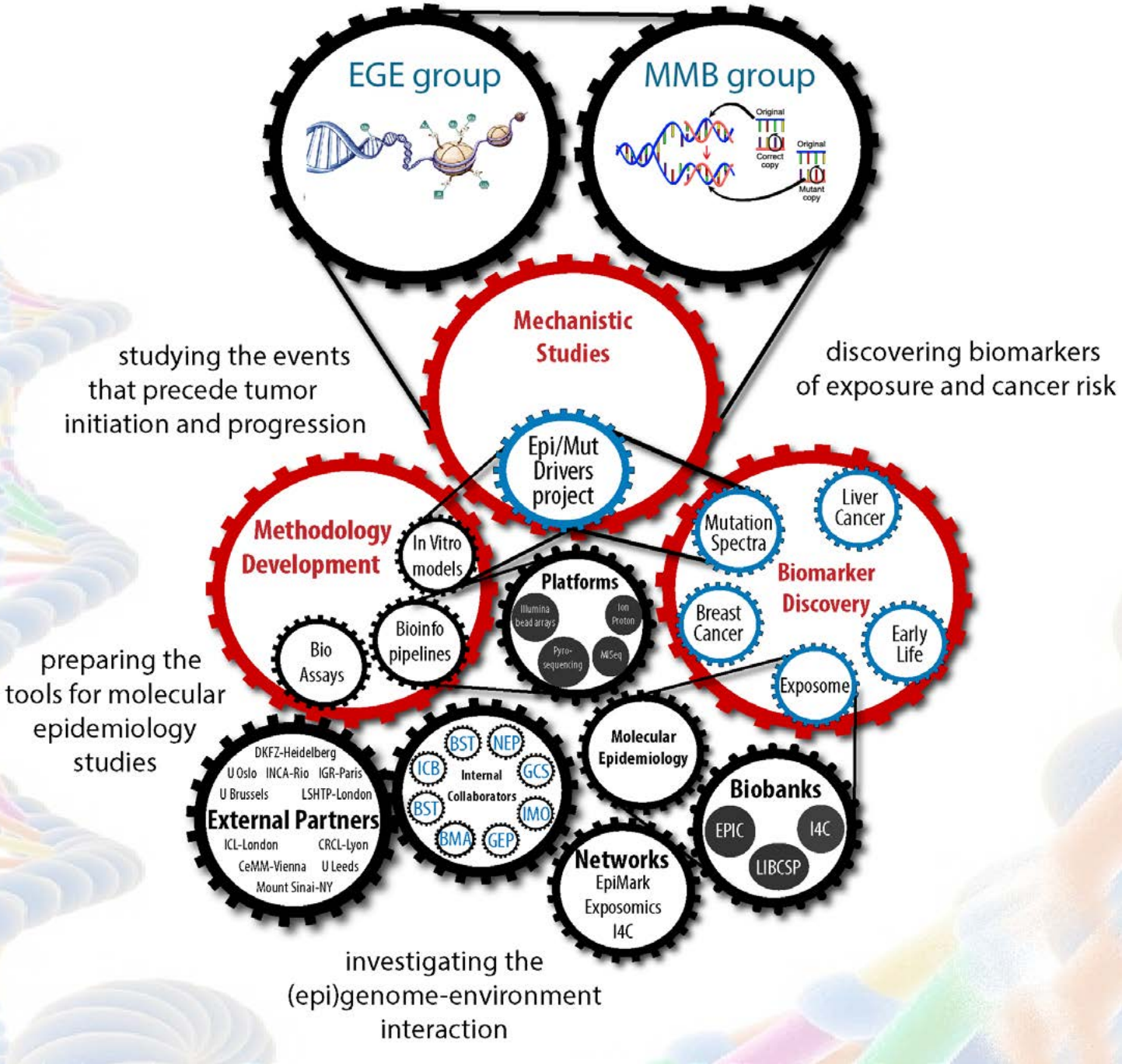
Understanding molecular mechanisms for cancer prevention



Section of Mechanisms of Carcinogenesis (MCA)

Section head
Dr Zdenko Herceg

Epigenetics Group (EGE)

Group head
Dr Zdenko Herceg

Scientists
Dr Akram Ghantous
Dr Hector Hernandez-Vargas
(until September 2017)

Secretariat
Ms Elizabeth Page

Visiting scientist
Dr Hae Dong Woo (until June 2017)

