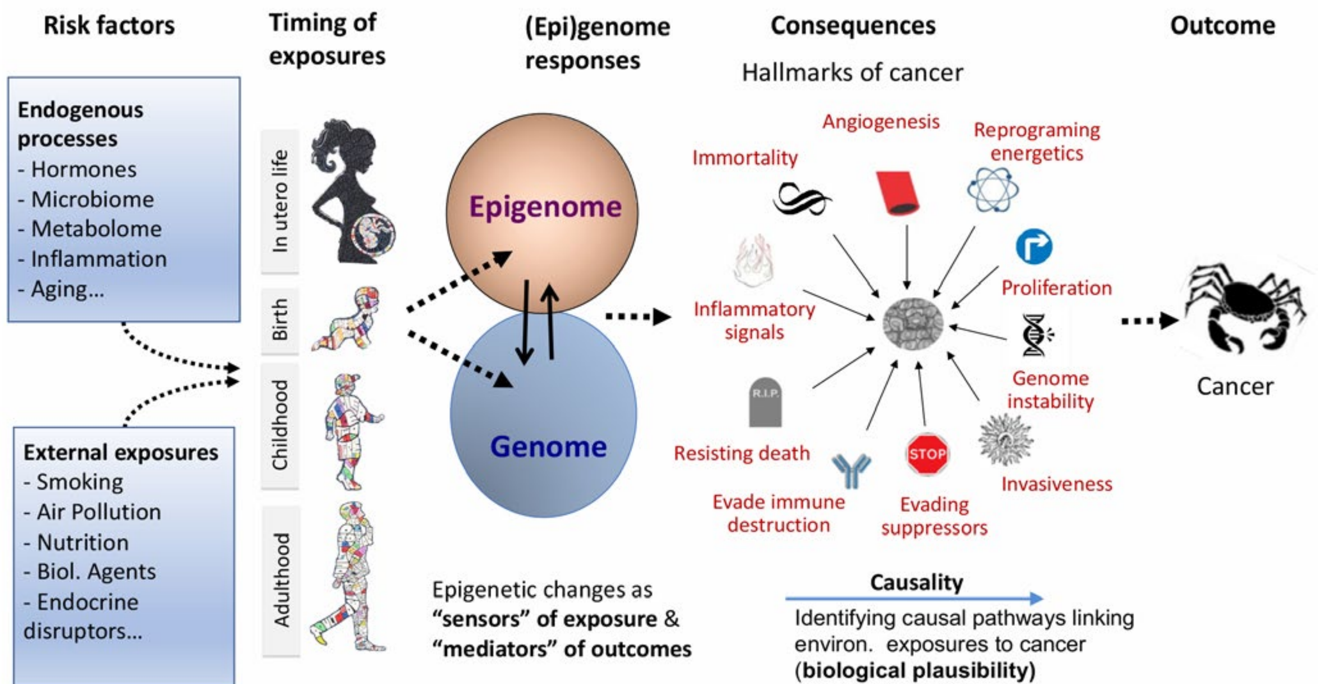

Studying (epi)genome deregulation and environmental origins of cancer
Risk factors
Timing of exposures
(Epi)genome responses
Consequences
Outcome
Hallmarks of cancer
Endogenous processes
- Hormones
- Microbiome
- Metabolome
- Inflammation
- Aging…
External exposures
- Smoking
- Air Pollution
- Nutrition
- Biol. Agents
- Endocrine disruptors…
In utero life
Birth
Childhood
Adulthood
Epigenome
Genome
Epigenetic changes as "sensors" of exposure & "mediators" of outcomes
Angiogenesis
Reprograming energetics
Immortality
Proliferation
Inflammatory signals
Genome instability
Resisting death
Invasiveness
Evade immune destruction
Evading suppressors
Cancer
Causality
Identifying causal pathways linking environ. exposures to cancer (biological plausibility)

© IARC

Section of Mechanisms of Carcinogenesis (MCA)

Section head
Dr Zdenko Herceg

Epigenetics Group (EGE)

Group head
Dr Zdenko Herceg

Scientists
Dr Rosita Accardi (died 13 May 2020)
Dr Akram Ghantous
Dr Rita Khoueiry

