EMBL

EMBL Conference

Cancer genomics

15 – 17 November 2023 | EMBL Heidelberg and Virtual

#EMBLCanGen

Abstracts of papers presented at the

EMBL Conference: Cancer Genomics

EMBL Advanced Training Centre, 15 - 17 Nov 2023

Scientific Organisers:

Isidro Cortes Ciriano
EMBL-EBI, United Kingdom

Christina Curtis
Stanford University, United States of America

Núria López-Bigas
ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Young Seok Ju
Korea Advanced Institute of Science and Technology, Republic of Korea

Conference Organiser:

Christopher Stocks
EMBL Heidelberg, Germany

Cover Illustration and Poster by
Christina Hof, EMBL Heidelberg

Layout and abstract book preparation by
Maria Mercedes Bacadare Goitia, EMBL Heidelberg

Thank you to our sponsors and media partners

Silver sponsors

Novogene

SCALE biosciences
CELL TO INSIGHT

Bronze sponsors

Parse BIOSCIENCES

SEEKGENE

tebubio
Facilitators of Life Sciences Research

Event sponsors

ACTIVE MOTIF
Enabling Epigenetics Research

BGI

VectorBuilder.com

vizgen

WATCHMAKER GENOMICS

Media partners

Molecular Oncology, a FEBS Press journal

International Union of Biochemistry and Molecular Biology

EMBL

The European Molecular Biology Laboratory wishes to thank the members of the EMBL Advanced Training Centre Corporate Partnership Programme.

Their support makes it possible to host conferences and courses on emerging topics and it enables us to continue to offer excellent scientific quality of our events.

If your organisation would like to receive information about the EMBL Advanced Training Centre Corporate Partnership Programme, please contact Nadine Ilk at nadine.ilk@embl.de.

Corporate Partners

10x GENOMICS

BD

BIO-RAD

Boehringer Ingelheim

eppendorf

EVIDENT
SEEING IS SOLVING

GSK

Leica
MICROSYSTEMS

NetApp

Nikon

resolve
biosciences

Roche

STILLA

ThermoFisher
SCIENTIFIC

vizgen

Associate Partners

ECLIPSE
BIOINNOVATIONS

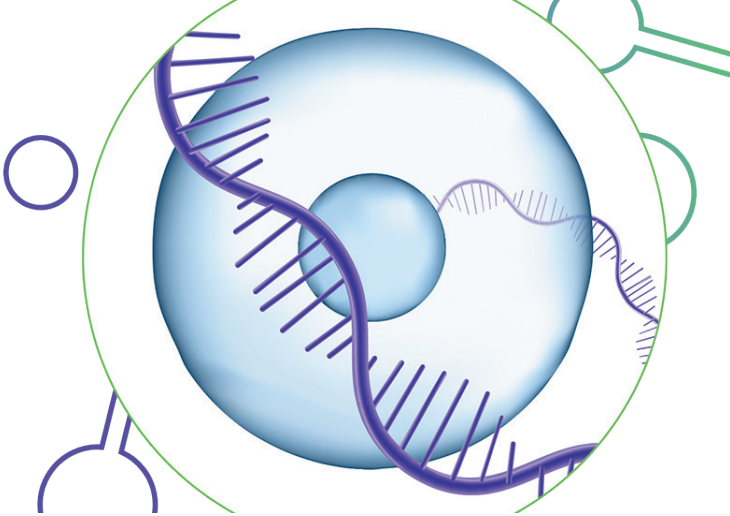
MERCK

NEW ENGLAND
BioLabs Inc.

Promega

sanofi

SARTORIUS



Welcome to the new single-cell omics.

At ScaleBio, we leverage the cell as the reaction compartment and use combinatorial indexing to perform cost-effective, instrument-free library prep on 100s of thousands of cells.

We currently support scRNAseq, scATACseq, scMethylation, and CROP-seq CRISPR screening applications.

Join us at booth #9 to learn more!

SCALE biosciences

Thank you to the EMBL Advanced Training Centre Corporate Partnership Programme for funding EMBL Conference fellowships for the following recipients:

Corleone, Giacomo	IRCCS Regina Elena National Cancer Institute, Italy
Fernández-Sanromán, Ángel	Spanish National Cancer Research Centre (CNIO, Spain
Fonseca, Adriana	University of Cambridge, United Kingdom
Harbers, Luuk	Karolinska Institute, Sweden
Mukherjee, Nivedita	National Centre for Biological Sciences, India
Paassen, Irene	Princess Máxima Center for Pediatric Oncology, The Netherlands
Pablo-Fontecha, Veronica	Institut d'investigacions biomèdiques august pi i sunyer (IDIBAPS), Spain
Pensch, Raphaela	Uppsala University, Sweden
Rajamani, Anantharamanan	Institute of Translational Cancer Research and Experimental Cancer Therapy, Klinikum rechts der Isar, Technical University of Munich, Germany
Rueda, Bertha	Universidad Nacional Autónoma de México / Instituto Nacional de Medicina Genómica, Mexico
Tomkova, Marketa	University of Oxford, United Kingdom
Volakhava, Anastasiya	CEITEC Masaryk University, Czech Republic

Wednesday 15 November 2023

10:30 - 12:00	**Registration and light refreshments** ATC Registration Desk	
10:45 - 11:45	**Pre-conference workshop "Single Cell Genomics at scale: Unlock single-cell workflows with combinatorial indexin" by Scale Biosciences** Klaus Tschira Auditorium	
12:00 - 12:15	**Opening remarks** Klaus Tschira Auditorium	
12:15 - 13:00	**Keynote lecture: Harnessing cancer evolution for clinical benefit: lessons from the oesophagus** Rebecca Fitzgerald *University of Cambridge, United Kingdom*	1
13:00 - 15:05	**Session 1: Mutation & clonal selection** **Chairs:** Núria López-Bigas – Barcelona Institute for Research in Biomedicine, Spain, Anna Poetsch – TU Dresden, Germany Klaus Tschira Auditorium	
13:00 - 13:25	**To be presented on-site** Christina Curtis *Stanford University, United States of America*	2
13:25 - 13:40	**Evolutionary dynamics of large-scale DNA copy number alterations in patient-derived adenoma organoids** Thomas van Ravesteyn *Hubrecht Institute, The Netherlands*	3
13:40 - 13:55	**Selective pressures of platinum compounds shape the evolution of therapy-related myeloid neoplasms** Jurrian de Kanter *Princess Máxima Center for Pediatric Oncology and Oncode Institute, The Netherlands*	4

13:55 - 14:20	**Reconstructing the embryology of childhood cancer from somatic mutations** Sam Behjati *Wellcome Sanger Institute, United Kingdom*	5
14:20 - 14:35	**Accurate de novo detection of somatic mutations in single-cell genomics and transcriptomics data** Francesc Muyas *EMBL-EBI, United Kingdom*	6
14:35 - 14:50	**Sperm sequencing reveals extensive positive selection in the human germline** Matthew Neville *Wellcome Sanger Institute, United Kingdom*	7
14:50 - 15:05	**Single-mitosis dissection of acute and chronic DNA mutagenesis and repair** Paul Ginno *German Cancer Research Center (DKFZ), Germany*	8
15:05 - 15:35	**Coffee break & meet the speakers** Auditorium Foyer	
15:35 - 17:35	**Session 2: Cancer Evolution** **Chairs:** Christina Curtis - Stanford University, USA Craig Anderson - MRC Institute of Genetics and Cancer, UK Klaus Tschira Auditorium	
15:35 - 16:00	**On the evolutionary history of metastatic cancer** Kamila Naxerova *Harvard Medical School, United States of America*	9
16:00 - 16:15	**The origin and evolution of pediatric secondary neoplasms** Abel Gonzalez-Perez *Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain*	10
16:15 - 16:40	**Darwinian epigenetic evolution and non-Darwinian cellular plasticity in colorectal cancer** Andrea Sottoriva *Human Technopole, Italy*	11

16:40 - 16:55 **Methylation dynamics in the decades preceding Acute Myeloid Leukaemia** 12
Adriana Fonseca
University of Cambridge, United Kingdom

16:55 - 17:10 **Inferring genomic trajectory of uveal melanoma** 13
Chang Hyun Nam
Korea Advanced Institute of Science and Technology, Republic of Korea

17:10 - 17:35 **Tumor evolution: a real-time and single-cell perspective** 14

Hugo JG Snippert
University Medical Center Utrecht, The Netherlands

17:35 - 18:05 **Flash talks**
Klaus Tschira Auditorium

18:05 - 19:35 **Poster session 1 (odd numbers)**
Helix A

19:35 - 21:00 **Dinner in Canteen**
EMBL Canteen

21:00 - 22:30 **Drinks and networking**
Auditorium Foyer

Thursday 16 November 2023

09:15 - 12:05	**Session 3: Cancer Genome Medicine** **Chair:** Young seok Ju - Korea Advanced Institute of Science and technology, South Korea Carolin Sauer - EMBL-EBI, UK Klaus Tschira Auditorium	
09:15 - 09:40	**Convergence of machine learning and genomics for precision cancer medicine** Eliezer Van Allen *Harvard Medical School, United States of America*	15
09:40 - 09:55	**Saturation mutational scanning uncovers druggability of all FGFR kinase point mutations** Carla Schmidt *University of Freiburg and German Cancer Consortium (DKTK), Germany*	16
09:55 - 10:20	**Understanding cancer biology through multi-omics genotype-phenotype tumour maps: applications in rare cancers** Lynnette Fernandez Cuesta and Matthieu Foll *IARC, France*	17
10:20 - 10:35	**Genomic analysis of skin cancers from Xeroderma Pigmentosum subgroups revealed new mechanisms of UV mutagenesis** Sergey Nikolaev *Gustave Roussy Cancer Campus, France*	18
10:35 - 11:00	**Coffee break** Auditorium Foyer	
11:00 - 11:25	**Genesis of uterine leiomyomas; lessons from a benign neoplasm** Lauri Aaltonen *University of Helsinki, Finland*	19

11:25 - 11:40	**Somatic xeroderma pigmentosum group D (XPD) mutations alter local and global landscape of mutational processes in bladder cancer** Jayne Barbour *The University of Hong Kong, Hong Kong*	20
11:40 - 12:00	**The genomic and transcriptomic profile of melanoma in Latin American patients** Carla Daniela Robles-Espinoza *Universidad Nacional Autónoma de México, Mexico*	21
12:05 - 13:30	**Meet the speakers and lunch** Auditorium Foyer	
13:30 - 15:40	**Session 4: Emerging technologies in cancer genomics** **Chairs:** Isidro Cortes Ciriano - EMBL-EBI, UK Claudia Arnedo-Pac - University of Cambridge, UK Klaus Tschira Auditorium	
13:30 - 13:55	**Clonal selection in normal tissues and its relation to cancer risk** Núria López-Bigas *ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain*	22
13:55 - 14:20	**Genomic landscapes of pre-malignant lesions in stomach cancer** Patrick Tan *Duke-NUS Medical School, Singapore*	23
14:20 - 14:35	**Cancer patient-derived organoids dependency map** Carmen Herranz-Ors *Wellcome Sanger Institute, United Kingdom*	24
14:30 - 15:00	**Charting the regulatory landscape of tumors using single-cell sequencing and spatial omics** Oliver Stegle *EMBL Heidelberg, Germany*	25

15:00 - 15:15 **Single-strand mismatch and damage patterns in cancer and cancer-predisposition syndromes revealed by single-molecule DNA sequencing** 26

Gilad Evrony
New York University Grossman School of Medicine, United States of America

15:15 - 15:40 **Characterizing the cellular, molecular, and genetic heterogeneity of solid tumors using in situ sequencing** 27

Mats Nilsson
Stockholm University, Sweden

15:40 - 16:10 **Coffee break**
Auditorium Foyer

16:10 - 17:00 **Session 5: Genomic instability & mutational processes**
Chairs: Martin Taylor - University of Edinburgh, UK
Axel Rosendahl Huber - Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
Klaus Tschira Auditorium

16:10 - 16:35 **Extrachromosomal DNA in cancer** 28
Anton Henssen
Charité - University Medicine Berlin; Max Delbrück Center for Molecular Medicine, Germany

16:35 - 17:00 **Chromothripsis in cancer** 29
Aurélie Ernst
German Cancer Research Center (DKFZ), Germany

17:00 - 17:30 **Flash talks**

17:30 - 19:00 **Poster session 2 (even numbers)**
Helix A

19:00 - 21:00 **Conference dinner**
EMBL Canteen

21:30 - 00:00 **Live band: The Chaotics**
Auditorium Foyer

Friday 17 November 2023

09:00 - 12:00	**Session 5 continued: Genomic instability & mutational processes** **Chairs:** Martin Taylor - University of Edinburgh, UK Claudia Arnedo-Pac - University of Cambridge, UK Klaus Tschira Auditorium	
09:00 - 09:25	**To be presented on-site** E. Alice Lee *Harvard University, United States of America*	30
09:25 - 09:40	**Geographic variation of mutagenic exposures in kidney cancer genomes** Sarah Moody *Wellcome Sanger Institute, United Kingdom*	31
09:40 - 10:05	**Impact of chromatin architecture on somatic mutational processes** Sabarinathan Radhakrishnan *National Centre for Biological Sciences, India*	32
10:05 - 10:20	**CpG>TpG mutagenesis: DNA polymerase errors as a new culprit of an old crime** Marketa Tomkova *University of Oxford, United Kingdom*	33
10:20 - 10:50	**Coffee break** Auditorium Foyer	
10:50 - 11:15	**Forecasting oncogene amplication** Geoff Macintyre *Spanish National Cancer Research Centre (CNIO), Spain*	34
11:15 - 11:30	**Single-cell evaluation of DNA damage in neonatal hematopoietic cells after prenatal exposure to chemotherapy** Ilana Struys *KU Leuven, Belgium*	35

11:30 - 11:45	**Mutation risk redistribution in cancer genomes linked to alterations in cell cycle genes** Marina Salvadores Ferreiro *Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain*	36
11:45 - 12:00	**Mechanisms of APOBEC3 mutagenesis in human cancer cells** Mia Petljak *NYU Grossman School of Medicine, United States of America*	37
12:00 - 13:00	**Meet the speakers and lunch** Auditorium Foyer	
13:00 - 14:30	**Session 6: Cancer ecosystems & microenvironment** **Chairs:** Sarah Aitken- University of Cambridge, UK Maria Andrianova - Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain Klaus Tschira Auditorium	
13:00 - 13:25	**To be presented on-site** Young Seok Ju *Korea Advanced Institute of Science and Technology, Republic of Korea*	38
13:25 - 13:40	**Homopolymer on/off switches mediate adaptive mutability in mismatch repair-deficient colorectal cancer** Marnix Jansen *UCL Cancer Institute, United Kingdom*	39
13:40 - 14:05	**Insights into lung cancer evolution and metastases in TRACERx and PEACE** Mariam Jamal-Hanjani *UCL Cancer Institute, United Kingdom*	40
14:05 - 14:30	**To be presented on-site** Isidro Cortes Ciriano *EMBL-EBI, United Kingdom*	41

14:30 - 15:15	**Keynote lecture: To be presented on-site** Jan O. Korbel *EMBL Heidelberg, Germany*	42
15:15 - 15:30	**Closing Remarks and Poster Prize**	
16:00	**Departure**	

Poster Removal - Please remember to collect your poster. Posters which have not been collected will be disposed of after the meeting.

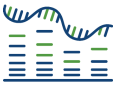
Feedback Questionnaire - Check your inbox when the meeting ends. You will find an email with the link to the online feedback questionnaire. Please take time to complete it.

Novogene | Your Trusted Genomics Partner

Novogene is a leading provider of genomic services and solutions with cutting edge Next Generation Sequencing (NGS) and bioinformatics expertise. With one of the largest sequencing capacities in the world, we utilise our **deep scientific knowledge, first-class customer service and unsurpassed data quality** to help our clients realise their research goals in the rapidly evolving world of genomics. With over 2,000 employees, multiple locations around the world and almost 20,000 publications in top tier journals such as Nature and Science, we are a world-leader in providing NGS services.

Genomics

Comprehensive genome coverage for identification of SVs, SNPs and InDels



Transcriptomics

Analysis of all RNA types to provide a global map of transcript levels and interactions

Epigenomics

WGBS, ChIP-seq & RIP-seq for methylation and protein interaction analysis

Metagenomics

Amplicon and shotgun-based approaches for microbial community analysis.

illumina® PacBio Oxford NANOPORE Technologies

Search 'Novogene Europe'

Find out more: novogene.com

1

Harnessing cancer evolution for clinical benefit: lessons from the oesophagus

Rebecca Fitzgerald[1]

[1] University of Cambridge, United Kingdom

Presenter: Rebecca Fitzgerald

To be presented on-site

2
To be presented on-site

Christina Curtis[1]

[1] Stanford University, United States of America

Presenter: Christina Curtis

To be presented on-site

3

Evolutionary dynamics of large-scale DNA copy number alterations in patient-derived adenoma organoids

Thomas van Ravesteyn[1], Lucijia Tomašic[4], Eloise van Kwawegen[2], Andrea van Santen Nieto[2], Alessandro Corsini[2], Anne Bolijn[3], Marianne Tijssen[3], Nico Lansu[2], Gerrit Meijer[3], Nenad Pavin[4], Beatriz Carvalho[3], Geert Kops[2]

1 Hubrecht Institute, The Netherlands
2 Hubrecht Institute, Oncode Institute, The Netherlands
3 Netherlands Cancer Institute, The Netherlands
4 University of Zagreb, Croatia

Presenter: Thomas van Ravesteyn

Colorectal cancer (CRC) is one of the most prevalent types of cancer and is expected to cause an increasing number of deaths as our life expectancy increases. While early detection of CRC has been improved through population screening programs, our inability to accurately predict which premalignant lesions will progress to cancer leads to overtreatment. Interestingly, CRC is characterized by wide-spread aneuploidy and recurrent patterns of chromosome copy number alterations (CNAs). We recently showed by live-cell imaging and single-cell DNA sequencing of patient-derived organoids that chromosomal instability (CIN) and karyotype evolution are ongoing in primary CRC. But how and when CIN and the resulting CNAs arise during tumorigenesis remains largely elusive. We aim to understand the mechanisms by which CIN arises, how recurrent CNA patterns are established, and how they contribute to CRC progression. Therefore, we generated a panel of patient-derived adenoma (PDA) organoid cultures from adenomas resected during a colonoscopy procedure. We transduced each PDA organoid line with H2B-mNeon lentivirus and determined chromosomal segregation error rates by live-cell imaging. While the majority of p53-proficient lines show relatively modest levels of CIN (>15%), two p53-deficient lines are highly unstable with ~40% error rates. Upon long-term culture of this PDA organoid panel we find that a large subset spontaneously evolves different genomic and phenotypic features analogous to in vivo CRC tumorigenesis. Besides acquisition of unique CNAs, single-cell DNA sequencing uncovered the evolution of CRC-recurrent chromosome gains 1q, 7pq, 8q, 13pq and 20pq in different combinations and lines. Interestingly we also find that multiple lines acquired and maintained large-scale structural alterations in a p53-proficient background. In addition, our panel contains PDA organoid lines that became p53-deficient and lost their dependency on extrinsic EGF for growth. Taken together, our findings establish PDA organoids as in vitro model for CNA evolution in premalignant cells and allows for detailed characterization and modeling of CNA dynamics in relation to CRC.