Postdoctoral fellows
Dr Srikant Ambatipudi
(until September 2016)
Dr Felicia Chung
Dr Davide Degli Esposti
(until December 2016)
Dr Cuong Duong
Dr Szilvia Ecsedi (until April 2017)
Dr Nora Fernandez-Jimenez
(until December 2016)
Dr Akram Ghantous
(until January 2016)
Dr Andrea Halaburkova
Dr Vibha Patil
Dr Fazlur Talukdar

Doctoral students
Mr Oscar Carmo Araujo
(until July 2017)
Ms Andrea Halaburkova
(until August 2017)
Mr Alexei Novoloaca
Mr Jesus Rodriguez-Aguilera
Ms Irati Romero-Garmendia
(until March 2016)
Ms Athena Sklias
Ms Anna-Luiza Vicente

Research assistants
Ms Marie-Pierre Cros
(until August 2017)
Mr Cyrille Cuenin
Ms Aurélie Salle

Senior research assistant, data management/analyst
Mr Vincent Cahais

Trainees
Mr Miroslav Bobrik
Ms Anne-Claire Boisson
(until August 2017)
Mr Andrés Cardona Echeverry
(until December 2016)
Ms Véronique Chauvet
(until July 2017)
Mr Thibaut Chemarin
(until May 2017)
Ms Diana Narvaez Noguera
(until September 2016)

Molecular Mechanisms and Biomarkers Group (MMB)

Group head
Dr Jiri Zavadil

Scientists
Dr Michael Korenjak
Dr Magali Olivier

Research assistants
Ms Marie-Pierre Cros
Dr Stéphanie Villar
(until August 2017)

Secretariat
Ms Sylvie Nouveau (until July 2017)
Ms Karine Racinoux

Visiting scientist
Dr Monica Hollstein
(until December 2016)

Postdoctoral fellows
Dr Maude Ardin (until March 2017)
Dr Pamela Melki
Dr Manuraj Pandey
Dr Sengkwawoh Lueong Smiths
(until December 2016)

Students
Ms Maude Ardin
(until November 2016)
Mr Thomas Grainger
(until June 2017)
Mr Thibaut Halegua
(until September 2016)
Ms Hana Huskova
(until October 2017)
Ms Souad Kolli
(until August 2017)
Mr Alexis Robitaille
(until June 2016)
Ms Clara Robitaille
(until February 2017)
Ms Shuhan Wang
(until August 2016)
Ms Maria Zhivagui

Improving the knowledge of mechanisms of carcinogenesis related to environmental exposures provides a foundation for studies of cancer etiology, cancer prevention, and carcinogen evaluation, the core activities of IARC. The overarching objective of the Section of Mechanisms of Carcinogenesis (MCA) is to provide the evidence base for the study of cancer causation and prevention by elucidating the molecular mechanisms by which genetic and epigenetic alterations alter critical molecular pathways and promote cancer development. Major emphasis is placed on discerning events that precede or drive tumour initiation and progression related to environmental exposures.

The research of MCA focuses on two priority areas. First, MCA studies are aimed at providing critical insights into mechanisms of carcinogenesis through the identification of molecular alterations and molecular pathways deregulated by specific cancer risk factors. This is achieved through mechanistic studies of functionally important (epi)genetic "driver" events and molecular pathways altered by specific cancer risk agents (with a focus on a set of genotoxic and non-genotoxic agents prioritized according to their relevance to cancer etiology and prevention), using in vitro models and state-of-the-art approaches including (epi)genome-wide screens and functional genomics. Second, MCA is involved in identifying molecular biomarkers of exposure and cancer risk. To this end, MCA uses cutting-edge (epi)genomics, population-based cohorts, and innovative bioinformatics tools to investigate (epi)genomic profiles of specific cancers and surrogate tissues and to identify signatures of cancer risk and exposures. The primary focus is on cancers of the breast, urinary tract, and liver and childhood malignancies. MCA also participates in an interdisciplinary approach aimed at characterizing exposures throughout the life-course (with a particular focus on the fetal exposome and childhood cancer) by building on unique samples from international birth cohorts and other population-based studies.

