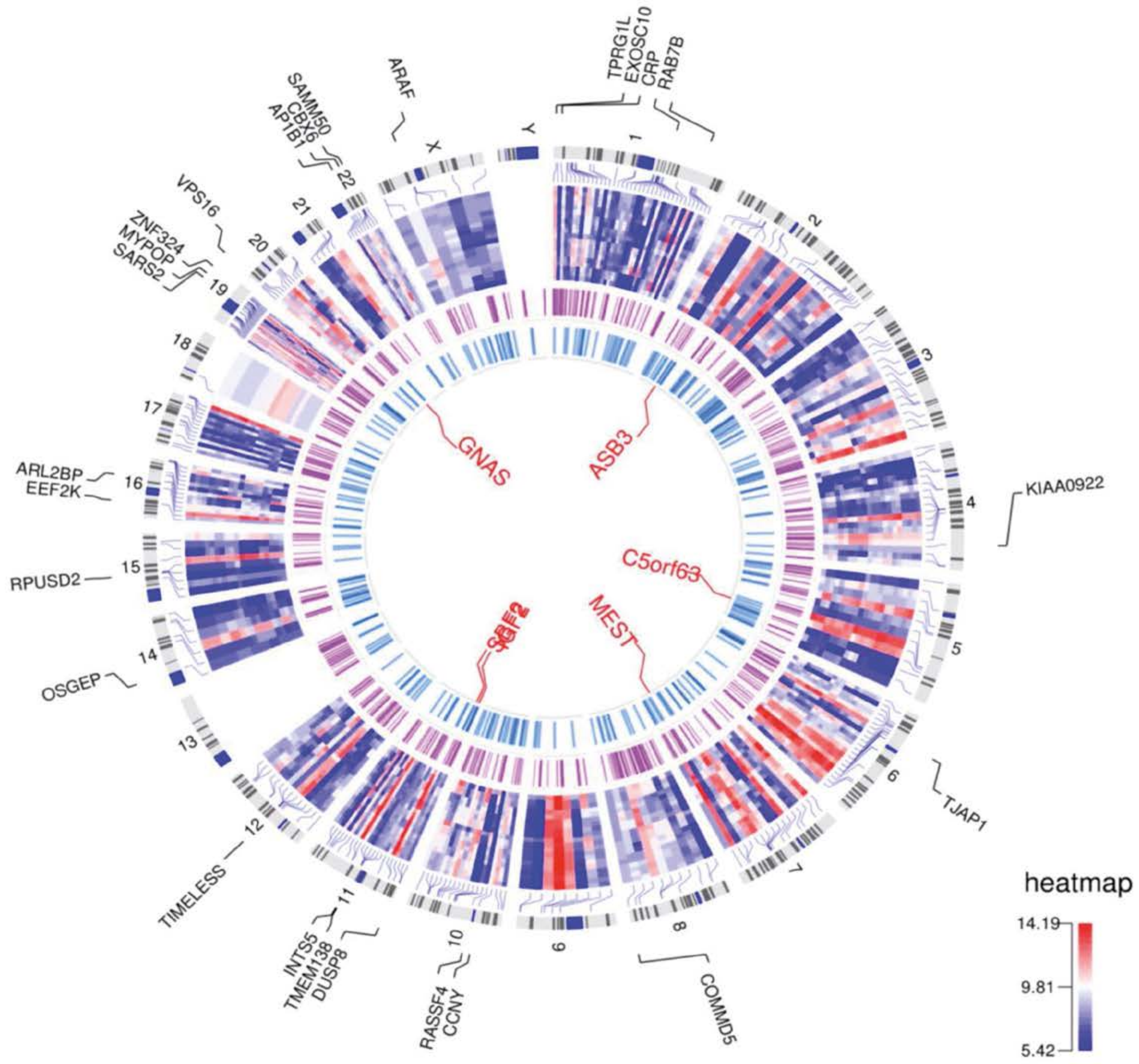

TPRG1L
EXOSC10
CRP
RAB7B
ARAF
SAMM50
CBX6
AP1B1
VPS16
ZNF324
MYPOP
SARS2
ARL2BP
EEF2K
RPUSD2
OSGEP
TIMELESS
INTS5
TMEM138
DUSP8
RASSF4
CCNY
COMMD5
TJAP1
KIAA0922
GNAS
ASB3
C5orf63
MEST
heatmap
14.19
9.81
5.42

© IARC

Section of Mechanisms of Carcinogenesis (MCA)

Section head
Dr Zdenko Herceg

The Section of Mechanisms of Carcinogenesis (MCA) conducts studies aimed at elucidating molecular mechanisms by which environmental exposures induce genetic and epigenetic alterations and deregulate molecular pathways critical for cancer development and progression, thus enhancing the evidence base directly relevant to studies of cancer causation and prevention. Emphasis is placed on events that precede or drive tumour initiation and progression. Key MCA strategies include innovative research and the development of genomic/epigenomic and screening methodologies and bioinformatics resources that are applicable to experimental models and biobanks associated with population-based and epidemiological studies. MCA also contributes to translational studies, through the discovery of mechanism-based biomarkers of exposure, early detection, and risk stratification. MCA studies are interdisciplinary in nature, and the synergistic collaborations with other IARC laboratory-based scientists and epidemiologists as well as external groups advance major IARC programmes. The Section comprises two groups, the Epigenetics Group (EGE) and the Molecular Mechanisms and Biomarkers Group (MMB), which work in close collaboration to create synergies and better exploit and further expand unique research tools and expertise.