4

Selective pressures of platinum compounds shape the evolution of therapy-related myeloid neoplasms

Jurrian de Kanter[3], Eline Bertrums[5], Mark Verheul[3], Markus van Roosmalen[3], Henrik Hasle[1], Evangelia Antoniou[7], Dirk Reinhardt[7], Marry M. van den Heuvel-Eibrink[6], C. Michel Zwaan[4], Bianca Goemans[2], Ruben van Boxtel[3]

1 Aarhus University Hospital, Denmark
2 Princess Máxima Center for Pediatric Oncology, The Netherlands
3 Princess Máxima Center for Pediatric Oncology and Oncode Institute, The Netherlands
4 Princess Máxima Center for Pediatric Oncology and Erasmus MC-Sophia Children's Hospital, The Netherlands
5 Princess Máxima Center for Pediatric Oncology, Oncode Institute and Erasmus MC-Sophia Children's Hospital, The Netherlands
6 Princess Maxima Center for Pediatric Oncology; UMCU-Wilhelmina Children's Hospital, The Netherlands
7 University Hospital of Essen and AML-BFM Study Group, Germany

Presenter: Jurrian de Kanter

Childhood cancer therapy does not only affect cancerous cells but also normal cells, thereby leading to late effects. However, the exact role of the treatment exposure in transforming normal cells into second cancers like therapy-related myeloid neoplasms (t-MN) remains unclear. It is thought that the induction of driver mutations is the primary role of chemotherapy in pediatric t-MN evolution, while in adults it is mainly thought to select pre-existing mutant clones.
We studied mutation accumulation and clonal evolution of t-MN in 43 pediatric patients by combining whole genome sequencing (WGS) of bulk t-MN blasts with single-cell WGS of hematopoietic stem and progenitor cells (HSPCs) and t-MN blasts using primary template-directed amplification (PTA).
Both normal and malignant cells had an increased mutation burden compared to a previously established baseline of healthy HSPCs. In some patients, this increase could be linked to, among others, thiopurines or platinum compounds which likely induced driver mutations. We used phylogenetic inference to time the expansion of the t-MN compared to platinum compound exposure. Generally, the DNA-crosslinking effect of platinum compounds inhibited the expansion of the pediatric t-MN, indicating that the cessation of this treatment is likely the rate-limiting step of leukemogenesis. In pediatric patients with Li-Fraumeni syndrome, characterized by a germline TP53 mutation, the t-MN lost the TP53 wildtype allele and already expanded during treatment. This suggests that platinum-induced growth inhibition is TP53-dependent. In addition, it strengthens previous findings in adults that indicated that platinum compounds select for TP53-mutated clones. Moreover, it could help to explain previous studies that report that in contrast to adults, the driving mutation of t-MN could not be found long before t-MN diagnosis in pediatric patients. Finally, our results demonstrate that germline aberrations can interact with treatment exposures, highlighting the importance of the study of more such interactions.

5

Reconstructing the embryology of childhood cancer from somatic mutations

Sam Behjati[1]

[1] Wellcome Sanger Institute, United Kingdom

Presenter: Sam Behjati

The roots of childhood cancer lie in human development which is a scientific "black box" impossible to study in vivo. Yet, we can gain insights into the developmental origins of childhood cancer through retrospective lineages tracing. That is, using barcodes of somatic mutations accumulated in cells, we can reconstruct the phylogenetic relation of tumours and normal tissues. In this lecture I will discuss this approach and some of the discoveries we have made in the recent past.

6

Accurate de novo detection of somatic mutations in single-cell genomics and transcriptomics data

Francesc Muyas[1], Carolin Sauer[1], Jose Espejo Valle-Inclan[1], Ruoyan Li[3], Raheleh Rahbari[3], Tom Mitchell[3], Sahand Hormoz[2], Isidro Cortes Ciriano[1]

1 EMBL-EBI, United Kingdom
2 Harvard Medical School, United States of America
3 Wellcome Sanger Institute, United Kingdom

Presenter: Francesc Muyas

Detecting somatic mutations at single-cell resolution is essential to study genetic heterogeneity, clonal mosaicism in non-neoplastic tissues, and to identify the mutational processes operative in malignant and phenotypically normal cells. However, identifying mutations in individual cells is still challenging from a technical and algorithmic standpoint. Here, we present SComatic, an algorithm designed to detect somatic mutations de novo in single-cell transcriptomic and ATAC-seq data sets. Using more than 2.6 million single cells from 688 single-cell RNA-seq and ATAC-seq data sets, we show that SComatic can detect somatic mutations not only in tumour samples, but also in differentiated cells from polyclonal tissues not amenable to mutation detection using existing methods. In addition, SComatic allows the accurate estimation of mutational burdens and de novo mutational signature analysis at cell-type resolution. Using matched DNA sequencing and single-cell RNA-seq data, we show that SComatic has higher precision (> 4-fold) than existing algorithms for detecting somatic mutations without compromising sensitivity. Overall, SComatic permits the study of somatic mutagenesis at unprecedented scale and resolution using high-throughput single-cell profiling data sets.

7

Sperm sequencing reveals extensive positive selection in the human germline

Matthew Neville[1], Alex Cagan[1], Andrew Lawson[1], Rashesh Sanghvi[1], Tetyana Bayzetinova[1], Federico Abascal[1], Pantelis Nicola[1], Inigo Martincorena[1], Matthew Hurles[1], Raheleh Rahbari[1]

[1] Wellcome Sanger Institute, United Kingdom

Presenter: Matthew Neville

Mutations that occur in the germline are the source of all heritable genetic variation. Understanding the mechanisms creating germline mutations and the selection pressures shaping them are therefore crucial to the study of evolution and disease. Although the majority (~80%) of inherited mutations in humans originate from the paternal germline, direct observation of mutations in sperm has been limited by the need for an extremely low error rate sequencing technology. In this study, we utilized NanoSeq, a duplex sequencing method with an error rate below five per billion base pairs, to sequence 104 bulk sperm samples from individuals aged 24 to 75 years. Our findings revealed a steady accumulation of 1.4 mutations per year, consistent with classical trio studies. By conducting deep targeted sequencing on the bulk sperm samples, we identified over 10,000 coding mutations, equivalent to sequencing the coding regions in hundreds of parent-child trios for each sample. Leveraging these variants, we detected 7 known and 18 novel genes subject to significant positive selection in the male germline. Interestingly, our study shows that germline positive selection, which occurs via a proliferative advantage of mutant spermatogonia in the testis, extends beyond the RAS/MAPK pathway to a number of pathways identified in cancer studies. Notably, nearly all positively selected genes identified in our study are associated with pathogenic or cancer predisposition disorders in children when inherited, likely leading to an increased birth prevalence of these disorders, especially to older fathers. Furthermore, we quantified the fraction of sperm carrying pathogenic variants and demonstrated that although no single pathogenic variant is typically found in a high fraction of an individual's sperm (all < 0.5%), estimates of the total pathogenic burden exhibited a strong correlation with age, reaching over 5% of sperm in certain older individuals. These findings shed light on the dynamics of germline mutations and have important implications for our understanding of human disease.

8
Single-mitosis dissection of acute and chronic DNA mutagenesis and repair

Paul Ginno[1]

[1] German Cancer Research Center (DKFZ), Germany

Presenter: Paul Ginno

How chronic mutational processes and punctuated bursts of DNA damage drive evolution of the cancer genome is poorly understood. We have devised a strategy to both disentangle and quantify distinct mechanisms underlying genome evolution in single cells, during single mitoses, and at single strand resolution in mouse liver tumours. To distinguish mutations caused by chronic (ROS) and acute (UV) processes, we microfluidically separate individual pairs of mammalian sister cells resulting from the first mitosis following a burst of UV mutagenesis. Our work shows that ROS mutagenesis in transcribed regions is reduced in a strand-agnostic manner, while burst UV mutations manifest as sister-specific events, revealing mirror-image mutation phasing profiles across all chromosomes. We furthermore show that successive rounds of genome replication over persistent UV damage drives multi-allelic variation at tandem CC sites, and finally resolve phased mutational patterns to single-strands across the entire genome of liver tumours from F1 mice. The strategy we present here can be widely used to resolve the contributions of overlapping cancer relevant mutational processes.

9

On the evolutionary history of metastatic cancer

Kamila Naxerova[1]

[1] Harvard Medical School, United States of America

Presenter: Kamila Naxerova

Metastasis is a process that in humans takes place completely unobserved, over long time periods, and we know very little about it. Animal models are valuable tools for studying metastasis mechanistically, but only if we have an accurate picture of the metastatic process in humans, so that we can select the most appropriate models and ask the right questions. One our lab's main interests are host organ-specific metastasis patterns. It is currently unknown whether metastases in different organs form through distinct evolutionary mechanisms, although different metastasis types display divergent clinical behavior. We have evidence to suggest that the characteristics of colorectal cancer metastases - in particular their genetic diversity and their phylogenetic relationships to the primary tumor – differ by host organ. We are taking advantage of a unique collection of multi-region sampled metastatic colorectal cancer cases that we have been building over years, encompassing more than 1000 distinct tumor samples from more than 100 patients. We have reconstructed phylogenetic trees that visualize the evolutionary history of these cancers and find recurrent, quantifiable differences in the position of different metastasis types on the phylogenetic trees. For example, liver metastases are homogeneous and diverge late in tumor evolution, indicating late metastatic seeding by single cells or small groups of cells. Their inter-lesion uniformity suggests selection for a liver-competent growth phenotype. Locoregional lymph node metastases, in contrast, are highly diverse and associate closely with the primary tumor, indicating continuous seeding and comparatively mild selection pressure. The unique phylogenetic patterns associated with different metastasis type encode important information about formation mechanisms and have implications for surgical and medical treatment.

10

The origin and evolution of pediatric secondary neoplasms

Abel Gonzalez-Perez[1], Mònica Sánchez-Guixé[1], Ferran Muiños[1], Morena Pinheiro-Santin[1], Víctor González-Huici[1], Carlos Rodriguez-Hernandez[2], Alexandra Avgustinova[2], Cinzia Lavarino[3], Abel González-Pérez[1,*], Jaume Mora[3], Núria López-Bigas[1,4,*]

1. IRB Barcelona, Baldiri Reixac, 10, 08028 Barcelona, Spain.
2. Institut de Recerca Sant Joan de Déu, Barcelona 08950, Spain.
3. Pediatric Cancer Center Barcelona (PCCB). Hospital Sant Joan de Déu. 08950 Barcelona. Spain
4. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
*Equal contribution

Presenter: Abel Gonzalez-Perez

Pediatric cancer patients without known germline predisposing mutations that present with a second malignancy during their young age are rare. When this occurs it is often assumed that the second malignancy could be related to the treatment received for the first one. Other possible explanations for the second tumor include shared early driver mosaic mutations and completely independent somatic origin. We investigated the origin and evolution of second neoplasms of four children from Sant Joan de Déu Barcelona Children's Hospital who were cured during infancy, childhood or adolescence of a primary solid tumor and presented, several years later, with a second solid or hematopoietic malignancy. No recognizable germline predisposing variant (or familial history) explained the origin of the second tumors. Through the analysis of the whole-genome sequence of the primary and secondary tumors of the four patients, we estimated the approximate time they diverged during the life of the patients. We found different explanations underlying the origin of the four secondary malignancies. In one case, the exposure to cytotoxic therapies was responsible for the emergence of a therapy-related AML. A common somatic mutation acquired early during embryonic development was the driver of the two solid tumors of another patient. Finally, in two cases, the two tumors had emerged and developed in a completely independent manner during the embryonic development. We also investigated the mutational impact of chemotherapies across normal tissues of three of the patients. Platinum-based therapies contributed at least one order of magnitude more mutations per day of exposure than aging to normal tissues.

11

Darwinian epigenetic evolution and non-Darwinian cellular plasticity in colorectal cancer

Timon Heide[1]

[1] Human Technopole, Italy

Presenter: Andrea Sottoriva

Cancers evolve following Darwinian rules. Predicting and potentially controlling disease evolution requires appropriate data modelling to measure evolutionary processes from human samples and model systems. This can only be done by combining appropriate data collection with computational and mathematical modelling of evolutionary dynamics. Moreover, cancer evolution is not only driven by genetic mutations, but also by epigenetic alterations that, when heritable, can fuel Darwinian selection. I will present our approach, which combines new data and computational methods to extract evolutionary biology from cancer genomic and epigenomic data. I will show how we deconvolute Darwinian genetic and non-genetic mechanisms from non-Darwinian cellular plasticity. Finally I will introduce the huge challenges we currently face in modelling cancer evolution in light of the emerging complexity of multifactorial evolutionary processes.

12
Methylation dynamics in the decades preceding Acute Myeloid Leukaemia

Adriana Fonseca[4], Caroline Watson[4], Ruslan Strogantsev[2], Jose Montoya Mira[2], Sophia Apostolidou[3], Aleksandra Gentry-Maharaj[3], Usha Menon[3], Paul Spellman[2], Hisham Mohammed[1], Jamie Blundell[4]

[1] Knight Cancer Institute; Oregon Health & Science University, United States of America
[2] Oregon Health & Science University, United States of America
[3] University College London, United Kingdom
[4] University of Cambridge, United Kingdom

Presenter: Adriana Fonseca

Altered DNA methylation is a hallmark of cancer development and is a promising tissue-specific biomarker for early detection. However, it remains unknown how early in tumour development methylation changes occur, how these early changes progress as the tumour develops, and how methylation dynamics are shaped by the acquisition of somatic driver mutations. Here, we develop a highly sensitive methylation sequencing strategy to probe patterns of methylation change at ~6000 AML-specific differentially methylated regions (DMRs) across the genome at high coverage (>1000x molecular depth). By applying this technique to >250 serial blood samples collected annually in the decades preceding 50 incidental Acute Myeloid Leukaemia diagnoses and age-matched controls, we reveal how DMRs clearly distinguish pre-AML cases from controls up to 14 years prior to a clinical diagnosis. In contrast to the observations in pre-AML cases, the dynamics of methylation in healthy controls are remarkably stable between people and through time. By combining longitudinal methylation sequencing with longitudinal somatic variants calls we can associate the abundance of specific somatic driver mutations with methylation levels at specific CpGs, thus providing a genotype-phenotype map for certain driver mutations. Remarkably, we find that driver mutations in diverse AML-associated genes are associated with a convergent CpG phenotype. These results highlight the rich levels of information in methylation patterns and show how they could be used for risk stratification very early in disease development.

Adriana Fonseca is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

13
Inferring genomic trajectory of uveal melanoma

Chang Hyun Nam[1], Yong Joon Kim[2], Christopher Seungkyu Lee[2], Young Seok Ju[1]

1 Korea Advanced Institute of Science and Technology, Republic of Korea
2 Yonsei University College of Medicine, Republic of Korea

Presenter: Chang Hyun Nam

Uveal melanoma (UM) is the most common primary eye cancer, frequently leading to metastasis-related mortality. It is characterized by specific genomic aberrations associated with tumor recurrence and survival, including loss in chromosome 3, gain in chromosome 8q, and BAP1 loss-of-function mutations. These aberrations serve as determinants of the TCGA class, which is known to correlate with the prognosis of UM. However, the timing of genomic aberrations during tumor evolution remains unclear, particularly in terms of chronological ages. Here, we performed whole-genome sequencing on 40 primary UMs with their matched blood samples and identified somatic genomic variations. Of note, gain in chromosome 1q, which is not well described in previous literature, is prevalent in our cohort, particularly among patients with metastatic recurrence. Indeed, modified TCGA classification, which takes into account the status of chromosome 1q gain, better predicts whether patients will eventually develop metastasis. Furthermore, we inferred the molecular timing of prevalent copy number gains by measuring the mutational burden preceding amplification. Our findings revealed that the earliest copy number alterations were acquired several decades prior to diagnosis, presumably in the early decades of life. Throughout the long latency to diagnosis, prevalent copy number gains accumulated in a specific sequential order: gain in chromosome 6p, followed by gain in chromosome 8q, and gain in chromosome 1q. Collectively, these findings provide valuable insights into the genomic evolution of UM and offer potential biomarkers for early detection and prognosis prediction.

14
Tumor evolution: a real-time and single-cell perspective

Hugo JG Snippert[1]

[1] University Medical Center Utrecht, The Netherlands

Presenter: Hugo JG Snippert

Cells within a cancer are highly heterogeneous with respect to their phenotype and can manifest distinct morphological, molecular and functional features. As a consequence, it is challenging to design treatment therapies that target all cancer cells as effectively.
Using human samples of colorectal cancers, the Snippert lab studies the causes and consequences of heterogeneity in cellular behavior during tumor growth, tumor evolution and the emergence of therapy resistance.

Primarily, we study tumor cells using cancer organoids and live-cell imaging experiments to assess changing cell states, real-time signaling dynamics and evolving genomes, all with single-cell resolution. From a cancer cell signaling point of view, we work on drug response measurements in patient-derived organoids where real-time and single-cell resolution helps us to understand their mode of response. Regarding tumor evolution, I will discuss our recently presented 3D Live-Seq technology, that integrates imaging of tumor organoid outgrowth and whole-genome sequencing of each imaged cell to reconstruct the extent and patterns by which genetic diversity is generated in tumors. Moreover, I will discuss our new approach to study the origin of malignant transformation in colon cancer.

15

Convergence of machine learning and genomics for precision cancer medicine

Eliezer Van Allen[1]

[1] Harvard Medical School, United States of America

Presenter: Eliezer Van Allen

Numerous advances in machine learning and cancer genome analysis are converging to guide the next wave of innovation in precision cancer medicine. In this presentation, we will share ongoing research initiatives at the inteface of cancer genomics, translational research, and artificial intelligence, and how these activities may ultimately impact patient care. These efforts including interpretable cancer genomics model to guide novel target discovery, multimodal data integration involving genomics, pathology, and radiology images for patient stratification, and clinical interpretation machine learning algorithms that match patients to therapies using these combined features for individualized care. Further convergence of these fields, with emphasis on algorithmic innovation paired with equitable data generation and sharing, will guide new innovations in these domains to improve patient care for cancer patients everywhere.

16

Saturation mutational scanning uncovers druggability of all FGFR kinase point mutations

Carla Schmidt[8], Avantika Ghosh[8], Jannis Stappenbeck[8], Marisa Riester[8], José Naveja Romero[6], Martin Ziegler[2], Catarina Reis Orcinha[8], Melanie Börries[3], Tilman Brummer[5], Sonja Loges[1], Katja Luck[4], Anna-Lena Illert[7], Sven Diederichs[8]

1 DKFZ-Hector Cancer Institute, Germany
2 German Cancer Research Center (DKFZ), Germany
3 Institute of Medical Bioinformatics and Systems Medicine, University of Freiburg and German Cancer Consortium (DKTK), Germany
4 Institute of Molecular Biology, Germany
5 Institute of Molecular Medicine and Cell Research (IMMZ) and Zentrum für Biochemie und Molekulare Zellforschung (ZBMZ), Germany
6 Integrative Systems Biology, Institute of Molecular Biology, Germany
7 Technical University of Munich and German Cancer Consortium (DKTK), Germany
8 University of Freiburg and German Cancer Consortium (DKTK), Germany

Presenter: Carla Schmidt

Functional genomics is essential to interpret cancer genome data to match patients with the most promising targeted therapy. However, for most non-hotpot mutations in cancer, their impact on drug sensitivity or resistance is unknown.
Aberrantly activated fibroblast growth factor receptors (FGFRs) are found in ca. 10% of all human cancers and several selective inhibitors (FGFRis) have recently been approved by the FDA for tumors with FGFR alterations like translocations or amplifications. However, the effect of point mutations frequently found in tumors on activation and drug resistance is largely unknown.
We implemented a saturation mutagenesis platform to screen all possible point mutations in the kinase domains of FGFR1-4. Overexpression of 11520 mutations in growth factor-dependent cells identified hundreds of both novel and known activating point mutations driving their proliferation. Saturation mutagenesis screening in FGFRi-sensitive cells uncovered hundreds of novel mutations conferring resistance against Pemigatinib and Futibatinib. Notably, we found mutations conferring drug sensitivity or general resistance or differential resistance to different inhibitors. Moreover, we identified mutations causing either similar or highly divergent effects in the different members of the FGFR family and correlated these with their structural position. By integrating these comprehensive datasets, we determined which point mutations are targetable by which inhibitor to have this information readily available for clinical decision support when FGFR point mutations are identified in patients.