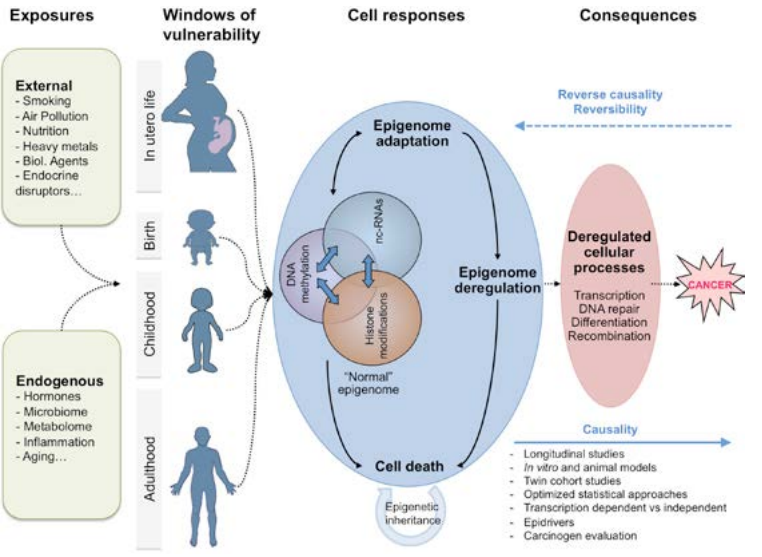
The expected outcome of these studies is the opening up of an opportunity to identify and characterize the key molecular events and pathways that underpin carcinogenesis, thereby elucidating important aspects of cancer etiology and opportunities for prevention.

The Section consists of two Groups: the Epigenetics Group (EGE) and the Molecular Mechanisms and Biomarkers Group (MMB).

Epigenetics Group (EGE)

The Epigenetics Group (EGE) conducts mechanistic studies and epigenetic profiling aimed at enhancing the understanding of epigenetic mechanisms underlying tumour development and progression as well as discovering new cancer biomarkers (Figure 1). EGE exploits new concepts in cancer epigenetics, the availability of unique population-based cohorts, and recent technological advances in epigenomics. 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. Outcomes of the recent studies are an improved knowledge of mechanisms of carcinogenesis associated with environmental factors and the provision of an evidence base for studies of cancer causation and prevention.

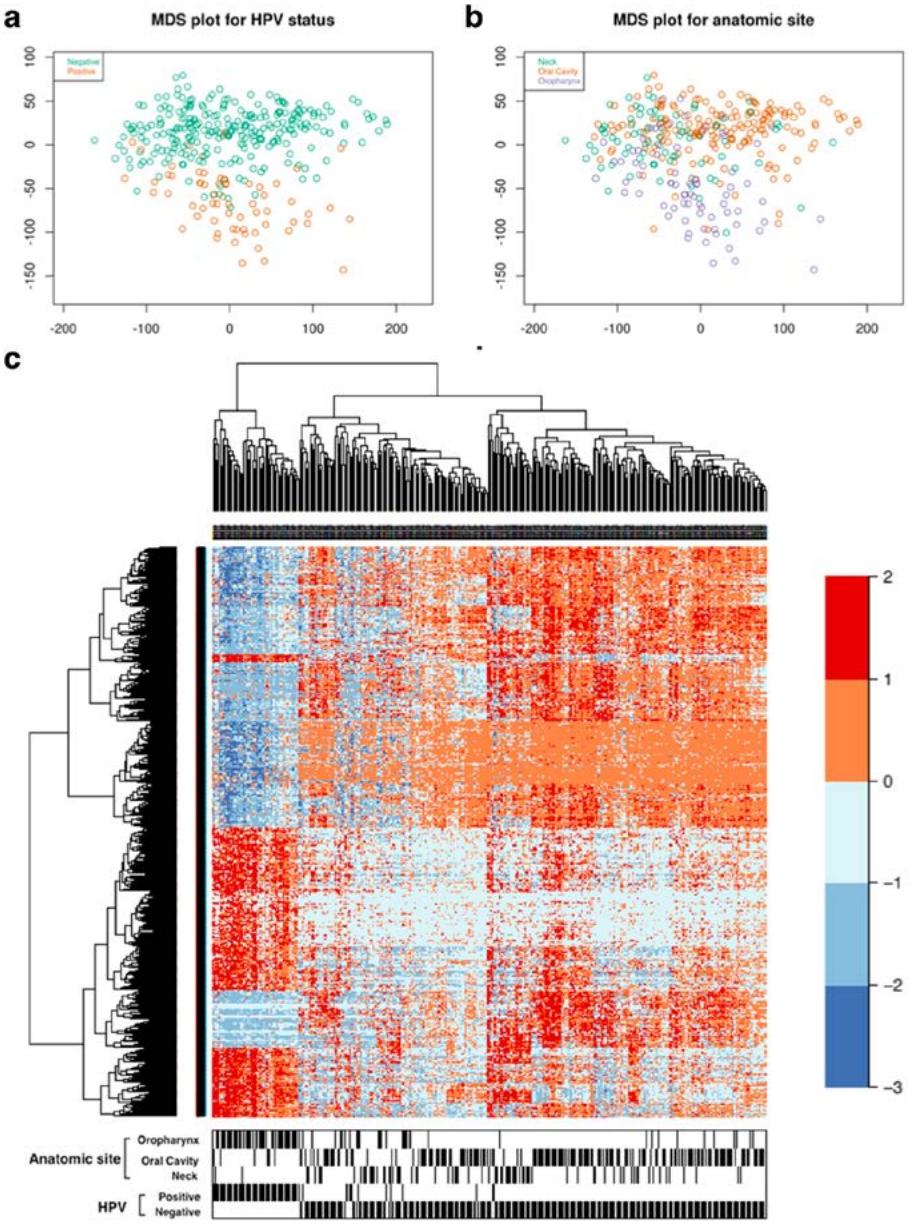
Figure 1. Studying epigenetic mechanisms and environmental origins of cancer. Exposures arising from external sources (e.g. environmental chemicals, air pollution, infectious agents, diet, tobacco use, alcohol consumption, and endocrine disruptors) and internal processes (e.g. metabolism, hormones, inflammation, gut microflora, and ageing) may induce stable and potentially reversible changes in the epigenome. The patterns ("signatures") and persistence of these alterations depend on multiple factors, including the types of epigenetic changes, the dosage and duration of the exposure, the tissue type, and the developmental stage. Thus, epigenetic mechanisms may represent "sensors" of exposure and "mediators" of the outcomes, including cancer development. Figure reprinted from Herceg et al. (2017). Roadmap for investigating epigenome deregulation and environmental origins of cancer. Int J Cancer. http://dx.doi.org/10.1002/ijc.31014 PMID:28836271. © 2017 IARC/WHO; licensed by UICC.