Epigenetics Group (EGE)

Group head
Dr Zdenko Herceg

Scientist
Dr Hector Hernandez-Vargas

Secretariat
Ms Asiedua Asante
(until February 2014)
Ms Elizabeth Page

Visiting scientists
Dr Fabienne Lesueur
(until January 2014)
Dr Terence Dwyer
(until June 2014)
Dr Zsuzsa Rakosy
(until July 2015)
Dr Hae Dong Woo

Postdoctoral fellows
Dr Srikant Ambatipudi
Dr Davide Degli Esposti
Dr Nora Fernandez-Jimenez
Dr Akram Ghantous
Dr Reetta Holmila
(until February 2014)
Dr Nawapol Kunkeaw
(until November 2015)
Dr Ho Sun Lee
(until September 2014)
Dr Vibha Patil
Dr Ruethairat Sriraksa
(until August 2014)
Dr Fazlur Talukdar

Students
Mr Pierre-Benoît Ancey
(until December 2015)
Ms Marine Benaissa
(until November 2014)
Mr Clément Guillot
(until February 2014)
Ms Andrea Hallaburkova
(until July 2015)
Ms Marija Klasic
(until May 2015)
Mr Quentin Mandier
(until March 2015)
Ms Athena Sklias
Mr Diego Uribe
(until November 2014)

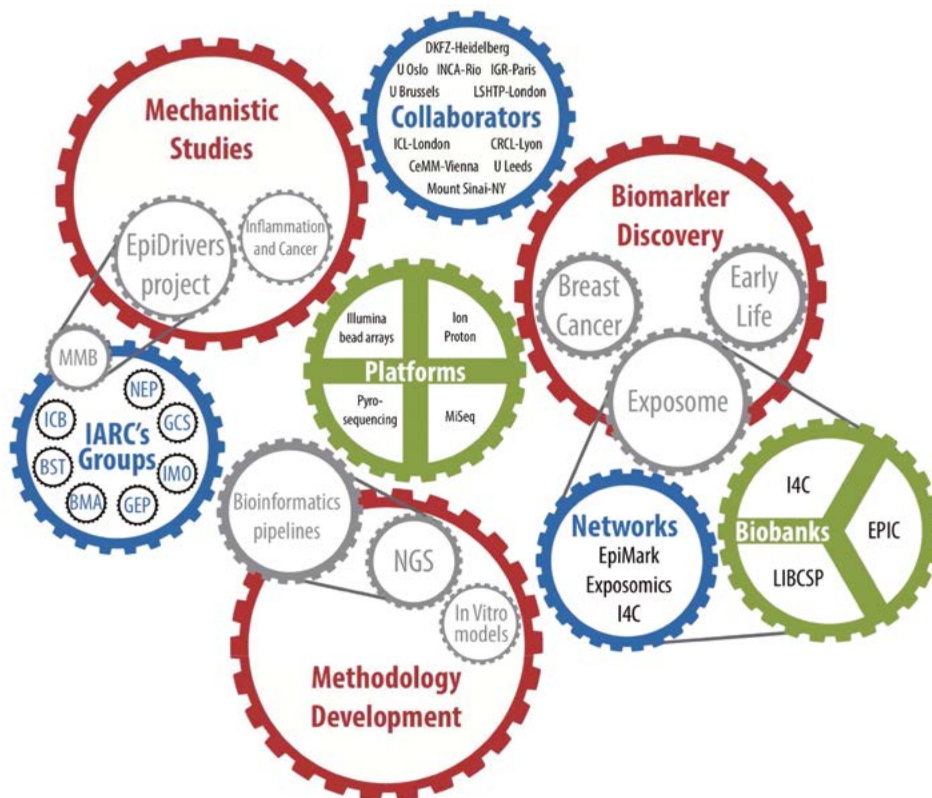
Laboratory technicians
Ms Marie-Pierre Cros
Mr Cyrille Cuenin

Bioinformatics assistant
Mr Vincent Cahais

Trainees
Ms Cansu Cirzi
(until September 2014)
Mr Antonin Jay
(until May 2015)

The Epigenetics Group (EGE) conducts studies aimed at providing critical insights into epigenetic mechanisms of carcinogenesis through the identification of epigenome alterations and molecular pathways deregulated by environmental exposures. Another focus of EGE is identifying epigenetic biomarkers of exposure and cancer risk and contributing to the characterization of key components of the exposome. This is achieved through mechanistic studies of functionally important epigenetic "driver" genes and molecular pathways altered by specific cancer risk agents and by the application of cutting-edge epigenomics in conjunction with unique biospecimens from population-based cohorts (Figure 1). EGE also develops epigenomic methodologies, profiling strategies, and bioinformatics tools, applicable to population-based cohorts and molecular epidemiology studies coordinated by IARC researchers and external collaborators.

Figure 1. Research themes, collaborators, technological platforms, and resources of the Epigenetics Group (EGE). EGE conducts studies aiming to characterize epigenetic alterations and molecular pathways deregulated by specific cancer risk factors and to identify epigenetic biomarkers of exposure, early detection, and risk stratification. It also develops epigenomic and profiling strategies and bioinformatics tools applicable to in vitro models and population-based studies. EGE capitalizes on recent conceptual and technological advances that have resulted in exciting opportunities for cancer epigenetics in understanding the causes of common cancers. EGE's programme is carried out in close collaboration with IARC scientists and epidemiologists as well as external collaborators, many of whom are part of international networks established to share technological platforms and biological resources. © IARC.