17

Understanding cancer biology through multi-omics genotype-phenotype tumour maps: applications in rare cancers

Lynnette Fernandez Cuesta[1]

[1] IARC, France

Presenter: Lynnette Fernandez Cuesta

Rare cancers collectively account for ~25% of all cancer diagnoses, but the lack of basic biological and clinical knowledge for these diseases translates into a worse prognosis for patients with a rare cancer than for those with common cancers. In this presentation we will introduce the Rare Cancers Genomics initiative, an international multidisciplinary open-science effort with the overarching goal of providing a turning point in the molecular characterization of rare cancers. We will show how we transcend the study of rare cancers from a mere genomic standpoint, advocating for a multi-dimensional understanding that encapsulates molecular, morphological, and immune responses using system biology approaches. By presenting our work on malignant pleural mesothelioma (the MESOMICS project) and lung neuroendocrine neoplasms (the lungNENomics project), we aim to highlight the role of integrated multi-omics in deciphering the complexity of rare cancers, ultimately contributing to precision cancer classification and care.

18

Genomic analysis of skin cancers from Xeroderma Pigmentosum subgroups revealed new mechanisms of UV mutagenesis

Andrey Yurchenko[2], Fatemeh Rajabi[3], Hiva Fassihi[5], Patricia Kannouche[1], Carlos Menck[8], Tirzah Lajus[7], Alain Lehmann[6], Chikako Nishigori[4], Alain Sarasin[1], Sergey Nikolaev[2]

1 CNRS UMR9019 Genome Integrity and Cancers, Institut Gustave Roussy, France
2 Gustave Roussy Cancer Campus, France
3 INSERM U981, Gustave Roussy Cancer Campus, Université Paris Saclay, France
4 Kobe University Graduate School of Medicine, Japan
5 National Xeroderma Pigmentosum Service, St John's Institute of Dermatology, Guy's and St Thomas' Foundation Trust, United Kingdom
6 St John's Institute of Dermatology, Guy's and St Thomas' Foundation Trust, United Kingdom
7 Universidade Federal do Rio Grande do Norte, Brazil
8 University of Sao Paulo, Brazil

Presenter: Sergey Nikolaev

Rare autosomal syndrome Xeroderma Pigmentosum (XP) is characterised by 1000 times increased risk of skin cancer due to the impaired Nucleotide Excision Repair (NER) pathway or translesion synthesis (polymerase eta). We assembled a unique collection of skin tumours (n=39) from five most frequent and cancer-prone XP subgroups (XP-A, XP-C, XP-D, XP-E, XP-V) and performed whole genome sequencing to characterise in detail their genomic mutational landscapes comparing with tumour type-matched sporadic cancers (n=139).

We found that heterogeneity of the mutation rates across skin cancer genomes is determined by the activity of NER, and that transcription-coupled NER extends beyond the gene boundaries reducing the intergenic mutation rate. The mutational profile in XP-V tumors revealed the role of polymerase η in the error-free bypass of (i) rare TpG and TpA DNA lesions, (ii) 3' nucleotides in pyrimidine dimers, and (iii) TpT photodimers. XP-V patients skin cancer mutational data are validated in-vitro. Specifically, we observed a 10-fold increase of mutagenesis in POLH-KO vs WT cells after treatment with UVA or UVC. Strikingly, mutations from TpG and TpA dinucleotides were most prevalent after UV-A exposure, and were observed after UV-C exposure. However, they were undetectable following 8 weeks of cell culturing without treatment, or with Reactive Oxygen Species inducing Potassium Bromate treatment. Overall, our study unravels the genetic basis of skin cancer risk in XP and provides insights into the mechanisms reducing UV-induced mutagenesis in the general population.

19

Genesis of uterine leiomyomas; lessons from a benign neoplasm

Lauri Aaltonen[1]

[1] University of Helsinki, Finland

Presenter: Lauri Aaltonen

To be presented on-site

20

Somatic xeroderma pigmentosum group D (XPD) mutations alter local and global landscape of mutational processes in bladder cancer

Jayne Barbour[3], Tong Ou[2], Hu Fang[3], Noel Yue[3], Xiaoqiang Zhu[3], Michelle Wong-Brown[1], Haocheng Yang[3], Yuen Wong[4], Nikola Bowden[1], Song Wu[2], Wong Jason[3]

1 Hunter Medical Research Institute, Australia
2 The Third Affiliated Hospital of Shenzhen University, China
3 The University of Hong Kong, Hong Kong
4 UNSW Sydney, Australia

Presenter: Jayne Barbour

Xeroderma pigmentosum group D (XPD) is a DNA helicase involved in transcription initiation and nucleotide excision repair. Missense mutations in XPD are putative drivers in bladder cancer (BLCA) and have been associated with nucleotide excision repair deficiency and a specific single base substitution mutational signature. However, the impact of the epigenome on XPD function and the consequence of mutations in XPD on the genome-wide distribution of somatic mutation formation remains unexplored. To further our understanding of the mutational phenotype of XPD mutants, we examined genomic and epigenomic determinants of mutation rate from whole-genome sequenced (WGS) BLCA samples with (n=39) and without (n=343) XPD mutations. XPD genotype significantly impacted the distribution of somatic mutations attributed to all major mutational processes in BLCA, including APOBEC-induced mutations. XPD mutant samples had increased mutation density at open chromatin, including striking mutation hotspots at CTCF-cohesin binding sites (CBS). XPD mutant BLCA displayed enrichment of APOBEC-associated TCW>TTW mutations (SBS2), with their distribution mirroring that of cells with uracil DNA deglycosylase (UNG) deficiency. Analysis of XPD ChIP-seq data and TFIIH repair sequencing data suggests that XPD is involved in protecting CBS from DNA damage. Base pair resolution profiles of genomic uracil in UNG deficient cell lines confirmed that uracil is enriched in CBS, specifically at the same positions within the CTCF motif as hotspot mutations. Finally, taking advantage of the variation in mutation rate at different genomic features, we develop a support vector machine that can distinguish XPD mutant and WT cancers with 99% accuracy and enable the identification of functional non-recurrent XPD mutations. Our study implicates XPD in genomic integrity maintenance of a range of mutational processes affecting both global and local mutation burden. It also raises an interesting question about the role that incorporation of dUTP into DNA may play in regulating epigenetics which we are currently exploring experimentally in our lab.

21

The genomic and transcriptomic profile of melanoma in Latin American patients

Patricia Basurto-Lozada[1], Patricia Basurto-Lozada[1], Martha Estefanía Vázquez-Cruz[1], Martha Estefanía Vázquez-Cruz[1], Carla Daniela Robles-Espinoza[1]

[1] Universidad Nacional Autónoma de México, Mexico

Presenter: Carla Daniela Robles-Espinoza

Melanoma is the most aggressive type of skin cancer. Acral lentiginous melanoma (ALM) is the most common type of melanoma in Mexico and other countries in Latin America, Africa and Asia. It arises in the palms of the hands, soles of the feet and subungual locations (under the nails), and its aetiology is unknown, but unrelated to UV exposure. It has worse 5- and 10-year disease-specific survival than other types of melanoma, and because it has been poorly studied, no specific targeted treatment options are available for its treatment. In this talk, I will discuss our research group's efforts to identify driver genes, risk factors and potential therapeutic targets for this type of cancer. Specifically, we have sequenced the exomes and transcriptomes of 128 tumours from 96 Mexican patients, and have identified that fewer than half of all tumours carry mutations in known melanoma drivers (BRAF, RAS genes, NF1), with KIT being one of the most commonly mutated genes. Analysis of the copy number alterations (CNAs) in these tumours show similarities to those reported in other studies, and nearly an absence of alterations in the MYC locus. Additionally, we found significant associations between the genomic profile and clinical characteristics, such as somatic mutations in NRAS and a younger age at diagnosis, and tumour location with number of CNAs. Unsupervised clustering of tumour transcriptomes shows three tumour groups, and deconvolution suggests that ALM tumours are characterised by an immunosuppressive microenvironment with an absence of NK-related and an abundance of CAF-related markers. We are focusing on characterising the specific markers associated with immune suppression through spatial proteomics. We are also working with our international colleagues in Brazil and the UK to create a collection of ALM PDX models to study tumour evolution and drug response. We hope that, in the future, the knowledge arising from this project deepens our understanding of ALM biology and that it can be used to create better treatment options for these patients.

22

Clonal selection in normal tissues and its relation to cancer risk

Núria López-Bigas[1]

[1] ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Núria López-Bigas

Clonal selection in normal tissues and its relation to cancer risk

Cancer appears when mutated cells with a proliferative advantage expand, invading the surrounding tissue and forming a tumor. Several studies have shown that cancer driver mutations are present in healthy tissues, implying that these mutations alone are not sufficient for tumor formation. Recent evidence also showed that many carcinogens are not mutagens. These results suggest that tumorigenesis is probably a two-step process, with a promotion step where certain pre-existing mutated clones have a growth advantage upon the exposure to a specific factor. Cancer risk factors could thus play a role in this second step by promoting the growth of specific clones. Understanding the differences in the clonal architecture of normal tissues from donors with different lifelong exposures to cancer risk factors is key to understanding their potential role as tumor promoters.

I will explain our advances regarding the effect of chemotherapies in normal tissue and clonal selection in normal bladder urothelium in relation to cancer risk.

23

Genomic landscapes of pre-malignant lesions in stomach cancer

Patrick Tan[1]

[1] Duke-NUS Medical School, Singapore

Presenter: Patrick Tan

Intestinal metaplasia (IM) is a pre-malignant condition of the gastric mucosa associated with increased gastric cancer (GC) risk. We analyzed 1256 gastric samples (1152 IMs) from 692 subjects through a prospective 10-year study. We identified 26 IM driver genes in diverse pathways including chromatin regulation (ARID1A) and intestinal homeostasis (SOX9). Single-cell and spatial profiling highlighted changes in tissue ecology and lineage heterogeneity in IM, including an intestinal stem-cell dominant cellular compartment linked to early malignancy. Expanded transcriptome profiling revealed expression-based molecular subtypes of IM associated with incomplete histology, antral/intestinal cell types, ARID1A mutations, inflammation, and microbial communities normally associated with the healthy oral tract. We demonstrate that combined clinical-genomic models outperform clinical-only models in predicting IMs likely to progress. By highlighting strategies for accurately identifying IM patients at high GC risk and clarifying a role for microbial dysbiosis in IM progression, our results raise opportunities for GC precision prevention and interception.

24

Cancer patient-derived organoids dependency map

Carmen Herranz-Ors[2], Shriram Bhosle[2], Alexandra Beck[2], James Gilbert[2], Clare Pacini[2], Gabriele Picco[2], Francesc Muyas[1], Jose Espejo Valle-Inclan[1], Isidro Cortes Ciriano[1], Agnieszka Swiatkowska[2], David Jackson[2], Emily Souster[2], Hazel Rogers[2], Maria Garcia-Casado[2], Cameron Collins[2], Laura Letchford[2], Sara Vieira[2], Jessica Davis[2], Mya Fekry-Troll[2], Jade Smith[2], Rizwan Ansari[2], Charlotte Beaver[2], Hayley E. Francis[2], Mathew Garnett[2]

[1] EMBL-EBI, United Kingdom
[2] Wellcome Sanger Institute, United Kingdom

Presenter: Carmen Herranz-Ors

A biobank of cancer patient-derived organoids has been established to study the heterogeneity and clonal evolution of the disease and help characterise potential new drug targets or cancer biomarkers. Here, we report the derivation of 215 organoids from colorectal, oesophageal, ovarian and pancreatic tumours. The models have been comprehensively annotated with clinical information, and whole-genome DNA and RNA sequencing have been performed. When available, we also sequenced the matched tumour samples. Organoids recapitulate the genomic diversity of the original lesions and stay stable over time, thereby confirming their suitability for clinically relevant research in contrast to traditional 2D in vitro cultures. In addition, organoids captured singularities related to patients' treatment. High-throughput drug screenings and genome-wide CRISPR-Cas9 screens were performed on a subset of organoids. These analyses identified colorectal organoid-specific cancer dependencies primarily associated with the EGFR-RAS-MAPK signalling pathway. This fully annotated and accessible platform offers a powerful tool for personalised medicine approaches and allows large-scale screenings in complex and diverse 3D models.

25

Charting the regulatory landscape of tumors using single-cell sequencing and spatial omics

Oliver Stegle[1]

[1] EMBL Heidelberg, Germany

Presenter: Oliver Stegle

To be presented on-site

26

Single-strand mismatch and damage patterns in cancer and cancer-predisposition syndromes revealed by single-molecule DNA sequencing

Mei Hong Liu[5], Benjamin Costa[5], Una Choi[5], Rachel Bandler[5], Emilie Lassen[3], Marta Gronska-Peski[5], Adam Schwing[5], Zachary Murphy[5], Daniel Rosenkjær[3], Vanessa Bianchi[4], Melissa Edwards[4], Lucie Stengs[4], Caitlin Loh[5], Tina Truong[5], Randall Brand[8], Tomi Pastinen[2], J. Richard Wagner[7], Anne-Bine Skytte[3], Uri Tabori[6], Jonathan Shoag[1], Gilad Evrony[5]

1 Case Western Reserve University School of Medicine, United States of America
2 Children's Mercy Kansas, United States of America
3 Cryos Sperm & Egg Bank, Denmark
4 Hospital for Sick Children, Canada
5 New York University Grossman School of Medicine, United States of America
6 The Hospital for Sick Children, Canada
7 Université de Sherbrooke, Canada
8 University of Pittsburgh School of Medicine, United States of America

Presenter: Gilad Evrony

Mutations accumulate in the genome of every cell of the body throughout life, causing cancer and other genetic diseases. Almost all of these mosaic mutations begin as nucleotide mismatches or damage in only one of the two strands of the DNA prior to becoming double-strand mutations if unrepaired or misrepaired. However, current DNA sequencing technologies cannot resolve these initial single-strand events. We developed a single-molecule, long-read sequencing method that achieves single-molecule fidelity for single-base substitutions when present in either one or both strands of the DNA. It also detects single-strand cytosine deamination events, a common type of DNA damage. We profiled more than 100 samples from diverse tissues, including from individuals with cancer-predisposition syndromes, and define the first single-strand mismatch and damage signatures. We find correspondences between these single-strand signatures and known double-strand mutational signatures, which resolves the identity of the initiating lesions. Tumors deficient in both mismatch repair and replicative polymerase proofreading show distinct single-strand mismatch patterns compared to samples deficient in only polymerase proofreading. In the mitochondrial genome, our findings support a mutagenic mechanism occurring primarily during replication. Since the double-strand DNA mutations interrogated by prior studies are only the endpoint of the mutation process, our approach to detect the initiating single-strand events at single-molecule resolution will enable new studies of how mutations arise in a variety of contexts, especially in cancer and aging.

27

Characterizing the cellular, molecular, and genetic heterogeneity of solid tumors using in situ sequencing

Mats Nilsson[1]

[1] Stockholm University, Sweden

Presenter: Mats Nilsson

I will present work on developing and applying targeted in situ sequencing (ISS) (Ke, R., et al. (2013) Nature Methods 10, 857-860) to build spatial maps of scRNAseq-defined celltypes in cm^2 sections of human tumors. The method allows in situ expression and mutation profiling of hundreds of genes with sub-cellular resolution. In collaboration with Mariella Filbin we have characterized the tumor cell states and spatial organization of H3-K27M mutant diffuse midline glioma (Liu, I., et al. (2022) Nat Genet 54, 1881-1894), and in ongoing studies we analyze glioblastoma in a similar fashion, together with Henk Stunnenberg (Ruiz-Monero, C., et al bioRxiv 10.1101/2022.08.27.505439). Targeted ISS can also be used to map somatic mutations in tumors. In a collaboration with Lucy Yates and Moritz Gerstung, we made use of this quality in a study to reveal the structure, nature, and evolution of the sub-clonal spatial landscape of some breast tumors (Lomakin, A. et al (2022) Nature 611, 594-602). I will summarize these studies and also describe ongoing work on characterizing the immune landscape of tumors.

28
Extrachromosomal DNA in cancer

Anton Henssen[1]

[1] Charité - University Medicine Berlin; Max Delbrück Center for Molecular Medicine, Germany

Presenter: Anton Henssen

Extrachromosomal DNA (ecDNA) amplification has been observed in at least 30 different cancer types and is associated with worse patient outcomes. This has been linked to increased oncogene dosage because both oncogenes and associated enhancers can occupy ecDNA. New data challenge the view that only oncogene dosage is affected by ecDNA, and raises the possibility that ecDNA could disrupt genome-wide gene expression. Recent investigations suggest that ecDNA localizes to specialized nuclear bodies (hubs) in which they can act in trans as ectopic enhancers for genes on other ecDNA or chromosomes. Moreover, ecDNA can reintegrate into the genome, possibly further disrupting the gene regulatory landscape in tumor cells. In this talk, I will discuss the emerging properties of ecDNA and highlight promising avenues to exploit this new knowledge for the development of ecDNA-directed therapies for cancer.

29
Chromothripsis in cancer

Aurélie Ernst[1]

[1] German Cancer Research Center (DKFZ), Germany

Presenter: Aurélie Ernst

Chromothripsis is frequent in cancer and linked with poor prognosis in a number of tumour types. The research in our group aims at achieving a better understanding of how chromothripsis arises and at assessing potential implications for cancer patients. More specifically, our goals are **1)** to characterize in which context chromothripsis occurs and which evolutionary dynamics chromothriptic tumors may have **2)** to decipher the mechanistic basis of chromothripsis **3)** to target vulnerabilities of tumor cells with chromothripsis. Defining the key genetic alterations as well as the biochemical and molecular factors leading to the formation and stabilization of chromothriptic chromosomes will lay the basis to identify causative factors and targets for intervention.