Identifying epigenomic signatures associated with exposures to risk factors

EGE plays a key role in several multidisciplinary studies aimed at testing the hypothesis that epigenetic changes may be risk factor-specific ("signatures") and that this may prove instrumental in the discovery of new biomarkers in cancer. By applying powerful epigenomic methodologies to unique case–control and population-based studies, EGE has led studies that resulted in important discoveries, including: (i) identification of an epigenetic signature of human papillomavirus (HPV) infection in head and neck cancer that is independent of the anatomical site (Figure 2), is functionally correlated with gene expression, and may be leveraged for improved stratification of prognosis (Degli Esposti et al., 2017b); (ii) comprehensive identification and cataloguing of the DNA methylation alterations associated with tobacco smoking (Ambatipudi et al., 2016, Joehanes et al., 2016); (iii) demonstrating that normal gastric mucosa from gastric cancer cases and healthy controls exhibits methylome-wide changes associated with current and past infection with *Helicobacter pylori*; (iv) demonstrating that specific DNA methylation changes in lung tumours are associated with asbestos exposure and identifying potential causal pathways induced by asbestos exposure (Kettunen et al., 2017); and (v) demonstrating that despite a marked reversibility of methylation changes after exposure removal (such as smoking cessation and *H. pylori* eradication), a significant number of genomic regions remained differentially methylated years later, suggesting the existence of a long-term "epigenetic memory" (Ambatipudi et al., 2016).