Secretariat
Ms Elizabeth Page

Postdoctoral fellows
Dr Zainab Awada
Dr Felicia Chung (until April 2021)
Dr Andrea Halaburkova
(until November 2020)
Dr Farah Nassar
Dr Alexei Novoloaca
(until December 2020)
Dr Grace Akinyi Odongo
Dr Alexandra Sexton-Oates
(until January 2020)
Dr Natalia Spitz Toledo Dias
Dr Fazlur Talukdar (until August 2020)

Doctoral students
Ms Maria Boskovic
Ms Mariana Gomes da Silva Araujo
Mr Antonin Jay
(until November 2020)
Ms Francesca Manara
Mr Thanos Mouchtaris-Michailidis
Mr Alexei Novoloaca
(until June 2020)
Ms Anna-Luiza Vicente
(until May 2020)

Research assistants
Mr Cyrille Cuenin
Ms Aurélie Salle

Senior research assistant, data management/analysis
Mr Vincent Cahais

Trainees
Ms Cédrine Milesi
(until September 2020)
Ms Farah Ramadan
(until December 2020)
Ms Ling Wei (until March 2020)

Molecular Mechanisms and Biomarkers Group (MMB)

Group head
Dr Jiri Zavadil

Scientists
Dr Michael Korenjak
Dr Magali Olivier (until August 2020)

Senior research assistant, data management/analysis
Dr Claire Renard

Senior research assistant
Ms Marie-Pierre Cros
(until June 2021)

Secretariat
Ms Karine Racinoux

Postdoctoral fellow
Dr Samrat Das

Students
Ms Liesel Claeys
(until November 2021)
Ms Shefali Thakur

Trainees
Ms Bérénice Chavanel
(until August 2021)
Ms Xinyue Jiang (until August 2020)
Mr Stéphane Keita
(until August 2021)

Identifying the causal pathways that link environmental or lifestyle exposures to tumorigenesis and gaining insights into the molecular mechanisms that underlie the associations observed in epidemiological studies (biological plausibility) provide a foundation for studies of cancer etiology, carcinogen evaluation and classification, and ensuring evidence-based cancer prevention. The Section of Mechanisms of Carcinogenesis (MCA) conducts hypothesis-driven and data-driven studies aimed at advancing the understanding of cancer causation and the mechanisms of tumorigenesis, as well as promoting international collaborations, as the core mission of IARC.

The studies of MCA are interdisciplinary in nature, and the major MCA programmes promote and advance synergistic collaborations with other IARC scientists and evidence synthesis experts. Key elements of MCA's strategy include developing innovative state-of-the-art molecular and cell biology and functional epigenomics research methodologies, and bioinformatics and biostatistics tools, applicable to experimental cancer models and human samples from population-based studies. The Section comprises two Groups – the Epigenetics Group (EGE) and the Molecular Mechanisms and Biomarkers Group (MMB) – whose corresponding research programmes are complementary with respect to methodological approaches and the shared ultimate objective of identifying causal links between environmental factors and cancer. With the start of the new IARC Medium-Term Strategy 2021–2025 and the new organizational structure as of 1 January 2021, MCA was renamed as the Epigenomics and Mechanisms Branch.

Epigenetics Group (EGE)

The overarching aim of the Epigenetics Group (EGE) is to advance the understanding of the role of epigenetic changes and pathways induced by environmental factors and endogenous processes in cancer causation, underpinning studies of etiology, carcinogen evaluation, and prevention. EGE exploits new concepts in cancer epigenetics, the availability of unique population-based cohorts, and recent technological advances in epigenomics (Halaburkova et al., 2020; Pashayan et al., 2020; Ghantous et al., 2021; Sklias et al., 2021). 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 (Merid et al., 2020; Karabegović et al., 2021; Sorroche et al., 2021; Talukdar et al., 2021).