30
To be presented on-site

E. Alice Lee[1]

[1] Harvard University, United States of America

Presenter: E. Alice Lee

To be presented on-site

31

Geographic variation of mutagenic exposures in kidney cancer genomes

Sarah Moody[3], Sergey Senkin[1], Marcos Díaz-Gay[2], Behnoush Abedi-Ardekani[2], Thomas Cattiaux[1], Aida Ferreiro-Iglesias[1], Jingwei Wang[3], Stephen Fitzgerald[3], Mariya Kazachkova[3], Raviteja Vangara[2], Anh Phuong Le[3], Laura Humphreys[3], Ana Carolina de Carvalho[1], Sandra Perdomo[1], Ludmil Alexandrov[2], Michael Stratton[3], Paul Brennan[2]

[1] IARC, France
[2] University of California San Diego, United States of America
[3] Wellcome Sanger Institute, United Kingdom

Presenter: Sarah Moody

The Cancer Grand Challenges Mutographs project aims to identify novel preventable causes of cancer by combining the fields of cancer genomics and cancer epidemiology. Cancer epidemiology has been able to identify many risk factors for cancer, however, for some cancer types there remains substantial variation in cancer incidence rates that currently cannot be explained. Evidence of mutagenic exogenous exposures in human cancers can be detected via the distinct patterns of somatic mutations (mutational signatures) they produce. By comparing the mutational signature landscape of cancers from countries of varying incidence rate, it may be possible to identify previously unknown causes of cancer.

Here we present the results of mutational signature analysis on 962 cases of clear cell renal cell carcinoma (ccRCC) from 11 countries of varying incidence, which identifies the presence of multiple mutagenic exposures in ccRCC genomes. Firstly, a ubiquitous mutational signature was identified which strongly associated with the age standardised rate of incidence. While the cause of this signature is unknown, it was found to be associated with biomarkers of impaired kidney function. Secondly, we found both established and novel mutational signatures which are caused by aristolochic acid in ccRCC primarily from Serbia, Romania, and Thailand. The relative proportions of these aristolochic acid associated signatures differed across Romania and Serbia, suggesting that there is variation in the type of aristolochic acid exposure across the Balkan region. Thirdly, we identified a likely exogenous exposure active in the majority of cases from Japan but rare elsewhere. This signature closely matched the COSMIC reference signature SBS12, which is of unknown cause but was previously identified in liver cancers. Reanalysis of existing datasets showed that SBS12 is also enriched in liver cancers from Japan compared to those from other countries. The results collectively demonstrate that there are multiple mutagenic processes which contribute to the global incidence of kidney cancer. Future studies will be needed to further investigate the sources and geographic extent of these exposures.

32

Impact of chromatin architecture on somatic mutational processes

Sabarinathan Radhakrishnan[1], Sabarinathan Radhakrishnan[1]

[1] National Centre for Biological Sciences, India

Presenter: Sabarinathan Radhakrishnan

CCCTC-binding factor (CTCF) and Cohesin play a major role in the formation of chromatin loops and topologically associating domains (TADs) that controls gene regulation and DNA replication. CTCF/Cohesin binding sites (CBS), which are present at the loop anchors and TAD boundaries, are frequently mutated in cancers. However, the molecular mechanisms underlying this remain unclear. In this study, we propose that the increased somatic mutations observed at CBS could result from replication constraints imposed by the CTCF/Cohesin complex on the DNA, and the consequent activation of error-prone repair pathways. By using chromatin fractionation and ChIP-sequencing, we find that CTCF and Cohesin remain bound to the DNA during the S phase in HeLa cells. Examining replication stress by immunostaining proteins associated with replication fork stalling (STN1 and MRE11) and DNA double-strand breaks (gH2AX) revealed that they are colocalized with CTCF/Cohesin in the S phase compared to the asynchronous population of cells. Further, with ChIP-sequencing we assessed the DNA occupancy of the above proteins in the S phase and found that they are highly enriched at CBS as compared to the flanks and unbound sites. Moreover, analysis of somatic mutations from cancer genomes supports that the enrichment of mutations at CBS sites is higher in samples having somatic alterations in STN1 and MRE11. Taken together, these results suggest that the binding of CTCF/Cohesin on the DNA during the S phase causes replication stress and genome instability, and this can lead to mutations due to the error-prone repair.

33

CpG>TpG mutagenesis: DNA polymerase errors as a new culprit of an old crime

Marketa Tomkova[2], Michael McClellan[2], Gilles Crevel[1], Jakub Tomek[2], Sue Cotterill[1], Benjamin Schuster-Boeckler[2], Skirmantas Kriaucionis[2]

1 St George's University London, United Kingdom
2 University of Oxford, United Kingdom

Presenter: Marketa Tomkova

Transitions of cytosine to thymine in CpG dinucleotides (CpG>TpG mutations) are the most frequent mutational pattern observed in cancer cells, normal somatic and germline cells, and genetic diseases. CpG>TpG mutations are also the defining feature of single-base substitution signature 1 (SBS1), one of the two most widespread mutational signatures that accumulate across many different cell types in a clock-like manner. The high rate of CpG>TpG mutations has been attributed to the presence of 5-methylcytosine (5mC) in CpG dinucleotides and the spontaneous deamination of 5mC into thymine.
Here, we show that also replication errors produce CpG>TpG mutations, independently of spontaneous deamination. We measured the direct error spectrum of human DNA Pol ε in vitro using Polymerase Error Rate Sequencing (PER-seq), a novel technique detailed in the associated back-to-back abstract by McClellan et al. We show that human DNA Pol ε has a particularly high error rate in CpGs, mis-incorporating A opposite template C, leading to CpG>TpG mutations. This error rate is further increased by the presence of 5mC, with 7-fold higher error rate of wild-type Pol ε when replicating 5mCpG compared to unmethylated C in other contexts.
The contribution of replication errors to CpG>TpG mutations is supported also by observations in over 16,000 cancer patients, including: 1) disproportionally high CpG>TpG frequency in hypermutated cancer patients with defective Pol ε proofreading and post-replication mismatch repair (MMR), 2) enrichment of CpG>TpG mutations in these patients on the leading-strand template, in line with the major role of Pol ε in the synthesis of the leading strand, 3) correlation of tissue-matched levels of 5mC with CpG>TpG mutation frequency in cancer patients both deficient and proficient in post-replication proofreading and repair, 4) the clock-like properties of SBS1 and its faster accumulation in tissues with high turnover rate, and 5) enrichment of CpG>TpG mutations in regions with lower activity of MMR, and the lack of this enrichment in MMR-deficient samples.
In conclusion, the presented results fundamentally change our understanding of the aetiology of CpG>TpG, a major mutational force in most cells, with broad implications for cancer, genetic diseases, and evolution.

Marketa Tomkova is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

34

Forecasting oncogene amplication

Ángel Fernández-Sanromán[32], Ángel Fernández-Sanromán[2], Alice Cadiz[2], Barbara Hernando[2], Barbara Hernando[2], David Gomez[2], Maxime Tarabichi[4], Tom Lesluyes[3], Juan Manuel Coya[1], Jon Zugazagoitia[1], Luis Paz-Ares[2]

[1] H12O-CNIO Lung Cancer Clinical Research Unit, Health Research Institute Hospital 12 de Octubre (imas12), Spain
[2] Spanish National Cancer Research Center (CNIO), Spain
[3] The Francis Crick Insitute, United Kingdom
[4] Université Libre de Bruxelles, Belgium

Presenter: Geoff Macintyre

The activation of an oncogene via amplification can underpin tumour initiation, progression and treatment resistance. The ability to predict future amplification using a genomic test offers new opportunities for improved disease management and treatment. Here, we present a method to forecast oncogene amplification for individual patients. We train predictive models for each oncogene using the different types of chromosomal instability found underlying oncogenic amplifications across 6,335 patients. We validate the approach using longitudinal pairs of glioma and prostate cancers. To demonstrate clinical utility we forecasted poor prognosis in gliomas via CDK4 amplification and osimertinib resistance in lung cancers via MET amplification. We provide guidelines and recommendations for clinical implementation paving the way for a new class of biomarker where the mutation generating processes in a tumour can be used to predict future genomic changes.

35

Single-cell evaluation of DNA damage in neonatal hematopoietic cells after prenatal exposure to chemotherapy

Ilana Struys[2], Carolina Velázquez[2], Joske Ubels[3], Manosij Ghosh[2], Jarne Geurts[2], Charlotte LeJeune[4], Katy Vandereyken[2], Wouter Bossuyt[2], Kristel van Calsteren[4], Rebecca Painter[1], Bernard Thienpont[2], Thierry Voet[2], Lode Godderis[2], Liesbeth Lenaerts[2], Ruben van Boxtel[3], Frédéric Amant[4]

[1] Amsterdam UMC, The Netherlands
[2] KU Leuven, Belgium
[3] Princess Máxima Center for Pediatric Oncology and Oncode Institute, The Netherlands
[4] UZ Leuven, Belgium

Presenter: Ilana Struys

At present, chemotherapy during pregnancy is offered as a standard treatment option and deemed safe in terms of the child's cardiac and neurocognitive development. However, based on the inherent genotoxic feature of chemotherapeutic drugs and their potential to cross the placenta, we hypothesized that prenatal chemotherapy exposure could induce a genotoxic signature in the unborn child's genome. In cancer survivors, certain chemotherapies have indeed been shown to damage DNA and leave a mutational fingerprint in both cancerous and healthy tissue, which may underpin potential long-term side effects.

To investigate this, we subjected cord blood cells, as a source of neonatal DNA, from chemotherapy-treated pregnant cancer patients (n=16), untreated pregnant cancer patients (n=8) and healthy pregnant women (n=19) to 3 genotoxicity assessments: cytokinesis-block micronucleus assay on lymphocytes, bulk sequencing of clonally expanded single hematopoietic stem and progenitor cells, and single cell sequencing of lymphocytes and monocytes.

Neonatal lymphocytes from chemotherapy-treated and untreated pregnant cancer patient groups showed elevated micronucleus frequencies, suggesting both the maternal cancer condition and prenatal chemotherapy exposure are associated with increased genotoxicity levels in neonatal blood. In contrast, an increased mutation burden, with higher de novo single nucleotide variant and indel counts, was only identified in neonatal hematopoietic stem and progenitor cells of chemotherapy-treated cases. Mutational signature analyses pointed to treatment-related effects, since prenatal exposure to platinum-based regimens was associated with direct DNA damage. Additionally, induction of clock-like processes was observed after prenatal exposure to alkylating agents plus anthracyclines, or to doxorubicin, bleomycin, dacarbazine and vinblastine (ABVD). Furthermore, copy number alterations were detected in neonatal lymphocytes, but not in monocytes, of a pregnant patient treated with ABVD for Hodgkin lymphoma. These data indicate the observed increase in mutagenesis in the fetal genome is due to both indirect and direct chemotherapy-compound related mutational mechanisms. Future long-term follow-up is needed to estimate the clinical impact of these findings.

36

Mutation risk redistribution in cancer genomes linked to alterations in cell cycle genes

Marina Salvadores Ferreiro[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Marina Salvadores Ferreiro

Somatic mutations in human cells have a highly heterogeneous genomic distribution, with increased burden in late replication time (RT), heterochromatic domains. This regional mutation density (RMD) landscape varies between cancer types, in association with tissue-specific RT or chromatin organization. We hypothesized that the RMD additionally varies between individual tumors, independently of cell type. Here, we identified three common RMD signatures that describe mutation risk redistribution across megabase-sized domains in >4000 tumor genomes. First, we identified a mutation redistribution RMD signature preferentially affecting facultative heterochromatin, enriched in B1 subcompartment and subtelomeric regions. This RMD mutation risk signature strongly reflects recurrent programs in plasticity in DNA RT, heterochromatin and DNA accessibility, observed across many tumors, tissues and cultured cells, robustly linked with a higher expression of cell proliferation genes. Consistently, this RMD redistribution pattern is associated with altered cell cycle control via loss of activity of the RB1 tumor suppressor gene. Second, another independant global RMD signature was associated with loss-of-function of the TP53 pathway, affecting the redistribution of mutation rates within late RT regions. Third, we quantify a known “flat” RMD pattern associated with DNA repair failures, additionally linking it to homologous recombination deficiencies. These three mutation risk redistribution signatures can change the local mutation supply towards 26%-75% cancer driver genes, potentially diverting the course of cancer evolution. Our study highlights that somatic mutation rates at the domain scale can vary across tumors in a tissue independent manner, but associated with cell proliferation and alterations of cell cycle genes, which may trigger the local remodeling of heterochromatin and RT program in human cancers.

Funded by EU Horizon2020 project DECIDER.
(https://www.biorxiv.org/content/10.1101/2022.10.24.513586v2.full)

37

Mechanisms of APOBEC3 mutagenesis in human cancer cells

Alexandra Dananberg[2], Mia Petljak[4], Kevan Chu[1], Erik Bergstrom[5], Josefine Striepen[2], Patrick von Morgen[2], Yanyang Chen[2], Hina Shah[2], Julian Sale[3], Ludmil Alexandrov[5], Michael Stratton[6], John Maciejowski[2]

[1] Cornell School of Medicine, United States of America
[2] Memorial Sloan Kettering Cancer Center, United States of America
[3] MRC Laboratory of Molecular Biology, United Kingdom
[4] NYU Grossman School of Medicine, United States of America
[5] University of California San Diego, United States of America
[6] Wellcome Trust Sanger Institute, United Kingdom

Presenter: Mia Petljak

The APOBEC3 family of cytosine deaminases has been implicated in some of the most prevalent mutational signatures in cancer. However, a causal link between endogenous APOBEC3 enzymes and mutational signatures in human cancer genomes has not been established, leaving the mechanisms of APOBEC3 mutagenesis poorly understood. To investigate the mechanisms of APOBEC3 mutagenesis, we deleted implicated genes from human cancer cell lines that naturally generate APOBEC3-associated mutational signatures over time. Analysis of non-clustered and clustered signatures across whole-genome sequences from 251 breast, bladder and lymphoma cancer cell line clones revealed that APOBEC3A deletion diminished APOBEC3-associated mutational signatures. Deletion of both APOBEC3A and APOBEC3B further decreased APOBEC3 mutation burdens, without eliminating them. Deletion of APOBEC3B increased APOBEC3A protein levels, activity and APOBEC3A-mediated mutagenesis in some cell lines. The uracil glycosylase UNG was required for APOBEC3-mediated transversions, whereas the loss of the translesion polymerase REV1 decreased overall mutation burdens. Together, these data represent direct evidence that endogenous APOBEC3 deaminases generate prevalent mutational signatures in human cancer cells. Our results identify APOBEC3A as the main driver of these mutations, indicate that APOBEC3B can restrain APOBEC3A-dependent mutagenesis while contributing its own smaller mutation burdens and dissect mechanisms that translate APOBEC3 activities into distinct mutational signatures.
In addition to this published work (Petljak et al, Nature 2022), I will present our current efforts to further dissect the mechanisms, and instigating factors, of APOBEC3 mutagenesis in cancer.

38
To be presented on-site

Young Seok Ju[1]

[1] Korea Advanced Institute of Science and Technology, Republic of Korea

Presenter: Young Seok Ju

To be presented on-site

39

Homopolymer on/off switches mediate adaptive mutability in mismatch repair-deficient colorectal cancer

Hamzeh Kayhanian[4], William Cross[4], Giulio Caravagna[6], Suzanne van der Horst[5], Luis Zapata[3], Panagiotis Barmpoutis[4], Arne Van Hoeck[5], Eszter Lakatos[2], Kevin Litchfield[4], Benjamin Werner[1], Trevor Graham[3], Hugo JG Snippert[5], Marnix Jansen[4]

1 Barts Cancer Institute, United Kingdom
2 Chalmers University of Technology, Sweden
3 The Institute of Cancer Research, United Kingdom
4 UCL Cancer Institute, United Kingdom
5 University Medical Center Utrecht, The Netherlands
6 University of Trieste, Italy

Presenter: Marnix Jansen

Bacterial species drive resistance evolution through adaptive mutability – transient increases in clonal mutation rates to accelerate selection of resistant daughter lineages. This commonly occurs through hypermutable homopolymer sequences in DNA repair genes, also known as 'contingency loci', which stochastically move in and out of reading frame to mediate rapid genomic evolution. By mapping the clonal topography of mismatch repair-deficient (MMRd) colorectal cancer, we show that microsatellite instability modulates subclonal DNA repair by exploiting hypermutable homopolymer runs in the mismatch repair genes MSH6 and MSH3 (C8 and A8, respectively) as genomic ON/OFF switches through stochastic frameshift switching. Spontaneous mutation and reversion at these evolvability loci modulates subclonal mutation rate, mutation bias, and clonal HLA diversity during MMRd cancer evolution. We combine clonal mapping with targeted laser capture microdissection studies to trace genomic tumour evolution in response to immune selection in situ. Dedicated multiplex immunolabelling studies reveal the microenvironmental architecture of MMRd immune adaptation. Molecular evolution studies quantify subclonal mutation rates and show that immune selection favours increased subclonal mutation rates. Finally, patient-derived organoids corroborate these observations and show that in the absence of immune selection MMR homopolymer sequences drift back into reading frame, suggesting a genotoxic fitness cost of elevated mutation rates. Overall, our data reveal that MMRd cancers fuel intratumour heterogeneity by adapting subclonal mutation rate and mutation bias to immune selection through a highly conserved evolutionary strategy. Our work reveals layers of mutational complexity and microsatellite biology in MMRd cancer evolution previously hidden in bulk analyses.

40

Insights into lung cancer evolution and metastases in TRACERx and PEACE

Mariam Jamal-Hanjani[1]

[1] UCL Cancer Institute, United Kingdom

Presenter: Mariam Jamal-Hanjani

TRACERx is a prospective longitudinal cancer evolution programme which investigates the relationship between intratumour heterogeneity and clinical outcomes in non-small cell lung cancer (NSCLC). In an interim cohort of 421 patients with extended follow up, matched primary and metastatic tumour analysis has demonstrated the impact of cancer evolution on patient outcomes. Large dominant subclones in primary tumours were associated with positive subclonal selection, and mutations in specific genes were found to be under significant subclonal, and not clonal, selection. Subclonal genome doubling was identified in 20% of tumours and multiple genome doubling events were identified in 10% of tumours. Patients with tumours containing large recent subclonal expansions had significantly shorter disease-free survival (DFS), and as previously demonstrated, chromosomal instability was associated with poor outcomes, in particular, extrathoracic disease relapse.
Metastatic relapse was found to occur late in tumour evolution in 75% of tumours, and early divergence (25% of tumours) was found to occur at predicted tumour volumes below the threshold for further investigation using radiological imaging, suggesting potential implications for clinical intervention. Metastasis-seeding subclones in primary tumours were subject to subclonal expansion, likely reflective of the acquisition of subclonal driver mutations and somatic copy number alterations. ctDNA analyses in this larger cohort demonstrated the value of ctDNA to predict the likelihood of future relapse, but also its use in tracking tumour growth dynamics and emergence of relapse.
Patients who develop relapsed disease in TRACERx after surgery are co-recruited into PEACE, a national research autopsy programme. Early findings from PEACE have demonstrated the genomic complexity of cancer relapse and metastatic heterogeneity and the different patterns of cellular migration that aid metastatic dissemination with both primary-to-metastasis and metastasis-to-metastasis seeding.