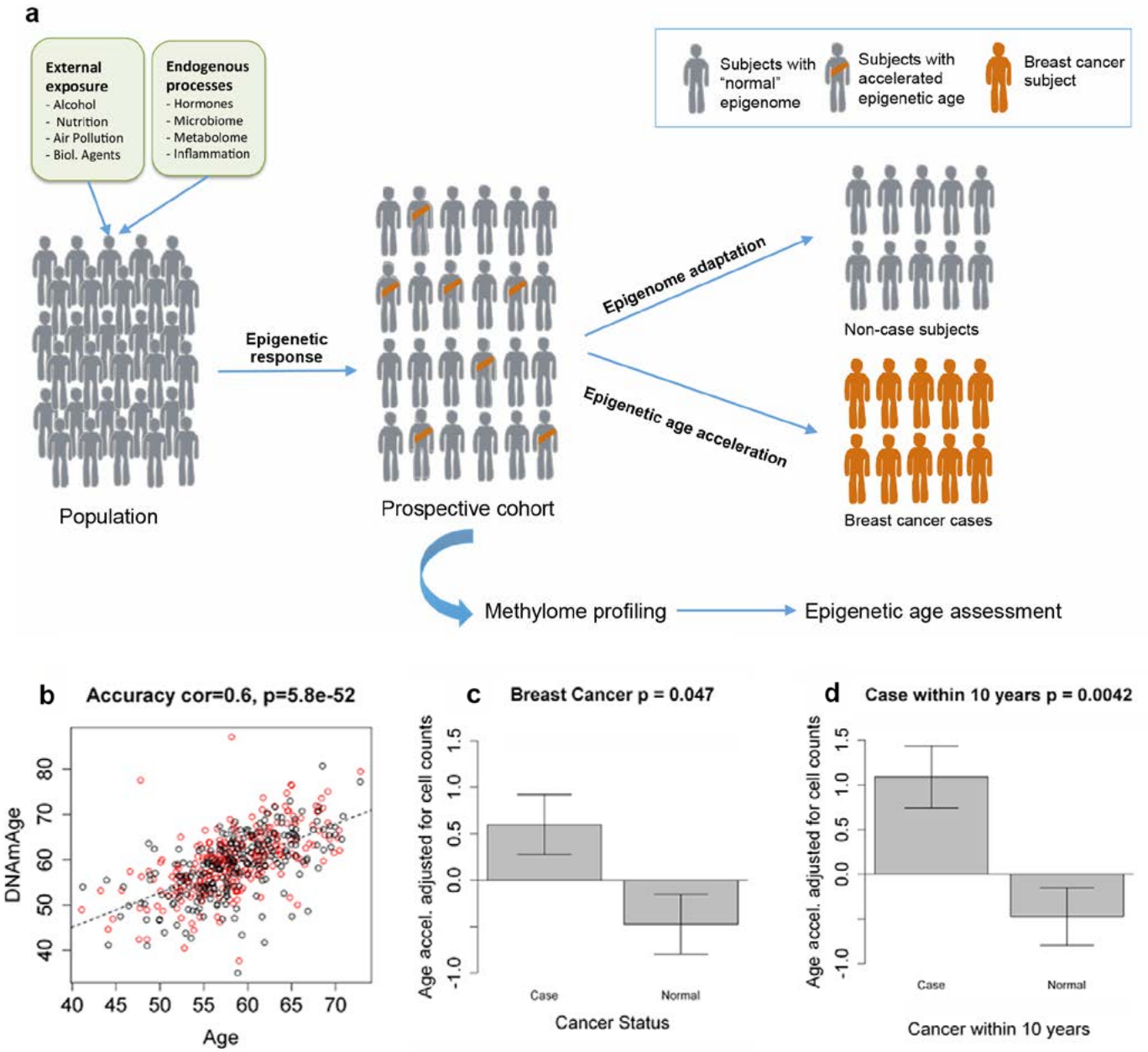
Figure 2. Human papillomavirus (HPV) infection leaves a clear DNA methylation signature in head and neck cancer. (a, b) Multidimensional scaling (MDS) plots showing sample clustering grouped by different variables: (a) HPV status; (b) organ site. (c) Heat map showing the 2410 differentially methylated positions associated with HPV status (false discovery rate < 0.05, differential methylation Δβ > 20%). Figure adapted from Degli Esposti et al. (2017b). © Degli Esposti et al., 2017.



DNA methylome-wide analysis of a prospective cohort identifies accelerated epigenetic ageing associated with cancer susceptibility

EGE coordinated a large study aimed at identifying the potential of epigenetic changes in peripheral blood as a marker of risk factor exposure and cancer risk. This approach combined the advantages of methylome-wide profiling and a large prospective cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC) study with adequate statistical power. The study revealed that higher epigenome-wide methylation at CpG islands was associated with breast cancer risk and that DNA methylation-based markers of ageing (known as the "epigenetic clock") are associated with susceptibility to postmenopausal breast cancer (Figure 3) (Ambatipudi et al., 2017). This study demonstrates that prospectively collected blood samples harbour epigenetic changes that may serve as potential markers of risk factor exposure and breast cancer risk.