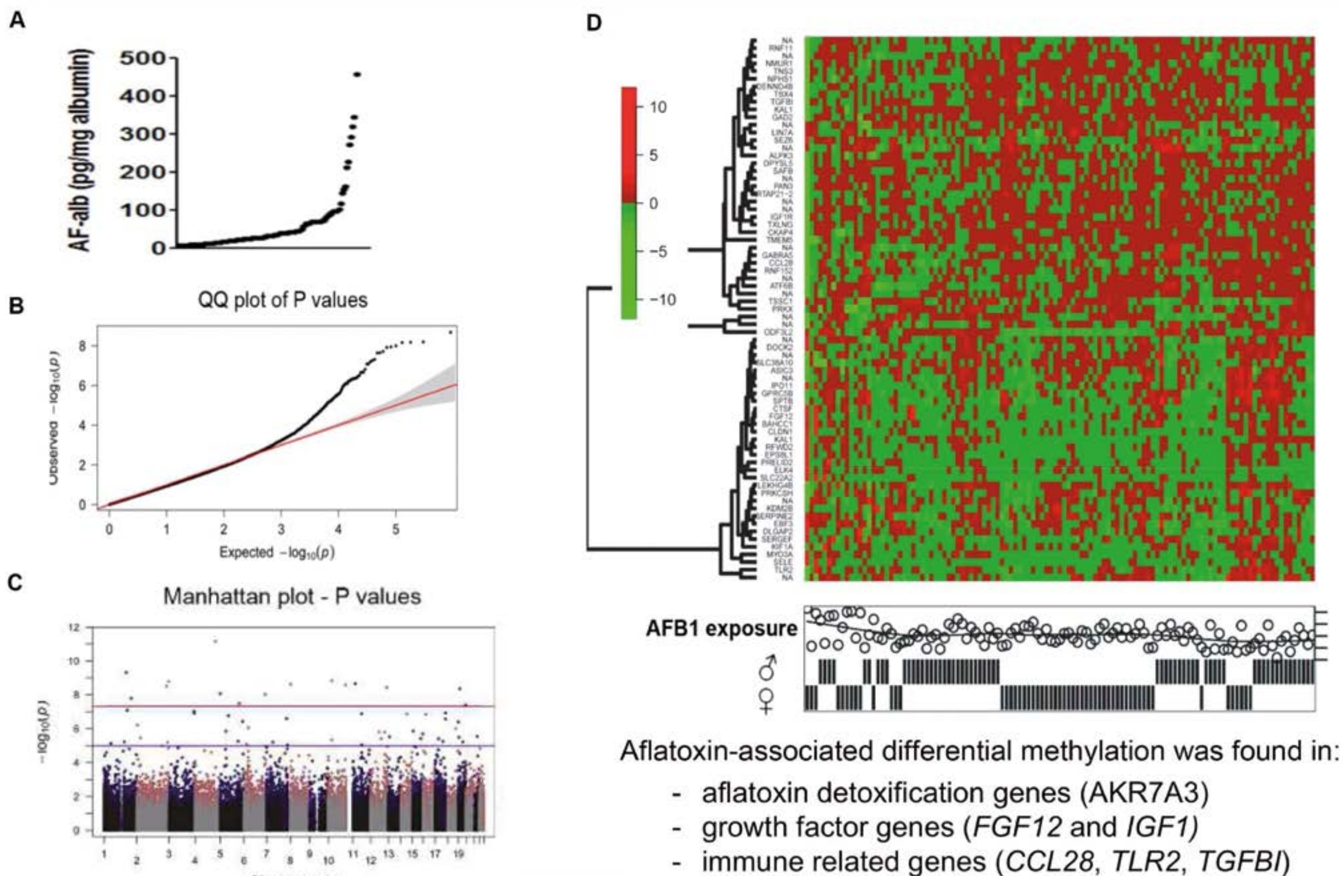
Exposure to aflatoxin B_1 in utero is associated with DNA methylation changes in white blood cells of infants in The Gambia

Exposure to environmental toxins during embryonic development may lead to epigenetic changes that influence disease risk in childhood and later life. EGE investigated the consequences of early-life exposure to aflatoxin at the epigenome (DNA methylation [DNAm]) level. Aflatoxin exposure in women from a rural region in The Gambia was assessed in plasma taken at 1–16 weeks of pregnancy, and global DNAm of white blood cells from their infants was measured using the Illumina Infinium HumanMethylation450 BeadChip. Aflatoxin exposure in the mothers was found to be significantly correlated with DNAm in their infants for a subset of CpG sites. Aflatoxin-associated differential methylation was observed in growth factor genes, immune-related genes, and a gene involved in aflatoxin detoxification (Figure 2) (Hernandez-Vargas et al., 2015). In addition, EGE identified that the effect of maternal nutrition on DNAm at specific genomic loci exhibits the hallmarks of "metabolic imprinting", including a critical window of sensitivity (in the pre-implantation embryo) and a dose–response relationship between exposure and outcome (Silver et al., 2015). These studies show that maternal exposure and diet during the early stages of pregnancy is associated with changes in epigenome patterns of infants. This reinforces the need for interventions, especially during critical periods of fetal and infant development.