41
To be presented on-site

Isidro Cortes Ciriano[1]

[1] EMBL-EBI, United Kingdom

Presenter: Isidro Cortes Ciriano

To be presented on-site

42
To be presented on-site

Jan O. Korbel[1]

[1] EMBL Heidelberg, Germany

Presenter: Jan O. Korbel

To be presented on-site

43

Detecting homologous recombination deficiency from exome sequenced breast and ovarian cancers

Ammal Abbasi[1], Ludmil Alexandrov[1]

[1] University of California San Diego, United States of America

Presenter: Ammal Abbasi

Repair of DNA double-strand breaks by homologous recombination (HR) is an essential cellular mechanism for maintaining genomic stability and preventing tumorigenesis. Prior studies have elucidated key genes in the HR pathway, including, BRCA1 and BRCA2, that commonly exhibit germline or somatic mutations in breast, ovarian, prostate, and pancreatic cancers. Defects in HR genes can disable the HR repair pathway, making cells vulnerable to double-strand breaks and, thus, rendering them sensitive to both PARP inhibitors and platinum therapies. Conventional identification of HR deficient (HRD) cancer patients involves screening for canonical genomic markers, including pathogenic germline variants and somatic copy number alterations in HR genes, thereby missing out on patients without these markers who could benefit from PARP inhibitors/platinum therapies. Recent analyses of cancer genomes have elucidated HRD-specific mutational patterns that can serve as a better alternative to conventional screening methods since they are phenotypic footprints of deficiency, independent of the mechanism(s) causing the deficiency. While prior clinical trials have shown that HRD biomarkers based on mutational patterns outperform standard-of-care approaches, these biomarkers have not been widely utilized in clinical settings as they require whole-genome sequencing (WGS). To address this gap, we present a machine learning approach termed Homologous Recombination Deficiency Profiler (HRProfiler) that incorporates the comprehensive molecular phenotypic footprint of failed HR pathway to detect HRD using whole-exome (WES) sequencing, more commonly used in the clinic. HRProfiler delivers comparable performance to current tools at the WGS resolution and outperforms them at the WES resolution. Patients treated with platinum therapy stratified as HRD by HRProfiler have a survival benefit compared to HR-proficient patients. The application of HRProfiler bridges the gap in using the comprehensive genomic footprint of failed DNA repair processes as clinical biomarkers. It provides future avenues for the reliable stratification of patients sensitive to PARP inhibitors and platinum therapies in breast and ovarian cancers.

44

Commonly mutated driver genes in primary mucinous ovarian carcinoma

Eldar Abdullaev[4], Francesca Tiso[2], Mihnea P. Dragomir[3], Peter Arndt[4], Kirsten Kübler[1]

1 Berlin Institute of Health at Charité, Germany
2Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Germany
3 Berlin Institute of Health at Charité-Universitätsmedizin Berlin; German Cancer Consortium (DKTK), Germany
4 Max Planck Institute for Molecular Genetics, Germany

Presenter: Eldar Abdullaev, Francesca Tiso

Ovarian cancer is the second most common gynecologic cancer. Histologically, approximately 90% of ovarian cancers are epithelial and among these primary mucinous ovarian cancers (PMOCs) account for approximately 3-4%. However, the accurate diagnosis of PMOCs remains challenging as current diagnostic tools often struggle to discriminate between primary mucinous tumors and mucinous metastases into the ovary. This has a significant impact on treatment decisions: curative for PMOCs versus palliative for mucinous metastases. Thus, we here describe the driver mutations most commonly present in PMOCs and investigate whether they are different from mucinous cancers that have metastasized to the ovary and can be used in differential diagnosis.
We performed targeted sequencing on a cohort of 12 PMOCs diagnosed at Charité. The gene panel included a selection of 50 commonly mutated cancer driver genes. Somatic point mutations were called and compared to targeted sequencing data derived from the AACR project GENIE database.
We found an enrichment of mutations in the KRAS and TP53 driver genes in our cohort, similar to results observed in 113 PMOCs from GENIE. 303 metastatic mucinous tumors of the gastrointestinal tract (20 gastric, 158 colorectal and 125 appendiceal) also showed a high mutation frequency in these two driver genes. This suggests that tumors with mucinous differentiation share genomic similarities, even if they originate from different organs, underscoring the need for novel genomic approaches to identify their cell-of-origin.

45

Transcriptomic landscape of muscle-invasive bladder cancer before and after neoadjuvant chemotherapy treatment

Ariadna Acedo-Terrades[3], Júlia Perera-Bel[3], Marta Bódalo[3], Maria Gabarrós[3], Miquel Claros[3], Núria Juanpere[4], Marta Lorenzo[4], Eulàlia Puigdecanet[5], Alejo Rodriguez-Vida[4], Oscar Buisan[2], Tamara Sanchueza[2], Eduardo Eyras[1], Lara Nonell[6], Joaquim Bellmunt[3]

1 Australian National University, Australia
2 Hospital Germans Trias i Pujol, Spain
3 Hospital del Mar Research Institute, Spain
4 Parc de Salut Mar, Spain
5 University of Vic - Central University of Catalonia, Spain
6 Vall d'Hebron Institute of Oncology, Spain

Presenter: Ariadna Acedo-Terrades

In muscle-invasive bladder cancer (MIBC), neoadjuvant cisplatin-based chemotherapy (NAC) has become a standard of care prior to cystectomy for eligible patients based on the improved disease-specific and overall survival. Downstaging to non-MIBC at cystectomy leads to an enhanced outcome with 5-year overall survival of 80–90%. High-throughput DNA and RNA profiling technologies might help to overcome the inability to predict responders. Since most MIBC patients undergo NAC followed by cystectomy, pre-treatment tumor biopsy and post-chemotherapy cystectomy specimens are clinically available, creating an ideal setting to study the genomic and transcriptomic effects of NAC.
Here we present RNA sequencing of a cohort of 113 MIBC patients treated with NAC from different hospitals. For each patient, FFPE pre (n=71) and post-treatment (n=29) samples were obtained from biopsy and cystectomy respectively. Response (n=58) was defined as downstaging to non-MIBC (<pT2) at cystectomy. Differential expression analysis, GSEA, deconvolution and weighted correlation network analysis (WGCNA) was performed to assess differences between responders (R) and non responders (NR) in pre-treatment samples. Differences between pre and post-treatment samples were also evaluated to understand the effect of treatment.
We found several differentially expressed (DE) genes (p.val < 0.05) upregulated in NR before treatment, associated with cancer growth and worse prognosis. On the other hand, R showed upregulated pathways related to the cell cycle. Interestingly, no differences were observed in immune cell proportions between the two groups. However, in the WGCNA, we identified a gene group negatively correlated with response, linked to crucial signaling pathways such as Wnt signaling and cell proliferation. Regarding differences between pre and post-treatment samples, we found many more DE genes with very large effect sizes (431 genes, adj.p.val < 0.05 and |logFC| > 2). Moreover, post-treatment samples showed several immune populations upregulated such as inflammation and immune cell infiltration. Future work will focus on the molecular subtyping of samples as well as profiling at DNA level, which, in combination with gene expression markers, might reveal further molecular mechanisms of response to NAC.

46

Application of long-read sequencing in chronic lymphocytic leukemia cases with complex karyotype

Sabina Adamová[4], Eva Ondroušková[4], Jan Svaton[5], Marie Jarošová[4], Michaela Bohúnová[4], Karol Pal[1], Karolína Cernovská[3], Kristýna Závacká[2], Jakub Pawel Porc[3], Jana Kotaskova[2], Karla Plevová[5], Kamila Stránská[5]

1 CEITEC, Masaryk University, Czech Republic
2 CEITEC, Masaryk University & University Hospital Brno, Czech Republic
3 Masaryk University & Central European Institute of Technology, Masaryk University, Czech Republic
4 University Hospital Brno, Masaryk University, Czech Republic
5 University Hospital Brno, Masaryk University & Central European Institute of Technology, Masaryk University, Czech Republic

Presenter: Kamila Stránská

Introduction: Complex karyotype (CK) typically involves various, often extensive numerical and structural chromosomal abnormalities. In chronic lymphocytic leukemia (CLL), it represents an established adverse prognostic marker. Common methods to detect CK include classical cytogenetics and genomic microarray, however, their resolution is limited. We aimed to explore the ability of long-read sequencing for the precise characterization of complex genomic variants in CLL patient samples.
Methods: CK cases were identified and characterized using classical (IL-2/CpG-stimulated chromosomal banding) and molecular (24×Cyte Multicolor FISH, CytoScan HD Array) cytogenomics. For long-read sequencing, high molecular weight DNA was isolated using chloroform-isopropanol extraction, fragmented by needle shearing, and short DNA fragments were eliminated. The sequencing libraries were prepared using the Ligation Sequencing Kit (Oxford Nanopore Technologies) and sequenced on the MinION or PromethION platform. Reads were aligned to the hg38 human genome reference, and breakpoints were identified with the SVIM variant caller.
Results: For 21 patients, we obtained sequencing data providing 10× (MinION; 4 patients), or >20× (PromethION; 17 patients) average coverage of the genome, with N50 15–25 kb. Identified breakpoints were compared with available cytogenomic results of classical karyotyping, mFISH and genomic microarray. The majority of these results were confirmed by long-read sequencing. In addition, we observed breakpoints in both CLL-associated genes and non-recurring genes affecting cell signaling pathways, which may impact biological processes in CLL cells.
Discussion & conclusion: Our study contributes to a better understanding of the structural variants present in complex CLL genomes and their impact on the leukemic cell phenotype. By confirming previously detected breakpoints using long-read sequencing, we demonstrated the reliability of this approach in characterizing selected chromosomal rearrangements. However, challenges remain in identifying breakpoints in highly repetitive regions. To address this, we intend to employ the T2T human reference.
Supported by MHCZ-AZV NU21-08-00237, MHCZ-DRO FNBr65269705, MUNI/A/1224/2022, NPO-NUVR LX22NPO5102, NCMG LM2023067

47

Remodeling of extracellular matrix by chromatin regulator UBR7 in Triple-negative breast cancer: insight into chemoresistance

Swagata Adhikari[1]

[1] Saha Institute of Nuclear Physics, Homi Bhaba National Institute, India

Presenter: Swagata Adhikari

Standard chemotherapeutic drugs such as Doxorubicin and 5-Flurouracil are used in treating triple-negative breast cancer (TNBC), though early relapse and induction of chemoresistance leading to a failure of the treatment regimen is of frequent occurrence. Hence, it is important to identify novel targets to overcome the challenges of therapeutic resistance. UBR7 (E3 Ubiquitin ligase) monoubiquitinates H2B at K120, inhibiting cancer metastasis by regulating the transcription of EMT genes. RNA-sequencing study reveals that UBR7 might play an essential role in ECM remodeling events in TNBC. We have identified that UBR7 transcriptionally regulates growth factor signaling along with other ECM component genes in cancer cells to exert paracrine signaling on fibroblast cells, thereby mediating their proliferation and deposition of ECM in the tumor microenvironment. UBR7 is instrumental in regulating the enrichment of the facultative heterochromatin mark H3K27me3 over the target gene promoters. Upon induction of chemoresistance, UBR7 gets altered and functionally regulates ECM stiffness and fibrotic phenotype. The altered ECM deposition and modifications during resistance conditions are monitored through LOX activity assay from a doxorubicin-resistant sphere model and total collagen assay from chemo-treated breast cancer patients. In silico analysis reveals a positive correlation between UBR7 and different ECM markers in chemotherapy-treated TNBC patients. Therefore, targeting UBR7 in therapy-resistant TNBC could be a potential tool to alleviate the challenges of treatment regimens by altering the architecture of the tumor microenvironment.

48
EAGLE: Predicting and evaluating per-nucleotide mutation susceptibility

Melissa Sanabria[1], Mario Aguilar[1], Yana Vassileva[1], Anna Poetsch[1], Jessica do Amaral-Andrade[1]

[1] Dresden University of Technology, Germany

Presenter: Mario Aguilar

Mutagenesis across the genome is highly heterogeneous and depends on active mutational processes. The distinct distribution of mutations in the genome is dependent on the direct sequence context, but also of larger range influences, such as genome function and the epigenome. Each nucleotide in the genome thus has a distinct context-dependent probability to mutate. By estimating the mutation probability of relevant sites in the genome, such as mutations in cancer genes, we can differentiate between effects of mutagenesis and selection, and thus study cancer evolution.
We developed EAGLE-MUT (Efficient Analysis with a Genome-wide LSTM to Evaluate per-nucleotide MUTation susceptibility), a deep learning Long Short-Term Memory (LSTM) model, a special type of recurrent neural network, for predicting the susceptibility to mutagenesis of the DNA in nucleotide resolution. Using over ranges of 500 nt, the DNA reference sequence and single-base-substitutions in an esophageal adenocarcinoma's cohort (n=423), we show that EAGLE-MUT is capable of capturing the specificity of mutagenesis in whole genome sequencing data and derives per-nucleotide probabilities of mutagenesis.
To evaluate the performance of EAGLE, we have developed a novel performance metric that evaluates performance over a random mutation distribution for both prediction of mutation and non-mutation. EAGLE not only outperforms other statistical models based on short range sequence context, but also provides the mutation susceptibility of each nucleotide in individual patients' genomes with up to 12-fold performance over random prediction. Clustering and applying EAGLE to subgroups underscores the individuality of patients, showcasing divergent susceptibilities both between and within clusters for the same driver genes. This allows to personalize the assessment of mutagenesis and can support the investigation of evolutionary trajectories in individual tumors.

49

Genetic encoding of per nucleotide single base substitution probabilities for individual patients

Anna Poetsch[1], Mario Aguilar[1], Yana Vassileva[1], Jessica do Amaral-Andrade[1], Melissa Sanabria[1]

[1] Dresden University of Technology, Germany

Presenter: Anna Poetsch

Cancer mutations distribute non-randomly over genomes. To understand the genome specificity of mutagenesis, we use DNA language models to investigate how stability is encoded in the DNA.
EAGLE-MUT (Efficient Analysis with a Genome-wide LSTM to Evaluate per-nucleotide MUTation susceptibility) learns sequence context of susceptibilities to somatic single base substitutions in individual patients. It is based on a bidirectional recurrent neural network with long short-term memory (LSTM) with an attention layer. In 500 nt windows it calculates the mutated genome from the unmutated genome. First, the model is trained on single based substitutions pooled from multiple tumors and then fine-tuned for individual patients. From the model we can extract the probabilities for each nucleotide to have mutated en route to the mutational pattern we see. With a model pooled over 423 esophagus adenocarcinoma samples, we achieve 12-fold performance over a random distribution. Fine-tuned for individual patients we reach a median performance of 19-fold with a peak performance of 93-fold. Since selection has limited impact on the imputed probabilities, the model output represents a background profile of mutagenic sensitivities for each nucleotide, which is computed purely on the larger sequence context in combination with the mutations that actually happened.
The obtained profiles of mutation probability reveal very heterogenous patterns. While some mutation types follow genome elements and epigenetic marks, others show very distinct hotspots that are defined by currently undescribed long-range sequence contexts. Relating the mutation probabilities to known driver mutations at different locations within the respective genes, we can differentiate between mutations that were selected despite relatively lower probability of mutagenesis and mutations that are enriched in specific patient subgroups due to favorable mechanisms of mutagenesis. Thus, EAGLE-MUT can also serve as a background model to infer selective forces for driver mutation identification.
Taken together, EAGLE-MUT shows that heterogeneity of mutagenesis over genomes is to a large extent encoded in the sequence, a code that is interpreted context dependently and thus leads to patient specific mutation probabilities for each nucleotide in the genome.

50

TREX1 and TREX2 as therapeutic targets to enhance cisplatin-induced cell death

Hashim Ahmed Nur[1]

[1] University of Cambridge, United Kingdom

Presenter: Hashim Ahmed Nur

Platinum-based chemotherapeutic agents, such as cisplatin, are commonly used as a first-line therapy for the treatment of a wide range of cancers. Although cisplatin is highly cytotoxic and many patients initially respond very well to platinum-based chemotherapy, most patients subsequently relapse with resistant disease. The DNA damage response (DDR) pathway, a complex system of biochemical mechanisms evolved to maintain genome integrity, has been implicated in resistance to platinum-based DNA damaging agents. In order to study the DDR to cisplatin, we performed unbiased transcriptome-wide and proteome-wide experiments and downstream computational analyses to identify candidates with potential for synthetic lethality.

In response to cisplatin treatment, we demonstrate that both TREX1 and TREX2 are upregulated at both the RNA and protein levels in cancer cells. TREX1 and TREX2 are non-processive 3'- 5' exonucleases that remove 3' mismatches from DNA. Although they are closely related, these proteins have distinct structural elements, suggesting that they play different biological roles. We show that in response to cisplatin-induced DNA damage, TREX2 translocates from the cytoplasm to the nucleus to execute its DDR effects. Individually, we have shown that that knocking down either TREX1 or TREX2 induces cisplatin resistance in cancer cells and causes cell cycle arrest. Additionally, we found that knocking down both proteins at the same time enhances cisplatin-induced cell death.

Taken together, our results indicate that knocking down either TREX1 or TREX2 offers a protective benefit to the cancer cells, thus allowing them to evade cell death. However, this effect is reversed when both are knocked down at the same time, resulting in enhanced cell death. Therefore, we postulate that there could be therapeutic benefit to combining inhibition of TREX1 or TREX2 with cisplatin treatment in patients with deleterious mutations in the other gene.

51

Indication of immune protection against cancer in obese younger men

Johanne Ahrenfeldt[1], Nicolai Birkbak[1]

[1] Aarhus University Hospital, Denmark

Presenter: Johanne Ahrenfeldt

A strong adaptive immune system is important to reduce the risk of cancer. Recent data indicates that women may have an overall stronger adaptive immune system than men, likely due to testosterone acting as an immunosuppressant. Obesity in men reduces the level of testosterone, which is linked to an improved immunological response to the influenza vaccine. To explore these aspects further, we investigated whether tumors found in younger (age < 60 years) obese (BMI > 30) men might be smaller in general relative to tumors from their non-obese counterparts, potentially indicating improved immune protection against cancer among these individuals.
For this endeavor, we assembled a cohort of 1,786 patients from 10 cancer types, obtained from the Cancer Genome Atlas, with annotations for BMI, age, sex, Tumor size (T-stage) and tumor gene expression data. Expression of immune pathways was calculated by gene set variation analysis and Fisher's exact test was used to determine enrichment between groups.
We compared the proportion of large (T3-T4) to small tumors (T1-T2) between sexes. We found that men in general were significantly enriched in large tumors relative to women (all: P=0.015, young: P=0.013, obese: P=0.091, young non-obese: P=0.094). Interestingly, the enrichment is lost when comparing young obese men to all young women (P=0.7), but remains in their non-obese counterparts (P=0.0046). Furthermore, we found an enrichment of large tumors in non-obese patients relative to obese patients (men: P = 0.031, women: P = 0.003), however, when stratified for age it was only significant in younger women (P = 0.047) and older men (P = 0.035). We trained a multivariate logistic model to predict large tumors based on BMI, including age, sex, fatty acid metabolism and cancer type as covariates, and were able to predict the presence of large tumors with an AUC of 0.79. Interestingly, we found that smaller tumors in younger, obese men, had a significantly higher expression of a pathway suppressing tumor immunity, compared to larger tumors (P = 0.017) and to smaller tumors from younger, non-obese men (P = 0.02).
Taken together, these results indicate that patient physiology may impact cancer development and immunosurveillance, resulting in faster growth patterns and later diagnosis, particularly in young non-obese individuals.