Figure 3. DNA methylome-wide analysis of a prospective cohort identifies epigenetic ageing associated with cancer susceptibility. (a) The general study design. (b) DNA methylation age (vertical axis) versus chronological age (horizontal axis). Points correspond to female subjects. Red indicates a breast cancer case, and black indicates a control. The dashed line indicates a regression line. (c) Epigenetic age acceleration versus breast cancer status. Each bar plot depicts the mean value and standard deviation and reports a non-parametric group comparison test *P* value (Wilcoxon test). (d) Epigenetic age acceleration versus breast cancer status (developed within 10 years after the blood draw). Each bar plot depicts the mean value and standard deviation and reports a non-parametric group comparison test *P* value (Wilcoxon test). Figure compiled from Ambatipudi et al. (2017). © IARC.



Identifying epigenetic changes induced by in utero and early-life exposures and their causal relationship with childhood cancer

One of the major focuses of EGE in recent years has been the development of a multidisciplinary study aimed at investigating the causal relationship between in utero and early-life exposures and increased risk of cancer in childhood and adulthood. EGE played a central role in developing an epigenetic epidemiology framework at IARC and several major international consortia focusing on the early-life period. In particular, EGE manages the International Biospecimen Coordinating Center of the International Childhood Cancer Cohort Consortium (I4C), the largest prospective investigation into childhood cancer, comprising about 500 000 mother–child pairs. This has led to exciting synergies and cross-interactions among I4C, the EXPOsOMICS Consortium (Vineis et al., 2017), and the Pregnancy and Childhood Epigenetics (PACE) Consortium, all of which comprise the early-life period. Building on these large and rich data resources, EGE has started cataloguing epigenetic signatures of early-life exposures (Joubert et al., 2016) and deciphering their effects on phenotypic

outcomes during this period (Table 1), with a primary focus on childhood cancer as an end-point. The prioritized exposures include tobacco smoking and air pollution, for which the effects on childhood cancer risk remain elusive. As for the phenotypes, EGE has focused on birth weight and associated covariates (pre-pregnancy body mass index, sex, and gestational age), because high birth weight (apart from ionizing radiation) is currently the only prospectively based risk factor for all childhood cancer. The current studies focus on epigenetic precursors of childhood cancer (Table 1) that may help to decipher complex exposure-to-phenotype patterns.

Table 1. Summary of epigenetic signatures of early-life exposures, phenotypes, and cancer identified to date

Exposure/phenotype/disease	Number of newborns	Number of CpGs identified agnostically after FDR [or Bonferroni] adjustment*	Major finding	Reference
Exposures during pregnancy				
Maternal smoking	6685 (13 cohorts)	6073 [568]	Many of these CpGs span genes associated with cancer, and all persist years later, throughout childhood.	Joubert et al. (2016)
Air pollution, NO_2	1508 (4 cohorts)	3 [0]	Although only a few CpGs were identified, all represented mitochondria-related genes.	Gruzieva et al. (2017)
Air pollution, $PM_{2.5}$ Air pollution, PM_{10}	1551 (8 cohorts) 1949 (7 cohorts)	5 [0] 8 [1]	$PM_{2.5}$ and PM_{10} have different impacts on the epigenome of the newborn.	In preparation
Phenotypes perinatally				
Birth weight	8365 (28 cohorts)	8620 [1071]	Birth weight is largely associated with epigenomic variations, ~5% of which remain significant until adulthood.	In preparation
Maternal pre-pregnancy BMI	9340 (19 cohorts)	Several thousand [9044]	Only 8 CpGs are due to a direct intrauterine effect of maternal BMI; the remaining CpGs are associated with blood cell proportions, genetics, and lifestyle factors.	Sharp et al. (2017)
Gestational age	6937 (19 cohorts)	12 799 [9515]	Gestational age has a major impact on the epigenome of the newborn, but only ~1.5% of the CpGs persist until age 7–9 years. The epigenome also accurately predicts gestational age.	In preparation
Sex	Ongoing	Ongoing	Ongoing	Ongoing
Type of cancer				
Childhood leukaemia	857 (3 cohorts)	3 regional clusters, each encompassing ~15 CpGs	Large, sex-specific effects were observed, replicable in three different continents; all regions encompassed imprinted and metastable epialleles.	In preparation
Childhood central nervous system tumours	1205 (4 cohorts)	Ongoing	Common signatures exist with childhood leukaemia.	Ongoing

BMI, body mass index; FDR, false discovery rate; NO_2, nitrogen dioxide; PM_{10}, particulate matter with particles of aerodynamic diameter < 10 μm; $PM_{2.5}$, particulate matter with particles of aerodynamic diameter < 2.5 μm.