Targeted deep DNA methylation analysis of circulating cell-free DNA in plasma using massively parallel semiconductor sequencing

Circulating cell-free DNA (cfDNA) isolated from the plasma of individuals with cancer has been shown to harbour cancer-associated changes in DNAm, and thus represents an attractive target for biomarker discovery. However, the reliable detection of DNAm changes in body fluids has proven to be technically challenging. EGE has developed a novel method that enables sensitive and targeted deep DNAm analysis in minute amounts of DNA present in body fluids using massively parallel semiconductor sequencing (the Ion Torrent PGM sequencer). This approach was applied to assess in plasma cfDNA the methylation of a panel of genes, including *FBLN1*, *HINT2*, *LAMC1*, *LTBP1*, *LTBP2*, *PSMA2*, *PSMA7*, *PXDN*, *TGFB1*, *UBE2L3*, *VIM*, and *YWHAZ*, and to evaluate the potential of these genes as novel biomarkers for hepatocellular carcinoma in two different case–control studies, one in France and one in Thailand. Methylation in cfDNA was detected for specific genes (*FBLN1*, *PSMA7*, *PXDN*, and *VIM*), with substantial differences in methylation patterns between cases and controls (Figure 3) (Vaca-Paniagua et al., 2015a, and unpublished data from

Figure 2. Differential methylation associated with early-life exposure to aflatoxin. (A) Distribution of aflatoxin exposure during pregnancy in all mothers. (B) Quantile–quantile plot of the *P*-values after the association between DNA methylation and aflatoxin exposure (as a continuous variable). (C) Manhattan plot to illustrate the distribution of *P*-values across somatic chromosomes. (D) Heat map of the 71 CpGs associated with in utero aflatoxin B_1 (AFB1) exposure. Annotations in the lower panel illustrate the corresponding aflatoxin exposure and sex. Reprinted with permission from Hernandez-Vargas et al. (2015), by permission of Oxford University Press.

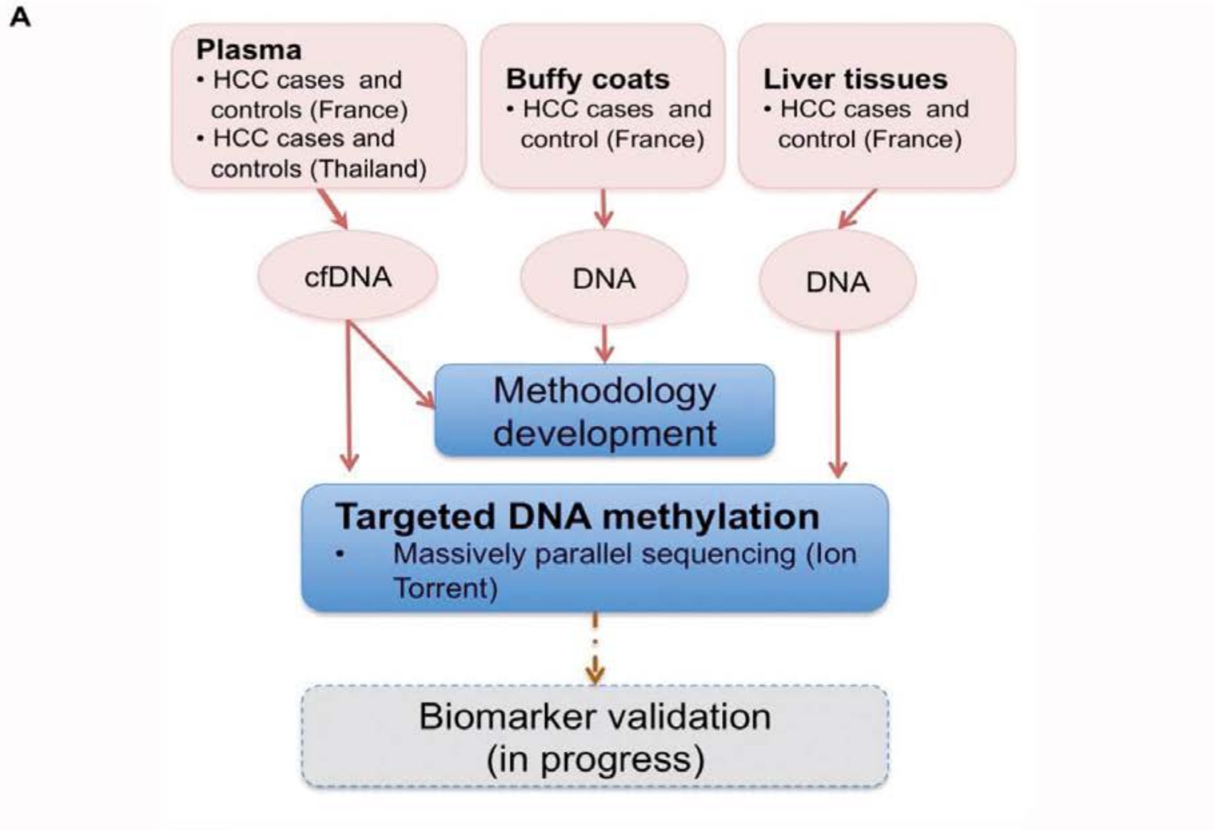


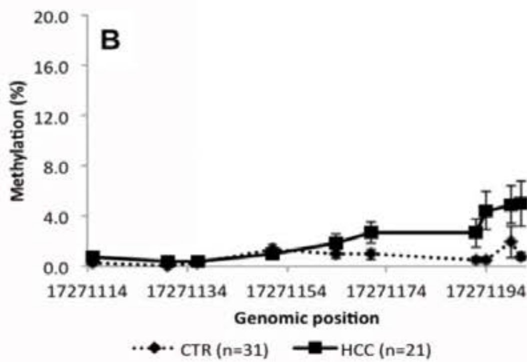
EGE). These results provide evidence that changes in methylation levels of *VIM* and *FBLN1* in cfDNA are associated with hepatocellular carcinoma and may represent useful plasma-based biomarkers for improved diagnostic accuracy and patient surveillance (Vaca-Paniagua et al., 2015a, and unpublished data from EGE). This study represents a proof of principle demonstrating the applicability of massively parallel semiconductor sequencing as a non-invasive, cost-effective, and time-efficient approach to identify, develop, and validate epigenetic biomarkers that are potentially translatable into epidemiological and clinical settings.