52

Rerunning tumour evolution reveals germline influences on mutagenesis and cancer susceptibility

Sarah Aitken[6], Juliet Luft[5], Liver Cancer Evolution Consortium[4], Núria López-Bigas[3], Colin Semple[5], Paul Flicek[1], Martin Taylor[5], Duncan Odom[2]

1 EMBL-EBI, United Kingdom
2 German Cancer Research Center (DKFZ), Germany
3 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
4 Liver Cancer Evolution Consortium, United Kingdom
5 MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, University of Edinburgh, United Kingdom
6 University of Cambridge, United Kingdom

Presenter: Sarah Aitken

Cancer evolution is a dynamic and multifaceted process, in which tumour phenotypes represent the net result of complex clonal dynamics of cellular damage and repair. Although great insights to cancer genome formation have been inferred from retrospective analysis of human tumours, including well-characterised mutational signatures caused by exogenous and endogenous toxicants, these cohorts are intrinsically confounded by a multitude of uncontrolled genetic and environmental variables.
Here, we prospectively analyse how genetic background and the epigenome of the tumour cell-of-origin shapes cancer initiation in vivo using a single burst of mutagenesis to transform healthy hepatocytes into liver tumours. We re-ran cancer evolution in four diverse inbred mouse species, capturing genetic variation equal to or greater than found in human patient populations, followed by whole genome and transcriptome sequencing of hundreds of tumours with matched histopathology (n=597).
Tumours from all species (C3H, BL6, CAST, CAROLI) demonstrate DNA lesion segregation, which results in chromosome scale asymmetry of mutations, thus extending our previous findings in C3H mice. There is also conservation of mutational signatures and convergence on a common transcriptional state across species. In contrast, the genetic background on which tumours evolved profoundly affects tumour latency, global genome stability, and driver gene preference during oncogenesis. Finally, we find species-specific differences in immune editing which may further shape the driver landscape.
We conclude that genetic background has critical implications for genome stability, oncogenesis, and cancer evolution, which is often neglected in population level studies of human disease.

53

Transposable elements are hotspots for copy number breakpoints in high grade serous ovarian cancer

Stuart Aitken[6], Colin Semple[6], Charlie Gourley[2], Patricia Roxburgh[3], Brian Dougherty[1], Simon Herrington[2], Alison Meynert[7], Ailith Ewing[5], Michael Churchman[4], Ryan Silk[7]

1 AstraZeneca, United States of America
2 CRUK Edinburgh Centre, University of Edinburgh, United Kingdom
3 Institute of Cancer Sciences, University of Glasgow, United Kingdom
4 Institute of Genetics and Cancer, United Kingdom
5 MRC Human Genetics Unit and CRUK Edinburgh Centre, University of Edinburgh, United Kingdom
6 MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, The University of Edinburgh, United Kingdom
7 MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, University of Edinburgh, United Kingdom
8 MRC Human Genetics Unit, University of Edinburgh, United Kingdom

Presenter: Stuart Aitken

Exploiting a deeply whole genome sequenced high grade serous ovarian cancer (HGSOC) dataset (N=324), we uncover the role of young LINE1 elements as hotspots for copy number variation (CNV) breakpoints. Breakpoints of consensus CNV calls of genomic deletion and amplification occur in the loci of transposable elements more than would be expected by chance, moreover, in particular classes of young LINE1 elements. Specific transposable element loci are recurrent breakpoints – hotspots – for deletions and amplifications across the cohort. We find deletions on chromosome X to be mediated by transposable elements. CNVs with breakpoints in transposable elements recurrently delete or amplify protein-coding genes and cancer genes with demonstrable impacts on survival. We also explore the role of sequence homology around breakpoints utilising the transposable element sequence to phase the CNV.

54

The deep subclonal landscape of clear cell renal cell carcinoma

Diana Alexander[2], Andrew Lawson[2], Sarah Moody[2], Yichen Wang[2], Laura O'Neill[2], Laura Humphreys[2], Behnoush Abedi-Ardekani[1], Paul Brennan[1], Michael Stratton[2], Tom Mitchell[2], Inigo Martincorena[2], David Adams[2], Peter Campbell[2]

[1] International Agency for Research on Cancer, France
[2] Wellcome Sanger Institute, United Kingdom

Presenter: Diana Alexander

Ongoing somatic mutation and selection drives tumour development and evolution and can cause the emergence of clones with more aggressive phenotypes. While common next generation sequencing approaches can reveal features of tumour evolution from mutations shared by a large proportion of cells in the tumour, they have insufficient resolution to accurately detect mutations private to single cells and small subclones. These mutations can provide insights into evolutionary trajectories and ongoing mutational processes.
In this study we investigated mutation and selection at this deep subclonal level in clear cell renal cell carcinoma (ccRCC), the most common form of kidney cancer. We sequenced tumours from 65 ccRCC patients to ~525x duplex coverage using targeted NanoSeq. The single molecule resolution of NanoSeq enables accurate detection of mutations private to single cells and small subclones. Using a targeted panel of 265 cancer driver genes we found key ccRCC drivers including SETD2, KDM5C, PBRM1 and BAP1 were under significant positive selection. This indicates detectable selection for the emergence of driver mutations in these genes at the deep subclonal level. Several patients had multiple distinct driver mutations in the same gene at increased cell fractions, suggesting there were multiple expanding subclones driven by different mutations in the same gene within the same tumour, consistent with previous observations in ccRCC.
Mutations in deep subclones are informative of ongoing mutational processes. In smokers and patients exposed to the potent mutagen aristolochic acid, a smaller proportion of mutations were attributed to each of these exposures at the deep subclonal level than in clonal mutations, suggesting their mutagenesis is reduced at later stages of tumourigenesis.
This study reveals positive selection on ccRCC driver genes in emerging small subclones and insights into ongoing mutational processes through identification of mutations previously undetectable by standard next generation sequencing.

55

Claudin genes: Prognostic and diagnostic significance in colon cancer via integrated bioinformatics

Rana Alghamdi[1], Maryam Al-Zahrani[1]

[1] King Abdulaziz University, Saudi Arabia

Presenter: Rana Alghamdi

The claudin gene family encodes tetraspan membrane proteins that are crucial structural and functional components of tight junctions. Tight junctions are the apical cell-cell adhesion that regulates paracellular permeability and is critical for epithelial cell polarity. The claudin family consists of 24 members in mammals, which exhibit complex tissue-specific patterns of expression. The extracellular loops of claudins from adjacent cells interact with each other to seal the cellular sheet and regulate paracellular transport between the luminal and basolateral spaces. The claudins interact with multiple proteins and are intimately involved in signal transduction to and from the tight junction. the claudins plays a significant role in the progression and growth of several cancers.The dysregulation of claudin genes has been linked to colon cancer, with studies indicating their crucial association with the invasion, metastasis, and prognosis of the disease.In this study, we conducted an integrated bioinformatic analysis to investigate the role of Claudin genes in colon cancer. Utilizing various methods, including differential gene expression analysis, gene set enrichment analysis (GSEA), protein-protein interaction (PPI) network analysis, survival analysis, and mutational variance analysis, we analyzed data from The Cancer Genome Atlas (TCGA) and public datasets.Our findings revealed dysregulated expression of several Claudin genes in colon adenocarcinoma (COAD). Specifically, CLDN2, CLDN1, CLDN14, CLDN16, CLDN18, CLDN9, CLDN12, and CLDN6 were found to be elevated, while CLDN8, CLDN23, CLDN5, CLDN11, CLDN7, and CLDN15 were downregulated in COAD 2.In order to validate the previous finding, we are currently conducting preliminary experiments using in vitro models. These studies will further investigate the functional roles of dysregulated Claudin genes in colon cancer progression and metastasis by gene knockout.
Understanding the role of claudins in cancer has important clinical implications that may offer potential therapeutic strategies for cancer treatment. Further research is needed to elucidate the specific mechanisms by which claudins contribute to cancer progression and metastasis.

56

Contrasts of local mutation rates elucidate DNA repair deficiencies

Marcel McCullough[1], Miguel M Alvarez[1], Maia Munteanu[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Miguel M Alvarez

Cells from the same cancer type can have varying genome-wide mutation rates and spectra, due to different mutational processes and/or deficiencies in DNA repair pathways that modulate the latter. Currently, these processes are summarized using SNV (SBS)-based trinucleotide signatures, such as those in the COSMIC catalog, which has identified over 100 presumed independent processes. However, the mechanistic basis of these signatures is often unclear, and their direct mapping to specific mechanisms is uncertain. For example, seven SBS signatures are reported for DNA mismatch repair (MMR) deficiency, raising questions about their precise correspondence to MMR mechanisms.Here, we introduce a methodology that contrasts local mutation rates across different genome regions to derive mutational signatures. In particular, each cancer genome is summarized as a vector of the associations between local mutation rates and chromatin features known to be relevant to DNA repair activity: locations of bound DNA repair factors (e.g. MSH6-binding in the case of MMR), or histone modifications (e.g. H3K36me3 is shown to be relevant to repair by recruiting MSH6 protein). The method accounts for additional topological features affecting mutagenesis, such as replication time and trinucleotide context at specific loci.We hypothesize that i) samples with a deficient DNA repair pathway will present an increase of the mutation rate in regions where there typically is a high activity thereof, ii) matrix factorization can extract these altered mutation rate patterns in the form of signatures, and iii) such local contrast signatures capture information complementary to standard trinucleotide signatures. We apply this framework to the analysis of tens of whole-genome sequences of MMR and BER deficient cell lines, demonstrating how mechanistically relevant patterns can be extracted from inter-individual variation in local mutation risk. This project is funded by the Plan Nacional project REPAIRSCAPE.

57

Plasma cell-free DNA profiling for deciphering cellular origins and immune competence in cancer patients

Laura Andersen[2], Jakob Pedersen[2], Søren Besenbacher[2], Lars Dyrskjøt[1], Claus L Andersen[2], Nicolai Birkbak[1]

1 Aarhus University Hospital, Denmark
2 Department of Molecular Medicine, Aarhus University Hospital; Department of Clinical Medicine, Aarhus University, Denmark

Presenter: Laura Andersen

Plasma cell-free DNA (cfDNA) is DNA released into the bloodstream from apoptotic cells, forming a composite of fragments from various cell-types. These fragments carry genetic and epigenetic information from the cells of origin; mutated fragments can infer genetic alterations, while fragment coverage patterns reflect chromatin status of the cells. Substantial research has focused on cancer-derived DNA fragments (ctDNA) for potential application in the clinical cancer management and early detection. Although healthy cell fragments constitute ~99% of the total pool of fragments in patients with early-stage tumors, these fragments have received little attention. Our study aims to explore the utility of plasma cfDNA to decipher the relative DNA fragment contribution from distinct cell types, particularly cell types of the immune system.
We performed whole genome sequencing of plasma cfDNA on a cohort of 263 colorectal and bladder cancer patients along with 45 healthy controls. We assessed fragmentomic profiles in three distinct genomic regions; regions of consistently positioned nucleosomes across cell types, promoter regions of housekeeping genes and immune related genes. In regions with consistent nucleosome positioning, we observed a significant increase in fragment coverage in non-relapse patients compared to healthy individuals (female: $p = 0.0043$, male: $p = 0.0063$). The same tendency was observed in promoter regions of housekeeping genes, although only significant in women (female: $p = 0.013$, male: $p = 0.099$). These results may suggest an elevated presence of cells in the bloodstream as a response to cancer. We further evaluated fragment coverage in promoter regions of immune related genes and found a significant increase in older men (age >65) compared to middle aged men (age 40-65, $p = 0.045$), suggesting a decreased expression of these genes. Future work may show how fragment coverage in differential accessible chromatin regions can be used to quantify the relative contribution of DNA from distinct cell types to the overall pool of cfDNA. We anticipate our analysis to illuminate the utility of plasma cfDNA as a biomarker for the immune competency of cancer patients. Such insight could potentially enable improved patient stratification for treatment, leveraging their distinct immune capabilities.

58

Using lesion segregation to infer multiple mutagen exposures in cancer

Craig Anderson[3], Martin Taylor[2], Michael Nicholson[5], Sarah Aitken[4], Duncan Odom[1]

1 German Cancer Research Center (DKFZ), Germany
2 MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, University of Edinburgh, United Kingdom
3 MRC Human Genetics Unit, University of Edinburgh, United Kingdom
4 University of Cambridge, United Kingdom
5 University of Edinburgh, United Kingdom

Presenter: Craig Anderson

Following exposure to a mutagen, strand-specific DNA damage necessarily segregates into daughter cells and can go unrepaired. This phenomenon, known as 'lesion segregation', manifests in chromosome-scale, strand-asymmetric distributions of mutations and is foundational to tumourigenesis. Inference of the lesion-containing strand offers novel insights into wide-ranging mutation aetiology, enabling us to compare strand asymmetric processes across the cancer genome, better resolution of mutational signatures and insights into cancer cell transformation.
As such, lesion segregation is a powerful tool, but is limited in its utility because signals of mutational asymmetry are convoluted with each additional exposure to a mutagen. Multiple exposures are typical across human cancers and associated laboratory models, including those impacted by UV, cigerette smoke or chemotherapeutics. As such, deconvolution of the patterns of lesion segregation across cancers that are the result of complex exposure scenarios is necessary. Here, we present a model for deconvolution of multiple exposures, supported by a range of in silico and in vivo exposure scenarios.
Initially, we present an approximate bayesian computation model for inference of exposure numbers throughout a range of complexities of in silico scenarios, which utilises a unique framework for quantifying the magnitude of DNA damage. We employ our model across in silico mixtures of exposures created from our diethylnitrosamine (DEN)-exposed C3H mouse cohort and compare them to a complementary, but previously unreported, WGS from a cohort of F334 rats uniformly exposed to 10 doses of DEN. We subsequently apply our model to previously published murine breast cancer models exposed to dimethylbenzathracene, as well as a human cancer cohort exposed to chemotherapeutics. These data form the basis for our efforts to match mutations to specific exposures in our bid to infer the lesion-containing strand more widely.

59

Translesion resynthesis induced oncogenic mutagenesis

Craig Anderson[6], Lana Talmane[5], Juliet Luft[5], John Connelly[5], Michael Nicholson[8], LCE Consortium[4], Núria López-Bigas[3], Paul Flicek[1], Colin Semple[5], Duncan Odom[2], Sarah Aitken[7], Martin Taylor[5]

1 EMBL-EBI, United Kingdom
2 German Cancer Research Center (DKFZ), Germany
3 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
4 Liver Cancer Evolution Consortium, United Kingdom
5 MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, University of Edinburgh, United Kingdom
6 MRC Human Genetics Unit, University of Edinburgh, United Kingdom
7 University of Cambridge, United Kingdom
8 University of Edinburgh, United Kingdom

Presenter: Martin Taylor

DNA base damage is a major source of oncogenic mutations (Alexandrov et al. 2020). Such damage can produce strand-phased mutation patterns and multiallelic variation through the process of lesion segregation (Aitken et al. 2020). Here, we exploited these properties to reveal how strand-asymmetric processes, such as replication and transcription, shape DNA damage and repair. The accumulation of multiple distinct mutations at the site of persistent lesions provides the means to quantify the relative efficiency of repair processes genome-wide and at single base resolution. At multiple scales, we show the pattern of DNA damage induced mutations is largely shaped by the influence of DNA accessibility on repair efficiency, rather than gradients of DNA damage. Finally, we reveal specific genomic conditions that can actively drive oncogenic mutagenesis by corrupting the fidelity of nucleotide excision repair. This translesion resynthesis induced mutagenesis (TRIM) can preferentially drive high variant allele frequencies and provides an opportunity for the production of oncogene proteins from a newly damaged gene before genome replication.

60

Estimation of minimum required number of drivers and their fitness from genealogy of cell divisions in cancer

Maha Shady[1], Maria Andrianova[3], Shamil Sunyaev[2], Vladimir Seplyarskiy[2]

1 Dana Farber Cancer Institute, United States of America
2 Harvard Medical School, United States of America
3 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Maria Andrianova

For a long time cancer was suggested to be a genetic disease that arises in proliferating cell lineages that acquired a crucial number of driver mutations during lifetime. This forced the investigation of the genetic component of cancer: estimation of minimum required number of driver events, building catalogs of driver mutations, deep investigation of patterns of somatic mutagenesis and role of the intrinsic and extrinsic processes in mutation accumulation. Mutation theory has much evidence of support: deficiency in DNA repair genes strongly enhances cancer predisposition and cancer incidence is associated with exposure to mutagens. However, recent studies accumulate more and more observations that can not be explained by genetic determinism. More complex theories of cancer origin were developed that consider cell identity as another important variable of tumorigenesis. This susceptibility of cells to develop cancer can depend on different intrinsic (epigenetic state, transcriptional program, etc) and extrinsic factors (microenvironment, mutagens, promoters).
Separation of the effects of genetic and non-genetic factors in each particular cancer type could be the next step in understanding of cancer. This basically means to answer two main questions: 1) how many drivers are necessary and sufficient to initiate malformation and 2) how many cells with this set of drivers could lead to malformation. Answering these questions is a difficult task because of variability of different factors and heterogeneity of the data. To get around this limitation we took advantage of available data from a specific controlled system with chemically induced liver tumors in mice (Aitken et al. 2019) that gives the opportunity to trace first divisions after mutagen exposure and reconstruct genealogies of tumors. Utilizing this system we have shown that most of the tumors resulted from one driver that prompts a very high fitness advantage >50%. Preliminary results also show that the expected number of liver cells that have such a driver event is much higher than the number of observed tumors, implying the limiting role of non-genetic factors.

61

Hotspot propensity across mutational processes

Claudia Arnedo-Pac[3], Ferran Muiños[2], Abel Gonzalez-Perez[2], Núria López-Bigas[1]

1 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
2 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
3 Institute for Research in Biomedicine Barcelona (IRB Barcelona) Spain; University of Cambridge, United Kingdom

Presenter: Claudia Arnedo-Pac

The sparsity of observed mutational events across the genome impairs our ability to study mutation rate variability at nucleotide resolution. Here, we investigate the propensity of 14 different mutational processes to form hotspots, this is, recurrent mutations in the exact same position across cancer genomes. We find that mutational signatures 17 (SBS17a and SBS17b) and 1 (SBS1) show the highest propensity to form hotspots, generating 5-78 times more than other common somatic mutational processes. After accounting for trinucleotide mutational probabilities, sequence composition and heterogeneity of mutation rates at 10 Kbp, most (89-95%) SBS17 hotspots remain unexplained. This suggests that local genomic features play a significant role in SBS17a and b hotspot propensity, among which we identify CTCF binding as a minor contributor. In the case of SBS1, we find that including genome-wide distribution of methylated CpGs sites into our models can explain virtually all (80-100%) of its hotspot propensity. The increased hotspot propensity of SBS1 is also observed in normal tissues from mammals, as well as in de novo germline mutations. We demonstrate that hotspot propensity is a useful metric to assess the accuracy of mutation rate models at nucleotide resolution and to investigate the covariates of such variability. We believe this new approach and the findings derived from it will contribute to a range of somatic and germline studies investigating and modelling mutagenesis.

62

GS-TCGA: Gene set-based analysis of the cancer genome atlas

Tarrion Baird[1], Rahul Roychoudhuri[1]

[1] University of Cambridge, United Kingdom

Presenter: Tarrion Baird

Most tools for analysing large gene expression datasets, such as The Cancer Genome Atlas (TCGA), focus on expression of individual genes or inference of the abundance of specific cell types from global gene expression data. While these methods provide useful insights, they can overlook crucial process-based information that could enhance our understanding of cancer biology. This poster describes the development of four novel tools, incorporated into a new online resource: Gene set-based Analysis of The Cancer Genome Atlas (GS-TCGA). GS-TCGA is designed to enable user-friendly exploration of the TCGA using pathway-based analysis, leveraging gene sets from the Molecular Signatures Database (MSigDB). GS-TCGA includes four unique tools. 1 GS-Surv performs survival analysis and generates Kaplan-Meier plots based on average gene-set expression, allowing users to explore how biological pathways predict survival in patient data. 2. GS-Surv (custom) performs the same analysis using a user defined gene list, enabling users to easily interrogate the clinical relevance of their own experimentally defined data. 3. CC-GSEA performs gene set enrichment analysis on co-correlated genes. As co-correlated genes may collectively contribute to cellular processes, CC-GSEA is a powerful method providing novel hypotheses of gene function. 4. GS-Corr calculates the correlation of average gene-set expression with expression of all other genes, suggesting gene functions or regulatory mechanisms. All tools are freely accessible, with all outputs available for download. These tools empower researchers to examine survival analysis linked to gene-set expression, explore the functional implications of gene co-expression, and identify potential gene regulatory mechanisms. GS-TCGA is available online at: https://gs-tcga.shinyapps.io/home/.