EEOC conference

In 2016, EGE organized the Epigenetics and Environmental Origins of Cancer (EEOC) conference, which brought together more than 20 leading scientists in the field and about 150 researchers from different disciplines and many countries around the globe. The leading scientists in the field reviewed the state of the science of epigenetics associated with environmental stimuli and cancer risk, highlighting key developments in the field. Critical knowledge gaps and research needs were discussed, as well as advances in epigenomics that may help in understanding the functional relevance of epigenetic alterations. The scientific exchanges during and after the meeting resulted in an opinion paper jointly written by the invited speakers, and the scientific exchanges promoted international collaboration in the field and contributed to the visibility of the Section and IARC.

Participants in the Epigenetics and Environmental Origins of Cancer (EEOC) conference, held at IARC in June 2016. © IARC/Roland Dray.

Molecular Mechanisms and Biomarkers Group (MMB)

The Molecular Mechanisms and Biomarkers Group (MMB) aims to identify critical molecular processes and markers of carcinogenesis associated with specific environmental and lifestyle risk factors, to facilitate evidence-based cancer prevention strategies. MMB focuses in particular on screening genomic alterations such as mutational signatures in experimental systems and in human and animal tissues, to reveal the impact of environmental factors on the genome and on tumour development. MMB also develops experimental methods and bioinformatics tools in this area that are applicable to molecular cancer epidemiological studies.

Multisystem approach for the identification of mutation spectra of human carcinogens

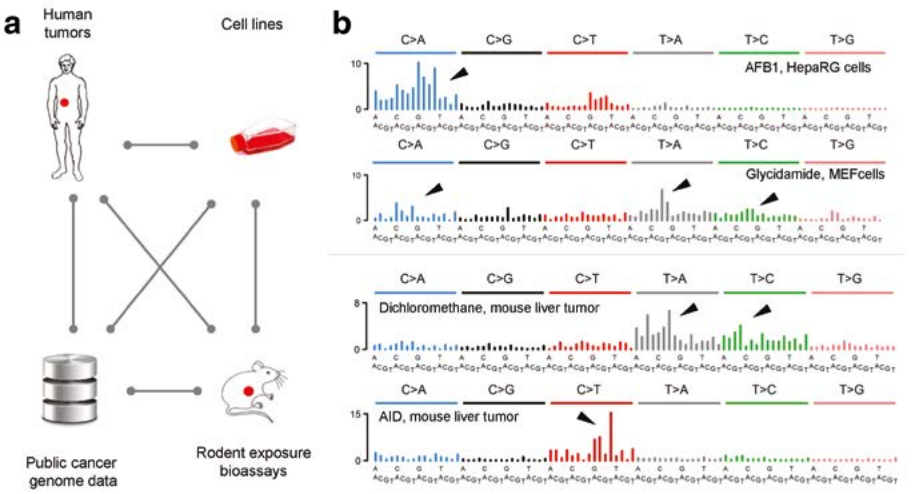
Many carcinogens are mutagenic and can induce specific alterations in the genome, in characteristic imprints that can be used to identify tumours that arise as a result of exposure to these carcinogens. MMB has devised a multisystem approach that includes genome-wide mutation screening in in vitro cell models of exposure, tissues from in vivo animal studies (such as from the United States National Toxicology Program), and tumour samples from exposed humans to characterize the genome-wide impact of several new

candidate human mutagenic factors (Figure 4).

In a recent example, MMB collaborated with the Duke-NUS Medical School, Singapore, to characterize the genome-wide mutational signatures of aflatoxin B_1 from in vivo and in vitro exposure assays. This innovative study showed that evidence of exposure to aflatoxin B_1 is present in 16% of liver cancer cases from Hong Kong Special Administrative Region, compared with 0.7% of cases from North America and 1% of cases from Japan. These results show that aflatoxin exposure apparently remains a substantial public health issue in some areas (Huang et al., 2017a). This integrative approach is thus a powerful strategy for the identification of human tumours linked to various environmental carcinogens, supporting the development of evidence-based cancer prevention measures.

Moreover, to streamline this type of analysis and make it available to the wider cancer research community, MMB has developed a user-friendly bioinformatics package for the analysis and interpretation of mutational signatures in various systems (Ardin et al., 2016).