Epigenome-wide profiling of oesophageal squamous cell carcinoma from high-incidence regions identifies crucial genes and potential cancer markers

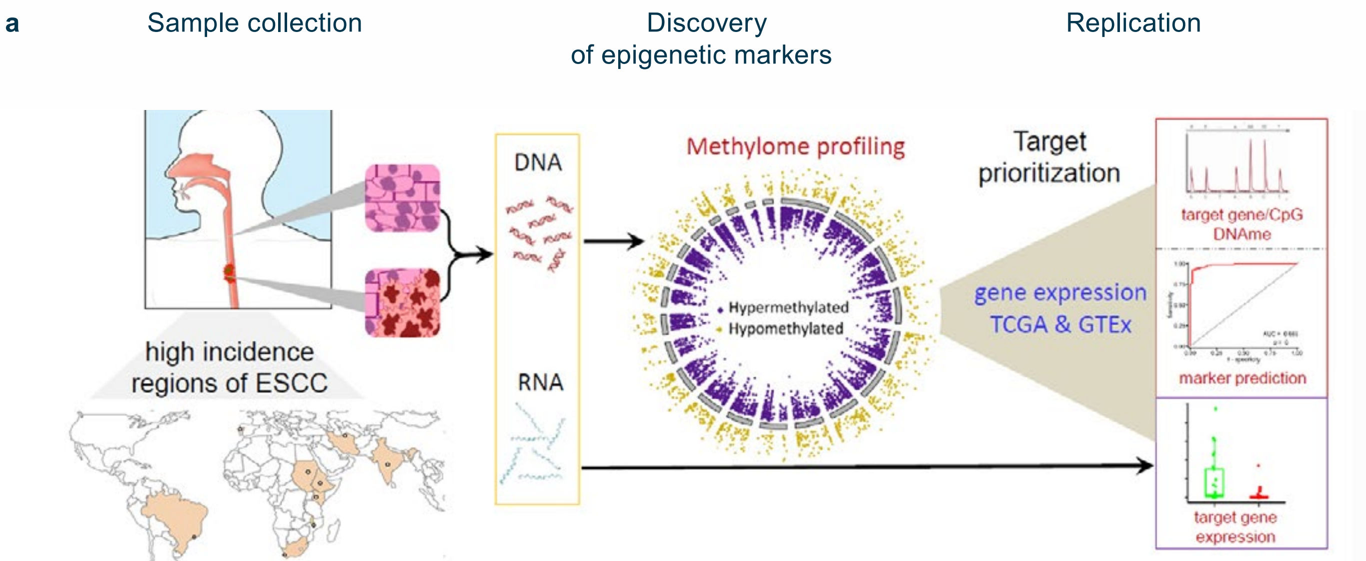
Oesophageal squamous cell carcinoma (ESCC) is one of the most aggressive and lethal forms of cancer in the world, with the highest incidence rates in low- and middle-income countries. EGE conducted the largest epigenome-wide (DNA methylome, DNAme) profiling of the collection of ESCC samples from high-incidence populations worldwide (Figure 1), with the aim of understanding ESCC etiology and identifying early biomarkers. DNAme changes in ESCC samples and normal tissue adjacent to the tumours (NAT) from patients with cancer in nine high-incidence countries in Africa, Asia, and South America were studied by using the Infinium MethylationEPIC array (Talukdar et al., 2021). Methylome analysis comparing tumour tissue and NAT identified 6796 differentially methylated positions (DMPs) and 866 differentially methylated regions (DMRs). Most of the identified DMPs and DMRs were hypermethylated in tumours. Top genes identified in the discovery phase were prioritized for validation, and their putative functional impact on gene transcription was analysed using RNA-seq data from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) project. The specificity and sensitivity of these DNAme events in discriminating tumours from NAT were then assessed. EGE identified and prioritized genes and pathways involved in the development of ESCC, and proposed an early detection marker panel, which could serve as a reference for tests to improve the early detection of this cancer type in low-resource settings (Talukdar et al., 2021).