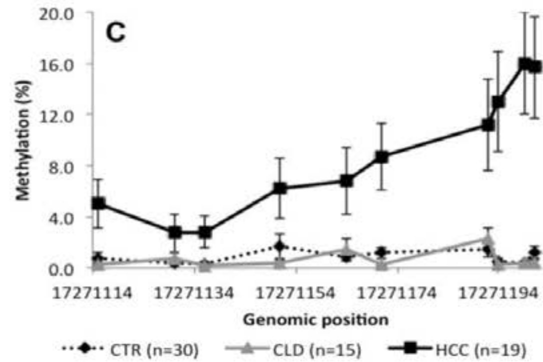
Development of epigenomic methodologies and bioinformatics tools applicable to population-based cohorts and molecular epidemiology

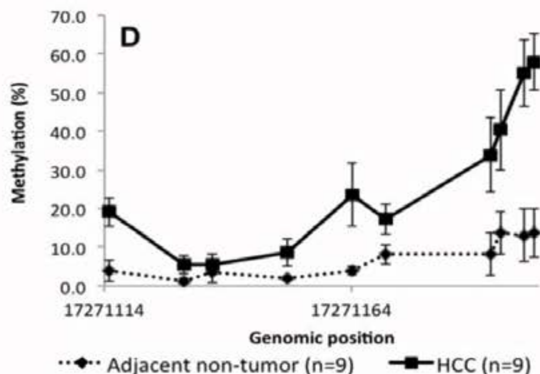
EGE has exploited improvements in the throughput and cost of methylation, histone modifications, and microRNA sequencing brought about by the recent establishment of a new-generation array platform (Illumina Infinium) for methylome and transcriptome profiling, and next-generation sequencing (NGS)-based platforms (Illumina MiSeq, Illumina Genome Analyzer, and Ion Torrent) at IARC and external collaborators (Figure 1). These methodologies have enabled EGE to move from focused approaches to comprehensive epigenome-wide approaches and to develop several new and original topics in cancer epigenetics (Ghantous et al., 2014; Hernandez-Vargas et al., 2015; Kuasne et al., 2015; Lambert et al., 2015; Martin et al., 2014; Silver et al., 2015; Vaca-Paniagua et al., 2015a). These developments have also motivated the building of bioinformatics capacity within EGE, with a first generation of data-mining tools specifically designed for epigenomic analyses.

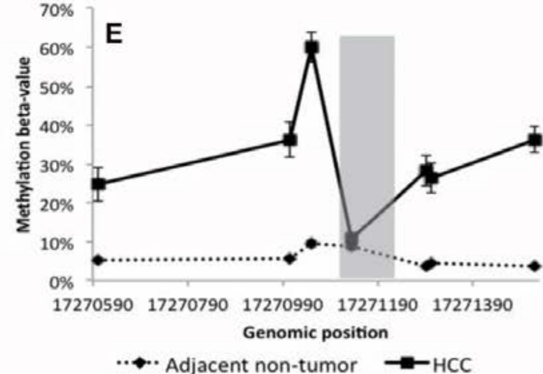
Figure 3. DNA methylation analysis of plasma circulating cell-free DNA (cfDNA) by targeted deep sequencing to evaluate potential epigenetic biomarkers of cancer. (A) General outline of the study and methodology development. (B–E) The DNA isolated from plasma of cases and controls was subjected to targeted deep DNA methylation analysis using massively parallel semiconductor sequencing. *VIM* methylation in circulating DNA in cases from France (B) and from Thailand (C), in tissue (D), and in The Cancer Genome Atlas (TCGA) data (E); the grey area represents the area analysed by massively parallel sequencing in this study. CLD, chronic liver diseases; CTR, controls; HCC, hepatocellular carcinoma. The error bars represent the standard error of the mean. Compiled from Vaca-Paniagua et al. (2015a).











EGE is grateful to the following for their collaboration:

Gabriella Tikellis, Victoria, Australia; Christoph Bock, Vienna, Austria; François Fuks, André Nogueira da Costa, Brussels, Belgium; Sheila Lima, Natalia Motta de Araujo, Felipe Pinto, Rio de Janeiro, Silvia Rogatto, São Paulo, Brazil; Anastas Gospodinov, Sofia, Bulgaria; Chantal Matar, Ottawa, Canada; Gordan Lauc, Nino Sincic, Zagreb, Croatia; Saadi Khochbin, Claire Vourc'h, Grenoble, Patrick Mehlen, Philippe Merle, Peter Mulligan, Isabelle Chemin, Jean-Yves Blay, Julien Marie, Alain Puissieux, Lyon, Ellen Obberghen-Schilling, Nice, Jacqueline Clavel, Paris, Thierry Frebourg, Rouen, France; Rudolf Kaaks, Christoph Plass, Heidelberg, Zhao-Qi Wang, Jena, Germany; Soterios Kyrtopoulos, Athens, Greece; Margit Balazs, Debrecen, Hungary; Bernardo Bonanni, Milan, Italy; Felipe Vaca Paniagua, Mexico City, Mexico; Anne-Lise Börresen-Dale, Vessela N. Kristensen, Siri Haberg, Oslo, Norway; Manel Esteller, Barcelona, Jose Ramon Bilbao, Bilbao, Spain; Nicole Probst, Basel, Switzerland; Temduang Limpaiboon, Khon Kaen, Thailand; Yun Yun Gong, Belfast, Michael Routledge, Leeds, Branwen Hennig, Andrew Pretince, Elio Riboli, Paolo Vineis, London, Terry Dwyer, Oxford, United Kingdom; Robert A. Waterland, Houston, Steve Horvath, Joseph Wiemels, Los Angeles, Jia Chen, New York, Stephanie London, Research Triangle Park, Bing Ren, San Diego, Martyn Smith, San Francisco, USA.