Figure 4. Identification of mutational signatures of human carcinogens in mutually cross-validating comparison of primary tumours and experimental systems. (a) Genome-wide sequencing data are generated from human tumours, experimentally established cell culture clones, and tumour tissues from rodents experimentally exposed to carcinogens. This framework can be used to study the mechanistic effects of numerous candidate carcinogens. (b) Examples of mutational signatures newly identified by MMB. The bar graphs depict frequencies of single base substitutions for each mutation type, shown as six condensed types (C>A = C:G → A:T, C>G = C:G → G:C, etc.), in a particular trinucleotide context (top row under each graph shows preceding base, lower row shows following base). From top to bottom: whole-genome signature of the carcinogenic mycotoxin aflatoxin B_1 (AFB1), observed in the human liver cancer cell line HepaRG; exome-scale signature of glycidamide, the genotoxic metabolic product of acrylamide, established in mouse embryonic fibroblasts (MEF); exome-scale signature of dichloromethane, an industrial solvent, identified in the liver tumours of exposed mice; and whole-genome signature of activation-induced cytidine deaminase (AID), identified as a result of transgenic activity driving the development of mouse liver tumours. © IARC.



Laboratory studies to elucidate the mechanisms of carcinogen-induced cell transformation

Because of the accumulation of genetic and epigenetic alterations, oncogenic stress can result in the evasion of critical biological barriers that protect normal cells from uncontrolled division. MMB is exploiting the resulting clonal expansion of immortalized cells by applying massively parallel sequencing strategies to study the relationship between genetic alterations and epigenetic changes in the context of specific carcinogen exposures. Therefore, integrated genome-wide analyses of mutations and epigenetic features are being used to identify characteristic profiles of carcinogen exposure and cell transformation in human and murine cell models. Cell-type-of-origin chromatin structure strongly influences mutation landscapes in corresponding tumours. In a second line of experiments, in vitro cell immortalization strategies are being applied to model this interplay in controlled exposure settings (Huskova et al., 2017). Findings from these mechanistic studies may be used to better understand carcinogen-driven tumour development and provide clues on cancer etiology.

Genomic features of premenopausal breast cancer in Latin American women: the PRECAMA study

MMB is actively collaborating with the Section of Nutrition and Metabolism (NME) on the Molecular Subtypes of Premenopausal Breast Cancer in Latin American Women (PRECAMA) study (precama.iarc.fr), a multicentre population-based case–control study on breast cancer in young Hispanic women, an understudied group. MMB is screening genomic alterations in tumour samples recruited in PRECAMA to better characterize the molecular features of breast cancer in this population. Preliminary results showed that a majority of tumours were hormone receptor-positive and that *TP53* and *PIK3CA* were the most frequently mutated genes. Interestingly, some unexpected mutational signatures were observed in the *TP53* gene and exome-wide. Further omics analyses of a larger number of cases in the near future will enable the investigation of relationships between genomic characteristics and risk factors.

Multiomics analysis of urothelial tumours of patients exposed to aristolochic acid

Ingestion of *Aristolochia* herbs containing aristolochic acid leads to aristolochic acid nephropathy, which is marked by severe renal damage and formation of cancer in the upper urinary tract urothelium, in the renal cortex, and at other anatomical sites. Focusing on

the upper tract urothelial tumours, MMB applied a highly integrative multiomics approach to profile the transcriptomes, the protein levels, and the mutations at both the DNA and RNA levels in these tumours, and to investigate urinary microRNAs as markers of tumour presence. This study generated insights into complex candidate mechanisms of carcinogenesis associated with aristolochic acid, and demonstrated the suitability of urine microRNAs as non-invasive biomarkers of early recurrence of urothelial cancer in patients with aristolochic acid nephropathy (Figure 5), which is potentially applicable to non-invasive surveillance of urothelial cancer development.

Figure 5. Tumour-specific microRNAs (miRNAs) can be detected in the urine of patients with upper tract urothelial carcinoma (UTUC). (a) The heat map shows relative abundance levels of urinary miRNA in urine samples collected before (Pre-op) and after (Post-op) tumour removal surgery. (b) Scatter plots show relative abundance levels of miRNA in the tumour and normal adjacent tissues for five distinct UTUC-specific miRNAs. Error bars present the mean and the standard error of the mean. (c) Scatter plots show relative abundance levels of miRNA in the urine collected before (Pre-op) and after (Post-op) the tumour removal surgery, for five UTUC-specific miRNAs. Error bars present the standard error of measurement around the mean. © IARC.