Epigenetic markers of breast cancer risk in a prospective cohort study

Epigenetic alterations are a near-universal feature of malignancy; however, much of the current evidence is based on findings in retrospective studies, which may reflect epigenetic patterns influenced by the onset of the disease. Studying breast cancer, EGE established genome-scale DNA methylation profiles of prospectively collected buffy coat samples from a case–control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) study using reduced representation bisulphite sequencing (RRBS) (Figure 2). For a subset of these individuals, EGE also profiled primary tumours and tumour-adjacent tissue samples. EGE observed cancer-specific DNA methylation events in both the breast tissue and buffy coat samples, each characterized by sample type-specific differences but with a shared enrichment for genes in specific biological pathways. Notably, increased DNA methylation in genomic regions associated with specific genes was linked to the length of time to diagnosis in prospectively collected buffy coat DNA from individuals who subsequently developed breast cancer

(Figure 2). Using machine learning methods, EGE piloted a DNA methylation-based classifier that predicted case–control status in a held-out validation set with high accuracy, in some cases up to 15 years before diagnosis. The findings suggest a model of gradual accumulation of cancer-associated epigenetic patterns in peripheral blood, which may be detected long before the clinical manifestation of cancer. Such changes may provide useful markers for risk stratification and, ultimately, personalized cancer prevention.

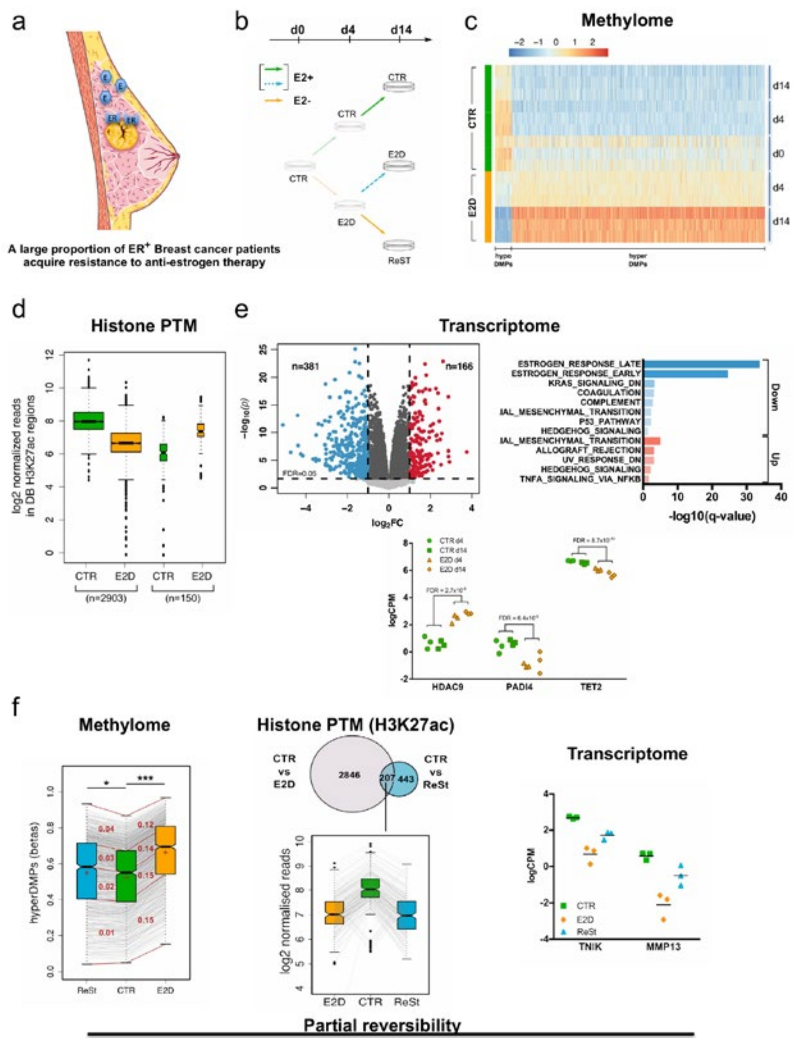
Figure 1. Genome-wide DNA methylation profiling of oesophageal squamous cell carcinoma (ESCC) from high-incidence populations of the world enables the identification of functionally relevant and robust DNA methylation markers for early detection in low-resource settings. (a) Overview of study design and sample characteristics. Country-wise sample distribution in percentages. Dots in the map showing sample collection sites and their respective countries are coloured. Reprinted from Talukdar et al. (2021), with permission from the American Association for Cancer Research. (b) Collaborators attending the Oesophageal Squamous Cell Carcinoma African Prevention research (ESCCAPE) network meeting (coordinated by Dr Valerie McCormack, IARC) held in Eldoret, Kenya. © IARC.