63

The development of mutational processes in primary and metastatic cancers

Toby Baker[1], Tom Lesluyes[1], Maxime Tarabichi[3], Peter Van Loo[2]

1 The Francis Crick Insitute, United Kingdom
2 The University of Texas MD Anderson Cancer Center, United States of America
3 Université Libre de Bruxelles, Belgium

Presenter: Toby Baker

Tumours tissues accumulate somatic mutations over the course of their development, from the fertilized egg to the time of biopsy. Many mutational processes induce single nucleotide variants with characteristic trinucleotide spectra, known as mutational signatures, which vary over tumour tissue development.
Here we present a novel method, GRITIC-SNV that leverages copy number gains and mutation clustering information to quantitatively time mutational signatures across both the clonal and subclonal evolutionary periods.
We first simulate a realistic tumour cohort and verify that GRITIC-SNV can accurately infer changes in mutational signature composition across tumour development.
We then apply GRITIC-SNV to time signature exposures across 7,071 tumours in the Pan-Cancer Analysis of Whole Genomes and Hartwig cohorts. As expected, we find that environmentally associated mutagens, such as UV light and tobacco smoke, are predominantly associated with earlier mutations while treatment-associated signatures generally occurred later in tumour development. Interestingly, mutations caused by aristolochic acid, a known carcinogen, are found to often arise after the accumulation of several copy number gains. We also find evidence for periodic APOBEC mutagenesis in many cancer types.
GRITIC-SNV can accurately measure changes in mutational signature composition over time, including the clonal period. This enables a better understanding of the role that individual mutational processes play in shaping different stages of tumour development and evolution.

64

Charting the mutational landscape of triple negative breast cancer tumours during treatment

Devin Bendixsen[4], Fiona Semple[1], Alastair Ironside[2], Natalie Wilson[1], Ailith Ewing[3], Colin Semple[4], Olga Oikonomidou[1]

1 Cancer Research UK, Scotland Centre, MRC Institute Genetics and Cancer, University of Edinburgh, United Kingdom
2 Edinburgh Cancer Centre, Western General Hospital, Department of Pathology, NHS Lothian, United Kingdom
3 MRC Human Genetics Unit and CRUK Edinburgh Centre, University of Edinburgh, United Kingdom
4 MRC Human Genetics Unit, MRC Institute Genetics and Cancer, University of Edinburgh, United Kingdom

Presenter: Devin Bendixsen

Triple Negative Breast Cancer (TNBC) is characterised by extensive intra-tumour heterogeneity (ITH) where clonal lineages diverge from distinct subpopulations over time, impacting treatment resistance and influencing metastasis. Most patients with early or locally advanced disease receive neoadjuvant chemotherapy (NACT). However, the mutational dynamics of these tumours over time is poorly understood. We investigated the mutational landscapes that occur in TNBC, how they develop in response to NACT, and throughout the progress of the disease, using a longitudinally sampled TNBC cohort. Here, we used whole-exome sequencing (WES) and an established inexpensive multiplexed assay (CUTseq) to profile the mutational landscape of FFPE samples from 22 TNBC patients, including samples from pre-treatment biopsies, samples from surgery post-neoadjuvant treatment and samples from recurrent tumours. DNA was extracted from homogenous tumour regions and was prepared for WES or CUTseq sequencing using respective protocols. In total, 96 samples were sequenced on an Illumina NextSeq 2000. Bioinformatic pipelines were then used to 'map' the mutational landscape of each tumour sample resulting in a library of consensus variant calls across mutational classes: short nucleotide variants (SNV), structural variants (SV) and copy number variants (CNV). Tumour subclonal reconstruction was then performed to track clones and mutational patterns through time during treatment. This allowed unusual insights into the mutational dynamics of TNBC during treatment. We have optimised and benchmarked a relatively inexpensive profiling technique to identify copy number variants in TNBC patients and study their dynamics over the course of the disease. WES and CUTSeq approaches provided a rich account of the structural evolution of these tumour genomes. We present mutational profiles from before and after NACT from the same patient, which can be compared to reveal commonalities and divergence in genome-wide patterns of structural variation. Ultimately, specific variants associated with survival, recurrence and metastasis can be identified. These data will direct larger scale follow-up studies for biomarker validation and treatment stratification.

65
Evolutionary Dynamics of Oligodendrogliomas

Sarah Benedetto[1], Verena Körber[1], Thomas Höfer[1], Peter Lichter[1], Bernhard Radlwimmer[1], Guido Reifenberger[2], Joerg Felsberg[2]

1 German Cancer Research Center (DKFZ), Germany
2 University Hospital Dusseldorf, Germany

Presenter: Sarah Benedetto

The development and progression of cancer is an evolutionary process. Deep genome sequencing data now allow us to time the origin and study the progression of individual tumors. We developed a model based on the birth-death process to infer tumor evolution from whole-genome sequencing data and applied it to six samples of an adult brain tumor, grade II oligodendrogliomas. Specifically, we used statistical and population-genetic approaches to infer the clonal evolution of primary tumors. We found that the tumors originated in early childhood and thereafter evolved slowly. Within the primary tumor, we found selected subclones in all samples with widely different selective advantages. In weakly selected subclones, we did not detect known tumor drivers whereas subclones harboring an activating mutation in the TERT promoter, which supports telomere maintenance, were strongly selected. These findings suggest that mutations in the TERT promoter may be acquired during oligodendroglioma evolution and confer a large selective advantage by which they eventually reach fixation in the majority of cases. Thus, TERT promoter mutations supporting telomere maintenance may be a key determinant of the aggressiveness of the disease.

66

Copy number losses of oncogenes and gains of tumor suppressor genes generate common driver events of human cancer

Elizaveta Besedina[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Elizaveta Besedina

Cancer driver genes can be under positive selection for various types of genetic alterations, including gain-of-function or loss-of-function point mutations (single-nucleotide variants, SNV), small indels, copy number alterations (CNA) and other structural variants. We studied the landscape of interactions between these different types of alterations affecting the same gene by a statistical method, MutMatch, which can test for significant differences in selection, while accounting for various causes of mutation risk heterogeneity. Analyzing 17,644 cancer exomes and genomes, we found that known oncogenes simultaneously exhibit signatures of positive selection and also negative selection, where the latter can mask the former. Consistently, focussing on known positively selected regions identifies additional tumor types where an oncogene is relevant. Next, we characterized the landscape of CNA-dependent selection effects, revealing a general trend of increased positive selection on oncogene mutations not only upon CNA gains but also upon CNA deletions. Conversely, we observe a positive interaction between mutations and CNA gains in tumor suppressor genes. Thus, two-hit events involving point mutations and CNA are universally observed on driver genes regardless of the type of CNA, and may signal new therapeutic opportunities that have been overlooked. An explicit focus on the somatic CNA two-hit events can identify additional driver genes relevant to a tumor type. By a global analysis of CNA-selection effects across many driver genes and tissues, we identified at least four independently varying signatures, and thus generated a comprehensive, data-driven classification of cancer genes by mechanisms of (in)activation by genetic alterations.
doi: https://doi.org/10.1101/2023.08.05.552104
Funded by EU Horizon2020 project DECIDER (965193)

67

Deciphering the molecular signatures of cancer promotion through the analysis of normal tissues

Raquel Blanco[2], Ferriol Calvet[2], Miguel Grau[2], Ferran Muiños[2], Brendan Kohrn[3], Elena Latorre-Esteves[3], Jeanne Fredrickson[3], Charlotte An[3], Abel Gonzalez-Perez[2], Rosa Ana Risques[3], Núria López-Bigas[1]

1 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
2 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
3 University of Washington, United States of America

Presenter: Raquel Blanco

Cancer in adults is thought to mainly originate through the accumulation of mutations over decades. However, some of the specific alterations driving tumorigenesis have also recurrently been found in normal tissues and associated with clonal expansions, highlighting the involvement of other non-mutagenic mechanisms in cancer initiation. In relation with this, recent studies have also proven many known carcinogens do not exert their effect through a mutagenic mechanism. Altogether, this raises the possibility that the rate-limiting step in carcinogenesis are the selective constraints acting on pre-existing mutant cells that persist in normal tissues until exposed to a promoting stimulus. We hypothesize that the exposure of individuals to different potential cancer promoting agents leaves differences in the clonal structure of their normal tissues, which may be detected through deep sequencing.

Here we show the preliminary results of a normal bladder pilot study on a cohort of 26 individuals (16 males and 10 females, ages 24 through 86) from the University of Washington. Samples from two regions in the urothelium were collected from each individual at the time of autopsy, one covering the bladder top (dome) and the other the floor (trigone). We performed bulk DNA error-correcting duplex-sequencing using a panel of 16 genes that are relevant in bladder carcinoma or known to be frequently mutated in normal urothelium. We identified the mutations using a pipeline developed in-house, reaching an average sequencing coverage of 6000x.

We found mutations in all the genes included in the panel, although their frequency varied greatly, with interesting differences across patients. As previously reported in epithelial urothelium, we found APOBEC signatures (SBS2 and SBS13) across the majority of samples, together with SBS40 and SBS5 clock-like signatures. We found strong positive selection in most of the genes in the panel, with remarkable values for KMT2D, RBM10, KDM6A, STAG2 and ARID1A at the cohort level. Again, inter-individual variability was observed in the genes under positive selection across patients. The results of this study will be extended with the analysis of samples from the IARC Mutographs repository, obtained from donors with different exposure to potential bladder cancer promoters.

68

Unveiling biomarkers to improved response prediction to immune checkpoint inhibitors in advanced bladder cancer

Lilian Marie Boll[1], Marta Espinosa Camarena[1], Joaquim Bellmunt[1], M.Mar Albà[2], Sergio Vázquez[1]

1 Hospital del Mar Research Institute, Spain
2 Hospital del Mar Research Institute, Catalan Institute for Research and Advanced Studies

Presenter: Lilian Marie Boll

Immune checkpoint inhibitors (ICI) can induce long-term complete remissions in cancer patients even at advanced stages. However, their efficacy remains limited to a small fraction of patients. To unravel the underlying factors driving this heterogeneous clinical response, we conducted a comprehensive analysis of six cohorts of advanced bladder cancer patients who received anti-PD-L1/PD-1 therapy. Leveraging computational techniques, we established a robust pipeline for analyzing RNA and whole-exome sequencing data. Through this pipeline, we extracted an array of biomarkers encompassing somatic mutations, clonality, neoantigen burden, mutational signatures, gene expression levels, immune infiltration, and bladder cancer subtypes. We identified nonstop mutations as a promising novel biomarker in predicting response to ICI treatment. We observed a higher ratio of nonstop mutations in responders compared to non-responders, resulting in the generation of a substantial number of putative non-self antigens. We are constructing a complete predictive model aimed at enhancing the accuracy of response prediction to ICI treatment in bladder cancer patients. Existing prediction models have been hampered by either limited sample sizes or their pan-cancer design, rendering them inadequate for identifying cancer type-specific biomarkers. By consolidating a large bladder cancer meta-cohort, our objective is to uncover new useful biomarkers and explore interactions that will augment the precision of response prediction to ICI therapy.

69

Profiling the genomic landscape of giant cancer cells in undifferentiated pleomorphic sarcomas: Hopeful monsters or an evolutionary dead end?

Amy Bowes[1], Tom Lesluyes[1], Maxime Tarabichi[4], Adrienne Flanagan[5], Nischalan Pillay[3], Peter Van Loo[2]

1 The Francis Crick Institute, United Kingdom
3 The University of Texas MD Anderson Cancer Center, United States of America
3 UCL Cancer Institute, Sarcoma Genetics, United Kingdom
4 Université Libre de Bruxelles, Belgium
5 University College London Cancer Institute, Royal National Orthopaedic Hospital, United Kingdom

Presenter: Amy Bowes

Background: Undifferentiated pleomorphic sarcomas (USARCs) are a malignant soft tissue tumour that demonstrate a high frequency of chromothripsis on WGS. USARCs often contain scattered giant polyploid tumour cells, or 'monster cells', which have an unknown role in tumour biology. As these polyploid cells frequently display bizarre nuclear atypia in the form multi-nucleation, micro-nuclei formation and the presence of extra-chromosomal DNA (nuclear features that have been linked to chromothripsis), we hypothesise that these atypical cells could harbour chromothriptic events, as well as additional complex structural variants. Single-cell sequencing approaches hold the potential to study the genomic landscape of these atypical cells in greater detail, as well as providing insight into how they might contribute to tumour heterogeneity, evolution and progression.
Methods: 112 polyploid cancer cells identified in 10 USARC cases by histology (i.e., nuclear size and shape, as well as multi-nucleation) were isolated using laser capture microdissection (LCM; Zeiss PALM Microbeam LCM microscope), thereby retaining important morphological and spatial information. Once isolated, single cell DNA was extracted and amplified using the Ampli1 whole genome amplification (WGA) kit (BioMenarini Silicon Biosystems). Single cell DNA was sent for WGS using 150 base pair, paired end sequencing with an average coverage of 3 X. All USARCs have also undergone matching bulk tumour and normal WGS. Allele-specific copy number aberrations (CNAs) were calculated for each single cell.
Results: After down sampling all BAM files to the lowest observed read depth, performance parameters such as coverage uniformity, coverage breadth, mapped reads and duplicated reads were found to be comparable across all laser-captured single cells. ASCAT.sc successfully inferred allele-specific copy number alterations in 85% of single cells, some of which encompassed chromothriptic events.
Conclusions: LCM permits single cell isolation of giant polyploid cancer cells whilst preserving the morphology and spatial topography of the tumour. Analysis is ongoing and future work aims to further assess the genomic landscape of polyploid cancers cells, as well as their potential impact on tumour progression and patient prognosis.

70

Effect of cancer risk factors on the clonal structure of normal tissues

Ferriol Calvet[3], Morena Pinheiro[3], Núria Samper[3], Erika López-Arribillaga[3], Miguel Grau[3], Raquel Blanco[3], Ferran Muiños[3], Maria Andrianova[3], Axel Rosendahl-Huber[3], Michael Olanipekun[1], Behnoush Abedi-Ardekani[1], Ana Carolina de Carvalho[1], Victor González-Huici[3], Abel Gonzalez-Perez[3], Paul Brennan[1], Núria López-Bigas[2]

1 IARC, France
2 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
3 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Ferriol Calvet

Cancer appears when mutated cells with a proliferative advantage expand, invading the surrounding tissue and forming a tumor. The accumulation of cancer driver mutations is what gives this advantage to our cells. However, several studies have shown that driver mutations are present in healthy tissues, implying that these mutations alone are not sufficient for tumor formation. Recent evidence also showed that many carcinogens are not mutagenic. These results suggest that tumorigenesis is probably a two-step process, with a promotion step where certain pre-existing mutated clones have a growth advantage upon the exposure to a specific factor that creates a bottleneck in which pre-existing mutations become advantageous to the cell. Cancer risk factors such as diabetes or hypertension could be thus playing a role in this second step by promoting the growth of specific clones. Understanding the differences in the clonal architecture of normal tissues from donors with different lifelong exposures to cancer risk factors is key to understanding their potential role as tumor promoters.To probe the differences in kidney clonal structure across donors from countries with different incidence of clear renal cell carcinoma (RCC), we designed a panel of 9 genes, frequently mutated in this malignancy. We then sequenced normal kidney samples of RCC patients available at the Mutographs biorepository at IARC. We followed the duplex sequencing protocol to prepare the samples and analyzed the sequencing reads with a pipeline developed in-house, reaching an average sequencing coverage of ~25,000x.We found mutations in all the genes included in the panel, and identified the mutational signatures active in these kidney samples, which coincided with those detected across tumors from the same donors. Using state of the art methods to detect positive selection in cancer, we identified PTEN and TP53 as positively selected across all samples analyzed, but also a degree of inter-individual variability. We performed calibration tests with neutrally generated sets of mutations to ensure that the signal observed in our samples was not artifactual. Currently, we are in the process of scaling up the sequencing and analysis of samples from selected groups of donors with different histories of exposure to potential kidney cancer promoters.

71

Aberrant transcriptome analysis of 3,760 hematologic malignancies reveals LRP1B as hairy cell leukemia variant biomarker

Xueqi Cao[6], Sandra Huber[4], Ata Jadid Ahari[6], Franziska R. Traube[8], Marc Seifert[1], Christopher C. Oakes[7], Polina Secheyko[3], Ines Scheller[6], Nils Wagner[6], Vicente A. Yépez[6], Piers Blombery[5], Torsten Haferlach[4], Matthias Heinig[2], Leonhard Wachutka[6], Stephan Hutter[4], Julien Gagneur[6]

1 Düsseldorf University Hospital, Germany
2 Helmholtz Center Munich, Deutsches Zentrum für Herz- und Kreislaufforschung (DZHK), Technical University Munich, Germany
3 Ludwig-Maximilians-University Munich, Germany
4 MLL Münchner Leukämie Labor, Germany
5 Peter MacCallum Cancer Centre, Australia
6 Technical University of Munich, Germany
7 The Ohio State University, United States of America
8 University of Stuttgart, Germany

Presenter: Xueqi Cao

Rare driver mutations are suspected to substantially contribute to the large heterogeneity of hematologic malignancies, but their identification remains challenging. To address this issue, we generated the largest dataset to date of matched whole genome sequencing and total RNA sequencing of hematologic malignancies from over 3,760 patients spanning 24 disease entities [1]. To discover rare and large regulatory aberrations, we analyzed aberrant expression and splicing events using an extension of DROP (Detection of RNA Outliers Pipeline) [2] and AbSplice [3], an algorithm that identifies genetic variants causing aberrant splicing. We found a median of seven aberrantly expressed genes, two aberrantly spliced genes, and two rare splice-affecting variants per sample. Each category showed significant enrichment for well-characterized driver genes, with odds ratios exceeding three among genes called in more than one sample. We next trained disease-specific driver gene prediction models integrating these data with recurrent mutation analyses [4]. On held-out data, integrative modeling significantly outperformed modeling based solely on genomic data and revealed promising novel candidate driver genes. Moreover, we found a truncated form of the low density lipoprotein receptor LRP1B to be aberrantly overexpressed in about half of hairy cell leukemia variant (HCL-V) samples and, to a lesser extent, in closely related B-cell neoplasms. This observation, which was confirmed in an independent cohort, suggests LRP1B be a novel biomarker and a yet unreported functional role of LRP1B within these rare entities. Altogether, this dataset and the companion computational workflow constitute unique resources to deepen our understanding of rare oncogenic events in hematologic cancers.