Financial support from the following bodies is gratefully acknowledged:

Association pour la recherche sur le Cancer (ARC), France
Agence nationale de recherches sur le sida et les hépatites virales (ANRS), France
Bill & Melinda Gates Foundation, USA
Canadian Institutes of Health Research, Canada
European Commission, Brussels, Belgium
Institut national du Cancer (INCa), Paris, France
INSERM, Paris, France
La Ligue contre le Cancer, Comité du Rhône, France
National Cancer Institute, National Institutes of Health, USA

Molecular Mechanisms and Biomarkers Group (MMB)

Group head
Dr Jiri Zavadil

Scientists
Dr Michael Korenjak
Dr Elisabetta Kuhn
(until September 2014)
Dr Magali Olivier

Research assistant
Dr Stéphanie Villar

Secretariat
Ms Sylvie Nouveau

Visiting scientists
Dr Monica Hollstein
Dr Zuzana Vojtechova
(until July 2015)

Postdoctoral fellows
Dr Xavier Castells Domingo
(until February 2015)
Dr Gihan Hosny
Dr Sengkwawoh Lueong Smiths
Dr Manuraj Pandey
Dr Yan Song
(until October 2014)

Students
Ms Maude Ardin
Ms Sarah Barrin
(until September 2014)
Ms Hana Huskova
Ms Claire Tissot
(until November 2014)
Ms Maria Zhivagui

The overarching objective of the Molecular Mechanisms and Biomarkers Group (MMB) is to establish an evidence base for cancer prevention, by identifying molecular mechanisms and biomarkers of carcinogenesis associated with specific environmental and lifestyle risk factors. MMB characterizes new biomarkers of exposure and tumorigenesis by mutational signature screens in experimental in vitro models, as well as in tumour tissues and plasma circulating cell-free DNA (cfDNA), taking advantage of existing epidemiological studies and also supporting new ones. MMB develops and validates screening methods and bioinformatics tools applicable to population-based and mechanistic studies. Collectively, MMB aims to advance the understanding of mechanisms of carcinogenesis and to facilitate evidence-based cancer prevention strategies.

Identification of cancer mutational signatures and driver mutations in vitro

Oncogenic stress in primary cells can result in a bypass of their finite lifespan, followed by clonal expansion due to the accumulation of mutations that support cellular immortalization. MMB exploits this property by combining carcinogen exposure of primary cells with barrier bypass–clonal expansion (BBCE) assays. Deep DNA sequencing of immortalized cell clones from carcinogen-exposed murine embryonic fibroblasts yielded genome-wide mutational signatures matching those found in human cancers (Olivier et al., 2014). It has also resulted in the identification of recurrent selected mutations in known cancer driver genes (Figure 1). MMB currently characterizes the mutational signatures of new candidate carcinogens and studies the roles of selected driver mutations in cell immortalization. In sum, the BBCE assays provide a powerful strategy for the identification of mutation spectra introduced by environmental chemicals and of driver mutations critical for cellular transformation.

Mutational signature of carcinogenic aristolochic acid in urological tumours

Exposure to aristolochic acid (AA) leads to severe nephropathies and urothelial cancers. MMB devised a customized low-coverage exome sequencing approach to identify the signature of AA exposure in formalin-fixed, paraffin-embedded upper tract urothelial carcinoma (UTUC) and renal cell carcinoma (RCC) tumours from the residents of endemic nephropathy regions in Croatia and Bosnia with a history of consumption of bread made from wheat contaminated by AA-containing seeds of *Aristolochia clematitis*. A mutational signature consistent with exposure to AA was observed in 5 RCC and 15 UTUC tumours (Figure 2) (Jelaković et al., 2015). In addition, MMB contributed to a study identifying AA signature-containing RCC tumours in Romanian patients (Scelo et al., 2014). The identification of multiple tumour types associated with AA exposure presents new epidemiological and public health implications for

Figure 1. Recurrent mutations in known cancer driver genes modelled in an in vitro clonal selection system. Concentric tracks represent 25 immortalized clones arising from murine embryonic fibroblast cultures harbouring a transgene expressing activation-induced cytidine deaminase (AID) or exposed to various mutagenic insults: AA, aristolochic acid; AFB1, aflatoxin B_1; B[a]P, benzo[*a*]pyrene; MNNG, *N*-methyl-*N′*-nitro-*N*-nitrosoguanidine; Spont, spontaneously immortalized (untreated); UVC, ultraviolet light class C. Dots represent enriched single base substitutions (mutations) located in particular chromosomal positions (chromosomes shown in the centre). The observed mutations were mostly exposure-specific (orange dots) and also non-specific (grey dots). On the perimeter, 86 recurrently mutated cancer genes are shown (red, oncogenes; blue, tumour suppressor genes; green, chromatin-associated factors; black, other cancer genes), with the observed number of mutations shown in parentheses. © IARC.