b

Figure 2. Identifying epigenetic markers of breast cancer risk in a prospective cohort (EpiMark study). (a, b) Schematic of the study design and analytical methods. (a) A nested case–control study was constructed within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Blood samples and lifestyle information were collected from apparently healthy participants upon enrolment. Participants who were diagnosed with breast cancer during the follow-up period were matched on a one-to-one basis with individuals who were cancer-free over the follow-up period. Buffy coat lysates derived from blood samples collected at enrolment as well as tumour tissue and adjacent normal tissue from cases were analysed by reduced representation bisulphite sequencing (RRBS). (b) Five-fold nested cross-validation was used to train and evaluate classifier models for their ability to discriminate individuals who developed breast cancer from those who were cancer-free over the follow-up period. The best-performing model was selected and trained on the full cross-validation data set of 340 pairs, to finalize model parameters. The final model was used to predict case–control status in a held-out validation set of 68 matched pairs. Differential methylation analyses and functional enrichment analyses were conducted in parallel. (c, d) The global DNA methylation landscape in breast tumours, and breast cancer-specific DNA methylation patterns in breast tissue. (c) Unsupervised clustering of methylation in the 6713 most variable regions according to the *k*-means algorithm and Euclidean distance. (d) Four consensus clusters of tumour tissue were identified and are denoted as clusters 1–4, in order of overall DNA methylation levels across CpG islands, gene bodies, and promoters compared with DNA methylation at the same sites in tumour-adjacent normal tissue (NT). (e–h) Prediction of case–control status in prospectively collected blood tissue using a prediction analysis for microarrays (PAM) classifier. (e) The receiver operating characteristic (ROC) curve and the corresponding area under the curve (AUC) statistics for the PAM classifier applied on the validation cohort, against a background of 100 label-shuffled control data sets that were subjected to the same model training and testing process. A *t*-distributed stochastic neighbour embedding (*t*-SNE) plot was generated using the 49 genomic regions used in the PAM classifier, coloured by (f) case–control status and (g) length of time from sample collection to diagnosis (by quartile). (h) Schematic of the classification results from the final PAM model on the held-out validation set alongside length of time to diagnosis (quartiles). (i) Diagram of documented functional and physical interactions between possibly perturbed transcription factors in both blood and tissue differentially methylated region (DMR) sets. © IARC.

Figure 3. Understanding estrogen receptor (ER) pathway regulation in breast cancer cells and revealing potential mechanisms underlying the roots of resistance to anti-estrogen therapy. (a) Breast cancers are classified into different molecular subtypes mainly according to the presence of ER, progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2), and their expression tends to determine the treatment approach. Most (70%) breast cancers are ER-positive (ER+) and are subjected to anti-estrogen therapy. However, patients regularly develop a non-reversible resistance to anti-estrogen therapy. (b) In vitro optimized protocol adapted for studying estrogen deprivation and re-stimulation using MCF-7 breast cancer cells. MCF-7 cells cultured in control conditions (CTR, charcoal-stripped serum + estradiol E2) were deprived of E2 for 4 and 14 days (E2D) or deprived of E2 for 4 days then re-stimulated with E2 for 10 days (ReSt, blue dashed line). (c) Heat map of differentially methylated positions (DMPs) between CTR and E2-deprived (E2D) MCF-7 cells with at least 10% differential methylation (Δβ) detected after methylome analyses using an 850K array, which measures the DNAm levels across more than 850 000 CpGs. (d) Analyses of histone post-translational modification (PTM) by ChIP-sequencing of H3K27 acetylation (H3K27ac), a mark associated with enhancer regions. Analysis showed a global decrease of histone acetylation in E2D compared with CTR. (e) Results of transcriptome analysis (RNA-sequencing). Upper left: Distribution of $-\log_{10}(P)$ of differentially expressed genes (DEGs) in E2D versus CTR according to $\log_2$ fold change (FC) of expression. Coloured dots represent downregulated and upregulated DEGs with an absolute $\log_2$(FC) > 1 (blue and red) and a false discovery rate (FDR) < 0.05 (dashed horizontal). DEGs that were differentially expressed with an absolute $\log_2$(FC) < 1 are shaded dark grey. Upper right: Gene set enrichment analysis of downregulated and upregulated DEGs (MSigDB, database H, hypergeometric test). Lower: Significant differential expression of three epigenetic remodelling factors, HDAC9, PADI4, and TET2, in response to E2 deprivation. (f) Partial reversibility of epigenomic and transcriptomic changes after E2 deprivation and re-stimulation. Left: Distribution of hypermethylated DMPs in response to E2 deprivation and re-stimulation for CTR, E2D, and ReSt at day 14 (CTR vs E2D, *n* = 950; FDR < 0.05, Δβ > 10%). Box plot: centre lines, median (Q2); box boundaries, 25% and 75% quartiles (Q1 and Q3); top and bottom whiskers, minimum and maximum (Q0 and Q4). For each pairwise comparison (ReSt–CTR and CTR–E2D), the quartiles are connected with red lines. In each interquartile range, the mean Δβ between the compared interquartile groups is shown in red. The mean of each group is shown with a red cross. Asterisks mark significant differences of ReSt and E2D means compared with CTR. Upper centre: Overlap of differential H3K27ac regions between CTR versus E2D and CTR versus ReSt regions, showing 207 common peaks. Lower centre: H3K27ac signal in $\log_2$-normalized reads of the 207 peaks. Right: Non-reversibility of expression of AP-1 transcription factor inducer TNIK and AP-1 target gene MMP13 for CTR condition (green squares), E2D (orange diamonds), and ReSt (blue triangles) on day 14. Differential expression analysis was performed after RNA-sequencing by contrasting CTR versus ReSt groups among E2-deprivation DEGs with FDR < 0.05 and $|\log_2 FC|$ > 1. Reproduced from Sklias et al. (2021). © 2021, Oxford University Press.



Epigenetic changes induced by estrogen hormones as a potential mechanism underlying endocrine resistance in ER-positive breast cancer

Estrogen hormones are implicated in the development of most breast cancers, and estrogen receptor (ER) alpha, the main nuclear factor that mediates estrogen signalling, orchestrates a complex molecular circuitry, which is poorly understood. EGE combined a novel in vitro protocol adapted for studying estrogen deprivation and re-stimulation with the latest epigenomics and bioinformatics tools, which enabled a genome-wide interrogation of the epigenome and transcriptome changes associated with modifications in ER pathways. The results showed that prolonged estrogen deprivation and re-stimulation result in time-dependent epigenetic changes across diverse genomic regions and changes in gene expression associated with specific biological pathways (Figure 3). Remarkably, many of the observed changes upon estrogen deprivation were also detected in breast cancer cells that developed resistance in response to anti-estrogen therapy (Sklias et al., 2021). Finally, the study revealed a selective reversibility and persistence of epigenetic and gene transcription changes observed after estrogen deprivation and re-stimulation, suggesting a potential mechanism underlying the roots of endocrine resistance that develops in response to anti-estrogen therapy (Figure 3) (Sklias et al., 2021).

Molecular Mechanisms and Biomarkers Group (MMB)

The overarching objective of the Molecular Mechanisms and Biomarkers Group (MMB) is to improve the knowledge base for mechanistic molecular studies of modifiable cancer causes and for relevant cancer prevention measures. Innovative experimental approaches assist in the discovery of molecular cancer markers (Melki et al., 2020). MMB conducts molecular cancer epidemiological studies (Karabegović et al., 2021; the MODARC study on the role of dietary acrylamide in renal carcinogenesis; the EVAMOVAIRE2 study) and participates in IARC's carcinogen evaluation (Samet et al., 2020) and cancer classification efforts (Cree et al., 2021a). During the biennium, MMB performed integrative toxicogenomic analyses of selected candidate carcinogens and their roles in oncogenesis (Claeys et al., 2020). MMB also collaborated with the IARC Monographs on systematic cancer hazard assessments (Barupal et al., 2021).