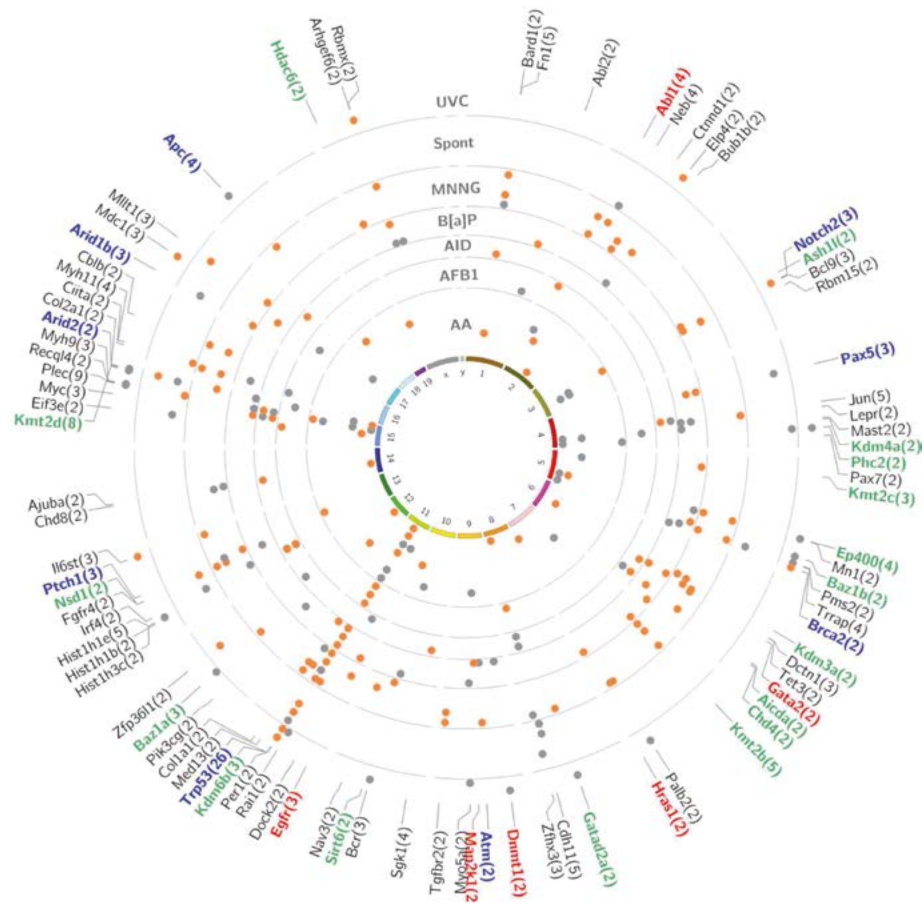
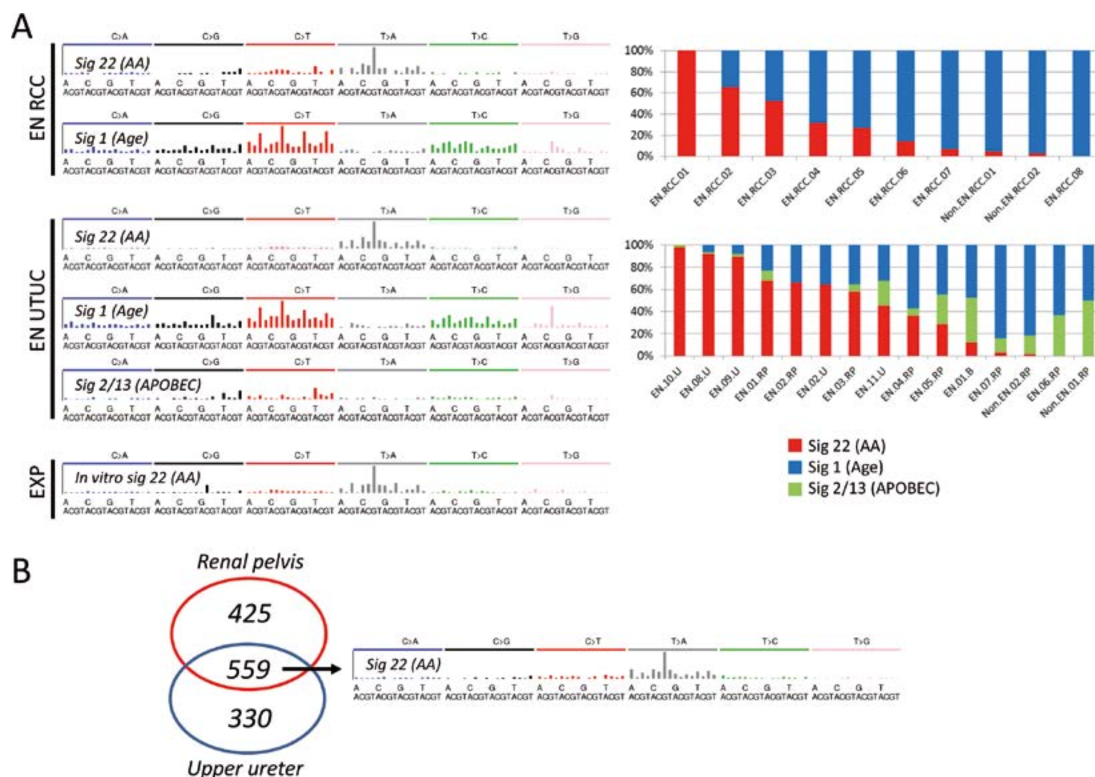