Evamovaire2: mutational signatures of asbestos exposure in ovarian tumours

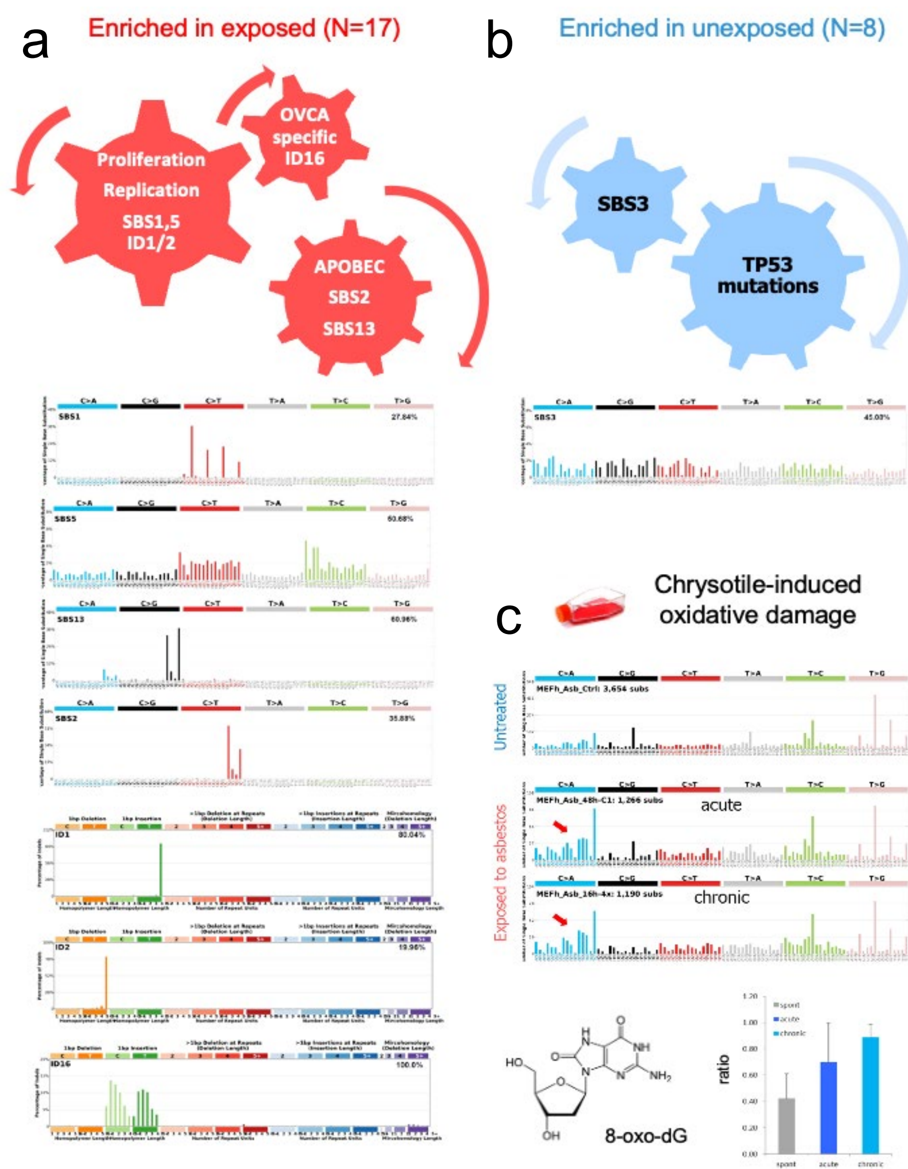
Epidemiological studies of geographical variations in ovarian cancer incidence have suggested a causal role for environmental factors, prompting the IARC Monographs to classify asbestos fibres as an ovarian carcinogen in 2009. MMB explored the link between asbestos exposure and ovarian cancer histological subtypes by integrating epidemiological data, exposure assessment, and whole-genome sequencing to determine the potentially carcinogenic effects of exposure to asbestos. Among 254 patients with ovarian cancer in the study, 13.4% had been exposed to asbestos occupationally and 16.5% had possibly been exposed indirectly, via a close relative. The direct exposure prevalence appeared to be higher than in the general population.

MMB conducted whole-genome sequencing of tumour–normal tissue pairs of 25 cases with established exposure mode, probability, and levels. Several exposure-specific mutational signatures were observed, all of an endogenous nature, alongside a signature of BRCA1/2 deficiency and lower rates of the *TP53* gene mutations associated with tumours of unexposed patients (Figure 4). Chrysotile asbestos treatment of cell model systems induced prominent, dose-dependent cytotoxicity and genome-wide mutagenesis indicative of oxidative DNA damage (Figure 4). The EVAMOVAIRE2 study has provided new insights into environmental and occupational exposure to asbestos as an important risk factor for ovarian carcinogenesis.

Genomic DNA damage induced by tobacco-specific nitrosamines

Epidemiological studies have linked tobacco use to numerous cancer types, including cancer of the lung, oral cavity, pharynx, larynx, oesophagus, pancreas, urinary bladder, and liver. The tobacco-specific nitrosamines (TSNAs) 4-methylnitrosamino-1-(3-pyridyl)-1-butanone (NNK) and *N′*-nitrosonornicotine (NNN) are recognized human carcinogens (IARC Group 1). However, the complex composition of tobacco smoke and a lack of molecular exposure markers mask the precise roles of TSNAs in human oncogenesis. NNK and NNN might contribute to the mutagenic effects of tobacco smoke, but the mutational signatures of TSNAs have not yet been described.

Figure 4. Mechanisms of asbestos-associated ovarian carcinogenesis investigated in 25 patients with epithelial ovarian cancer, using whole-genome sequencing of DNA derived from the formalin-fixed, paraffin-embedded (FFPE) tumour specimens. (a) In the included exposed patients (*n* = 17), enrichments of mutational signatures were observed, indicating increased cell proliferation (SBS1, SBS5), fingerprints of APOBEC-driven mutagenesis (SBS2, SBS13), higher rates of DNA replication errors (ID1, ID2), and ovarian-specific signature ID16. (b) Mutational signature SBS3 and TP53 mutations were enriched in the unexposed patients compared with the exposed patients. (c) Acute and chronic effects of chrysotile asbestos treatment on mutagenesis in cultured cells manifested through elevated mutational signature SBS18 (red arrows) formed by oxidated deoxyguanosines (8-oxo-2′-deoxyguanosine, 8-oxo-dG), consistent with increased exposure-related oxidative DNA damage. The top panel shows a background signature SBS17, and the bar graph shows the increase in the SBS18 versus SBS17 ratio after treatment with chrysotile. © IARC.



MMB characterized such mutational signatures induced by TSNAs in a human lung cell line and in a rat bioassay. DNA adduct analysis revealed major damage on thymidine (O2-POBdT) and guanine (7-POBG) residues (Figure 5). Genome-scale sequencing of TSNA cell clones and of rat tumours yielded highly similar mutational signatures (Figure 5), indicating the convergent effects of the two related nitrosamine compounds. The hallmark T > N mutations enriched on the untranscribed strand are consistent with the thymidine damage via O2-POBdT adduct formation. The newly identified signatures provide a valuable molecular marker for follow-up in silico studies of the mutagenic effect of TSNA exposure in thousands of human cancer genomes, to ultimately address the contribution of TSNAs to the mutation spectra of tobacco-related cancers.

Figure 5. (a) Levels of pyridoxobutyl DNA adducts in the A549 lung cell line exposed to 25 µM or 50 µM 4-[(acetoxymethyl) nitrosamino]-1-(3-pyridyl)-1-butanone (NNKOAc) and in rat oesophageal tumour tissue from animals treated with 4 ppm or 8 ppm *N′*-nitrosonornicotine (NNN) in the drinking-water (*n* = 3). The DNA adducts analysed are shown on the left. O6-POBdG was not measured in 4 ppm NNN-treated rat samples. DNA was isolated and analysed by liquid chromatography–tandem mass spectrometry (LC-MS/MS). (b) Mutational signature analysis in cells and animal tumours, showing the 192-channel mutational signatures extracted by non-negative matrix factorization from the NNKOAc-exposed A549 lung cell line (top panel) and from the NNN-induced rat oesophageal tumours. © IARC.