Figure 2. Mutational signatures in urological tumours of endemic nephropathy (EN) patients. (A) Signature (Sig) 22 corresponding to the mutagenic effects of aristolochic acid (AA) was observed in renal cell carcinoma (RCC) and upper tract urothelial carcinoma (UTUC) tumours of Croatian and Bosnian patients with EN. The bar graphs on the left show the individual signatures found in the urological tumours studied and the AA single base substitutions (SBSs) are represented by colours and labelled on top of each graph, and the frequencies of each of the possible combinations of an SBS in a particular trinucleotide context (listed under each graph) are shown; the predominant T > A in the C_G context is the typical feature of the AA signature. The UTUC samples also harboured the signature for increased APOBEC enzyme activity. The bar graphs on the right show the relative percentage contribution of each signature to the mutation load in individual tumour samples. (B) Concurrent UTUC tumours found in distinct anatomical sites (the renal pelvis and upper ureter) of one EN patient harbour a high number of overlapping AA-specific mutations, suggesting a possible mechanism of tumour spread by tumour cell seeding along the urinary tract. © IARC.



the incidence and prevention of AA-associated cancers worldwide. The screen developed by MMB can address the role of AA in cancers observed in high-risk populations exposed to the compound through the widespread use of alternative herbal remedies.

Assessing the performance of deep sequencing for the identification of clinically relevant somatic tumour mutations in circulating cfDNA in lung cancer

Circulating cfDNA extracted from the plasma of cancer patients may contain a significant fraction of tumour DNA. Somatic mutation analysis is part of the standard management of metastatic lung cancer to select gene-targeted therapies. Biopsy samples are often the only material available to access the tumour DNA, but they provide limited amounts of DNA and may not be representative of the entire tumour mass. To investigate whether cfDNA could be used as a surrogate tissue for detecting clinically relevant mutations in lung cancer from non-smokers, MMB used deep sequencing that enables highly sensitive mutation detection (Couraud et al., 2014). The results demonstrate that this method is suitable for the detection of tumour mutations in cfDNA with good sensitivity and specificity. Thus, cfDNA may be a promising resource for diagnosis and follow-up of lung cancer.

Bioinformatics tools for molecular cancer research

The recent interest in genome-wide mutational signatures observed in human cancers uncovered a need for user-friendly tools that would enable streamlined data analyses accessible to scientists with limited expertise in bioinformatics. To fill this gap, MMB developed MutSpec, an open-source software package embedded in the popular, user-friendly bioinformatics platform Galaxy. MutSpec includes tools performing variant annotation and advanced statistics for identifying mutational signatures present in cancer genomes and comparing the signatures obtained with those in the COSMIC database and other sources. MutSpec can analyse data from whole-exome, whole-genome, or targeted sequencing performed in human or mouse samples. The results are organized in tabular and rich graphical summaries. MutSpec facilitates systematic analyses of mutation spectra by a wider range of scientists with basic bioinformatics skills, to promote new studies on cancer etiology.

MMB is grateful to the following for their collaboration:

Richard Cotton, Melbourne, Heather Thorne, Victoria, Australia; Shahrokh F. Shariat, Vienna, Austria; Joëlle Nortier, Sandrine Rorive, Thierry Roumeguère, Brussels, Belgium; David Malkin, Toronto, Canada; Min Dai, Jian Huang, Guohui Li, Yamei Niu, Yan Song, Weimin Tong, Beijing, China; Damir Dittrich, Krešimir Karlović, Maja Mišić, Karla Tomić, Slavonski Brod, Fran Borovecki, Bojan Jelaković, Sandra Karanović, Neda Slade, Zagreb, Croatia; Tomas Stopka, Ruth Tachezy, Zuzana Vojtechova, Prague, Czech Republic; Gerard Zalcman, Caen, Jean-Paul Bringuier, Barbara Charbotel, Isabelle Chemin, Sebastien Couraud, Mojgan Devouassoux-Shisheboran, Charles Dumontet, Béatrice Fervers, Claire Tissot, Lyon, Géraldine Cancel-Tassin, Olivier Cussenot, Yann Neuzillet, Francois Radvanyi, Evanguelos Xylinas, Paris, France; Thangarajan Rajkumar, Rajaraman Swaminathan, Chennai, Ganesh Bakshi, Rajesh Dikshit, Mumbai, India; Mardhiah Mohammad, Arifin Bin Kaderi Mohd, Kuantan Pahang, Malaysia; Rosa María Álvarez Gómez, Hector Aquiles Maldonado, Enrique Bargalló, David Cantú de León, Veronica Fragoso Ontiveros, Luis Alonso Herrera Montalvo, Alejandro Mohar Betancourt, Carlos Perez Plasencia, Felipe Vaca Paniagua, Mexico City, Mexico; Peter E.M. Taschner, Leiden, The Netherlands; Anne-Lise Borrensen-Dale, Anita Langerod, Oslo, Norway; Steve Rozen, Patrick Tan, Bin Tean Teh, Singapore; Philipp Bucher, Lausanne, Shana J. Sturla, Zurich, Switzerland; Mark LaBarge, Martha Stampfer, Berkeley, Adriana Heguy, New York, Ronald A. Herbert, Arun Pandiri, Research Triangle Park, Kathleen G. Dickman, Arthur P. Grollman, Thomas Rosenquist, Viktoria S. Sidorenko, Stony Brook, USA.

Financial support from the following bodies is gratefully acknowledged:

European Commission, Brussels, Belgium
Intergroupe Francophone de Cancérologie Thoracique (IFCT), Paris, France
National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), Research Triangle Park (NC), USA
INSERM/Institut national du Cancer (INCa), Paris, France