References:
[1] https://www.mll.com/en/science/5000-genome-project.html
[2] Yepez et al., Nat Protoc (2021)
[3] Wagner et al., Nat Genetics (2023)
[4] Martínez-Jiménez et al., Nat Rev Cancer (2020)

72

Multi-species reference free transcriptomic immune estimation with ConsensusTME

Oliver Cast[3], Christabel Boyes[2], Anne Machel[2], Chrysa Kapeni[2], Michael Gill[2], Isidro Cortes Ciriano[3], Martin Miller[1]

1 Astrazeneca, United Kingdom
2 CRUK Cambridge Institute - University of Cambridge, United Kingdom
3 EMBL-EBI, United Kingdom

Presenter: Oliver Cast

The Tumour Microenvironment (TME) plays a critical role in shaping cancer initiation, progression, and response to therapies. While single-cell RNA sequencing (scRNA-seq) offers unparalleled resolution, bulk RNA-seq remains a cost-effective and widely used approach for analysing large, clinically annotated cohorts. To bridge this gap, we present ConsensusTME, an ensemble-based cell estimation method that integrates features from various existing methods and datasets to generate a comprehensive gene set compendium. Our results demonstrate that, when combined with single-sample Gene Set Enrichment Analysis (ssGSEA), these gene sets reliably estimate immune and stromal cell abundance in both human and mouse datasets. Comprehensive benchmarking shows ConsensusTME provides accurate and robust estimation, particularly in the context of the TME. Furthermore, we highlight the utility of these gene sets in annotating scRNA-seq and spatial transcriptomics data, using 10x Visium as a case study. ConsensusTME is made freely available as an R package (https://github.com/cansysbio/ConsensusTME) and through an online web portal (http://www.consensustme.org/), which also serves as a repository for exploring existing deconvolution methods, benchmarking datasets, and relevant review articles.

73

Differential oncogenic pathway expression within cell cycle phases when comparing breast cancer subgroups at a single-cell level

Miguel Castresana Aguirre[1], Nick Tobin[1]

[1] Karolinska Institute, Sweden

Presenter: Miguel Castresana Aguirre

Background: Breast tumours have traditionally been categorized into four widely recognized molecular subtypes. Recent single-cell advancements have however shown that individual breast cancer tumors can contain more than one subtype, highlighting their intrinsic heterogeneity. Similarly, broad heterogeneity in the cell cycle phase of individual epithelial cells is a common feature of breast tumours. Considering these concepts together, it is currently unclear how oncogenic signalling pathways differ between breast cancer subtypes within the context of individual cell cycle phases. As such, our aim is to elucidate the differences in oncogenic signaling pathways across distinct phases of the cell cycle (G1, S and G2M) for each breast cancer subtype (Luminal A, B, Her2-enriched and Basal-like), using single-cell data.
Methods: For this study we analyzed RNA-sequencing data from 100 064 single cells taken from a public dataset of 26 breast tumours. After filtering based on cell type, cell malignancy and single-cell breast cancer subtype annotation, 24 489 cancerous epithelial cells remained. Differential gene expression analysis between the four breast cancer subtypes per cell cycle phase was carried out using Limmatrend. Pathway analysis was then performed using FGSEA with a focus on hallmarks of cancer gene sets taken from MSigDB.
Results: We found 23 significant hallmark pathways differentially expressed ($P \leq 0.05$) in at least one comparison. The most frequent included TNFA signaling via NFKB and Interferon gamma response, which were enriched in all cell cycle phases of Luminal A and B subtypes respectively. A cluster of hallmarks including E2F-targets, G2M checkpoint, Mitotic spindle and MYC targets were enriched for all subtypes except Luminal A during G2M was also observed. Finally, and notably, the Epithelial to Mesenchymal transition hallmark was enriched for the Basal subtype in G1 only.
Conclusions: This study provides a comprehensive depiction of the biological pathways altered in each cell cycle phase when comparing single cell breast cancer subtypes. More broadly, our research aims to lead to a better understanding of the underlying biology within breast cancer tumors in order to better identify potential drug targets for the treatment of patients with precision medicine strategies.

74

Resolving chromatin organization in chronic lymphocytic leukemia

Karolína Cernovská[2], Jakub Pawel Porc[5], Kristýna Závacká[1], Kamila Stránská[1], Sabina Adamová[4], Eva Ondrouskova[6], Marie Jarošová[4], Šárka Pospíšilová[3], Karla Plevová[7]

1 CEITEC, Masaryk University & University Hospital Brno, Czech Republic
2 Masaryk University & CEITEC, Masaryk University, Czech Republic
3 Masaryk University & CEITEC, Masaryk University & University Hospital Brno, Czech Republic
4 Masaryk University & University Hospital Brno, Czech Republic
5 Masaryk University & Central European Institute of Technology, Masaryk University, Czech Republic
6 University Hospital Brno, Czech Republic
7 University Hospital Brno, Masaryk University & Central European Institute of Technology, Masaryk University, Czech Republic

Presenter: Karolína Cernovská

Introduction: The spatial chromatin organization has been recognized as essential for fundamental cell processes. Disruption of its hierarchical structure can be caused by chromosomal rearrangements, which, especially in cancer cells, lead to aberrant molecular signaling. We aim to study altered genome structure in chronic lymphocytic leukemia (CLL), the most common type of leukemia in adults. We are particularly interested in the analysis of long-distance chromatin interactions and the formation of topologically associated domains (TADs), which might play a role in CLL pathogenesis.
Methods: Chromatin crosslinking was carried out on B cells separated from the peripheral blood of 8 CLL patients. Further sample processing was performed using Micro-C (Dovetail Genomics) and NEBNext Ultra II DNA Library Prep (New England Biolabs) kits to generate comprehensive whole genome contact maps. Then we used bioinformatic tools to analyze copy number alterations (CNAs), structural variants (SVs) and neoTAD formation.
Results: We obtained data allowing for a detailed analysis of DNA interactions. Genomic rearrangements identified in contact maps as interchromosomal were primarily compared with mFISH results and CNA plots were compared with genomic array-based results. Overall, using the chromatin conformation analysis, we identified abnormalities detected by conventional methods with quantitative and qualitative differences related to the method resolution, sensitivity, and other molecular (e.g., epigenetic) mechanisms. Additionally, we described neoTAD formation in all CLL samples as a consequence of chromosome shuffling. Detailed analysis of chromatin reorganization and its effects on CLL pathophysiology is ongoing.
Conclusion: By analyzing chromatin interactions in CLL samples, we want to contribute to the knowledge of leukemia genomic complexity and related abnormal signaling. Our data show that the Micro-C method has great potential for detecting SVs, including balanced and cryptic translocations that remain hidden from other techniques. The advantage of the method lies in the additional information about DNA interactions in specific regions and novel TAD formation.

Supported by MHCZ-AZV NU21-08-00237, MHCZ-DRO FNBr65269705, MUNI/A/1224/2022, NPO-NUVR LX22NPO5102.

75

Mutational signatures of ethanol and acetaldehyde in oral cancer: an experimental modelling approach

Bérénice Chavanel1, Eva Tibaldi2, Cécilia Sirand1, François Virard1, Christine Carreira1, Foster Jacob3, Vincent Cahais1, Behnoush Abedi-Ardekani1, Daniele Mandrioli2, Silvia Balbo3, Zdenko Herceg1, Michael Korenjak1, Jiri Zavadil1

1 International Agency for Research on Cancer WHO, France
2 Cesare Maltoni Cancer Research Center, Ramazzini Institute, Italy
3 Masonic Cancer Centre, University of Minnesota, United States of America

Presenter/Co-author: Bérénice Chavanel

Alcohol use is associated with cancer development at various anatomical sites including oral cavity, and is responsible for ~13% of all cancers worldwide. Despite strong epidemiological evidence, the mechanisms of ethanol carcinogenicity in the oral cavity remain unclear. Ethanol's main metabolite acetaldehyde (AcA) may play a crucial role in head and neck cancers by forming covalent DNA adducts which, if not repaired, may ultimately lead to mutagenesis and cancer development. The COSMIC mutational signature SBS16 has been tentatively linked to alcohol drinking, yet it is also found in pathologies unrelated to alcohol exposure. The direct effects of ethanol/AcA in well-controlled experimental settings have not been clearly established. We hypothesized that the role of AcA in alcohol-related oral cancer is based on the formation of specific mutational signature(s) which can be modelled in suitable experimental systems.
Using a multi-system experimental approach, we aimed to characterize the mutagenic modes of action of ethanol and AcA at the genome scale level. Firstly, oral tumour tissues derived from ethanol- and AcA-driven carcinogenesis studies in longitudinally exposed rats are analyzed by whole genome-sequencing. The animal study is complemented by in vitro chronic exposure of non-tumour hTERT-immortalized oral cell lines to AcA, followed by clonal expansion and whole-genome sequencing. Mutational signatures are identified in the genomes of the rat tumours and exposed cells and are matched with LC-MS/MS DNA adductomics data generated from the cell exposure models.
Preliminary results were obtained from oral squamous cell carcinomas collected from the cheeks of rats exposed to 10% ethanol in drinking water. We observed a possible role for inflammation and oxidative DNA damage processes linked to cell keratinization, and the presence of COSMIC signature SBS17. Signature SBS16 has not been observed in the analyzed rat oral tumours, and the findings are being extended to additional cancer sites, including the zymbal gland or forestomach, where ethanol/AcA exposure-associated tumour formation had been observed.
We anticipate that this study will improve our understanding of the mechanisms by which ethanol and AcA induce oral carcinogenesis, to ultimately support cancer prevention measures.

76

RNAi and chemogenomic screens to identify novel modulators of the nonsense-mediated decay pathway

Ken Cheng[1]

[1] National Center for Advancing Translational Sciences, National Institutes of Health, United States of America

Presenter: Ken Cheng

The nonsense-mediated mRNA decay (NMD) pathway is the primary mechanism to eliminate aberrant transcripts containing premature termination codons. Many genetic diseases, including some cancers result from the introduction of a premature termination codon, which potentially cause loss-of-function or dominant-negative phenotypes. In addition to mRNA quality-control surveillance, studies have shown that the NMD machinery regulates the expression of ~5-10% of the mammalian transcriptome, highlighting its importance in modulating gene expression. To search for novel human NMD components and compounds that modulate the NMD pathway, we screened a well-characterized, luciferase-based β-globin NMD assay against a whole-genome RNAi (>21,000 genes) and an annotated small molecule (>5,000 compounds) library. From the hit list of the siRNA screen, core NMD and EJC components such as SMG1 and UPF1 were ranked as top candidates, and several potential NMD factors were also being identified. Using biochemical and genomic approaches, we have confirmed some of the hits as the novel factors that modulate NMD and splicing. In addition, compounds that regulate NMD or splicing were being selected from the small molecule screening. Currently we are at the stage of validating the molecular targets and investigating the mode of action of these chemicals.

77

MHC class II genotypes are independent predictors of anti-PD1 immunotherapy response in melanoma

Arne Claeys[1], Jimmy Van den Eynden[1]

[1] Ghent University, Belgium

Presenter: Arne Claeys

Immune checkpoint blockade is a highly successful anti-cancer immunotherapy. Both CTLA4 and PD1 checkpoint blockers are clinically available for melanoma treatment, with anti-PD1 therapy reaching response rates of 35-40%. These responses, which are mediated via neoantigen presentation by the polymorphic MHC complex, are hard to predict and the tumor mutation burden (TMB) is currently the best available biomarker. While responses are expected to be strongly determined by the MHC genotype, single allele association studies have remained elusive, likely due to the low population allele frequency and resulting lack of statistical power. To solve this problem, we quantified the complete MHC genotype based on the functional MHC binding properties and demonstrate that the MHC class I (MHC-I) and class II (MHC-II) genotypes are TMB-independent predictors of anti-PD1 immunotherapy responses in melanoma. Opposite responses were found for both classes, with strong MHC-I and MHC-II binding genotypes predicting good and bad responses, respectively. Interestingly, MHC-II binding genotypes were mainly associated with treatment response failure in a subgroup of anti-CTLA4 pretreated patients. As MHC genotyping does not require tumor material and/or advanced sequencing, it has clinical potential as a non-invasive and tumor-independent biomarker to guide anti-cancer immunotherapy in melanoma.

78

Decoding the evolution and heterogeneity of Peripheral T-cell Lymphoma through deep multiomic sequencing

Ruben Cools[2], Jonas Demeulemeester[2], Marlies Vanden Bempt[3], Daan Dierickx[1], Koen Debackere[1]

[1] KU Leuven - Laboratory of Experimental Hematology, Belgium
[2] VIB - KU Leuven Center for Cancer Biology, Laboratory of Integrative Cancer Genomics, Belgium
[3] VIB - KU Leuven Center for Cancer biology, Laboratory of Molecular Biology of Leukemia, Belgium

Presenter: Ruben Cools

Peripheral T-cell Lymphomas (PTCLs) are a heterogeneous group of non-Hodgkin Lymphomas that originate from mature T lymphocytes. So far, several subtypes of PTCL have been identified but ~30% of cases defy classification, missing distinct features to allocate them to another subtype. For these patients chemotherapy is the only available treatment option, resulting in a 75% relapse rate. Leukemias and lymphomas often harbour multiple subclones with distinct genetic alterations, which may contribute to this high degree of treatment resistance. The relative rarity of PTCL has however precluded characterization of its intratumour heterogeneity and evolution, which in turn has hampered development of novel therapies.
We profiled a unique cohort of 28 PTCL patients with deep whole-genome (150-300x) and transcriptome sequencing to chart the mutational landscape of the disease and identify distinct clinical entities. We observed a diverse driver landscape, with gene fusions and point mutations activating T-cell receptor signaling (e.g., CD3E); multiple hits to epigenetic regulators such as TET2; homozygous deletions of CDKN2A/B and high-level amplifications of known oncogenes. Analysis of mutational processes revealed surprising contributions from COSMIC signatures SBS7 and/or 17 in the majority of cases, associated with UV damage and upper gastrointestinal tract tumours, respectively. This suggests that precursors of the malignant clones resided in the skin, or gut-associated lymphoid tissues. Performing subclonal reconstruction, we discovered extensive genetic intratumour heterogeneity in these rare lymphomas, with multiple large-scale subclonal copy number aberrations being the norm.
Future work, including evolutionary analyses will reveal the timing of the mutagenic processes and their contribution to mature T cell transformation. This will allow us to generate an in-depth reconstruction of the genomic trajectories of PTCL, enhancing our understanding of the disease biology, aiding diagnosis, and provide new potential therapeutic targets in the long run.

79

Disentangling the chromatin accessibility landscape of Multiple Myeloma patients identified the transcription factor NRF1 a key driver for neoplastic development and progression

Giacomo Corleone[1], Tiziana Bruno[1], Stefano Di Giovenale[1], Clelia Cortile[1], Francesca De Nicola[1], Maria Chiara Cappelletto[1], Ludovica Ciuffrida[1], Francesco Marchesi[1], Svetlana Gumenyuk[1], Andrea Mengarelli[1], Maurizio Fanciulli[1]

[1] IRCCS Regina Elena National Cancer Institute, Italy

Presenter: Giacomo Corleone

Multiple myeloma (MM) is a hematological malignancy characterized by the clonal proliferation of plasma cells within the bone marrow. Despite significant efforts in understanding the underlying pathogenesis of MM and the development of novel therapeutic approaches, it remains an incurable disease. Recent evidence suggests that MM carcinogenesis involves epigenetic alterations, in addition to genetic factors. We conducted a comprehensive analysis of chromatin accessibility in 198 patients at various stages of the disease, including diagnosis, post-therapy, relapse, and the pre-malignant state (MGUS). We found that pervasive open chromatin loci in the majority of samples at diagnosis and relapse are enriched for the transcription factor NRF1. ChIP-seq analysis of NRF1 in a large cohort of patients confirmed that it is a key mediator in sustaining the activity of pervasive de novo promoter and enhancers. These elements are active only in the neoplastic condition and promote cellular growth, proliferation, and survival pathways. Interestingly, these elements are not bound by NRF1 and are in a closed chromatin state in pre-malignant lesions, suggesting that they are specific to the active disease state. We used pan-cancer integrative datasets and HiC of MM cell lines to assign target genes to the identified NRF1-bound elements. We then identified a distinctive NRF1 dependent transcriptional signature strongly correlated with aggressive disease and deadly outcome using ~700 patients from the CoMMPASS dataset. Collectively, our findings provide new insights into the role of chromatin variability in driving the development and progression of MM. We have shown that NRF1 is a key mediator of chromatin accessibility changes in MM, and that these changes are associated with the activation of numerous proliferative mechanisms that are advantageous to the cancer. Our findings suggest that targeting NRF1 may be a promising therapeutic strategy for MM.

Giacomo Corleone is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

80

Impaired formation of a stable RPA/RNAseH1 complex in senescent cells leads to uncontrolled processing of R-loops and unsuccessful DNA repair

Ylenia Cortolezzis[1], Vanessa Tolotto[1], Gabriele Magris[1], Luigi Xodo[1], Eros Di Giorgio[1]

[1] University of Udine, Italy

Presenter: Ylenia Cortolezzis

The maintenance of genome stability ensures cellular fitness. Cells are constantly exposed to DNA-damaging agents and DNA integrity is maintained by successful activation of the DNA damage response (DDR). Senescent cells accumulate unrepaired DNA due to a defective DDR. Noncanonical chromatin structures (NCSs) enable the maintenance of genome integrity by controlling genome replication, transcription, and DNA repair. Among NCSs, R-loops are three-stranded DNA-RNA hybrids that have been reported to recruit repair factors and maintain successful DNA repair when formed in trans and to increase genome instability when formed co-transcriptionally in cis.
Here, we used BJ/hTERT- RAS-ER as an inducible cellular model of oncogene-induced senescence (OIS) to investigate the contribution of R-loops to the maintenance of genome stability. Cells undergoing OIS accumulate more R-loops compared to proliferating fibroblasts, but only a small fraction of them support BRCA1 loading. By crossing DRIP -seq and DRIVE -seq data, we found that impaired loading of DDR proteins at double-strand breaks (DSBs) correlates with unsuccessful recruitment of RNAseH1 by the RPA complex to R-loops. By adopting the LacO/LacR tethering system, we found that forced recruitment of both catalytically active and inactive RNAseH1 to DSBs delays DNA repair, suggesting that both premature and delayed removal of R-loops affects DDR.
Finally, our research identified phosphorylation of RPA32 as the signal that controls the release of RNAseH1 from DSBs. Hyperphosphorylation of the RPA complex observed in pre-senescent cells alters this mechanism, leading to further accumulation of irreparable damage and onset of senescence.

81

Deciphering mechanisms of T-ALL relapse on a single cell multi-omic level

Julia Costea[3], Jan Korbel[1], Andreas Kulozik[2]

[1] EMBL Heidelberg, Germany
[2] German Cancer Research Center, Hopp Children's Cancer Center Heidelberg (KiTZ), University Clinic Heidelberg, Germany
[3] Molecular Medicine Partnership Unit, EMBL Heidelberg, Germany

Presenter: Julia Costea

The mechanisms underlying relapse in paediatric T cell acute lymphoblastic leukaemia (T-ALL) are incompletely understood. While initial diagnoses of T-ALL patients typically see a high success rate with long-term survival rates of approximately 80%, relapsed patients demonstrate treatment resistance and have a much lower survival rate. We have compared groups of relapsing and non-relapsing patients to follow the evolution of individual clones throughout the disease by integrating data from 1) single cell strand-specific sequencing (Strand-Seq; N=15 samples) which can detect structural variations and correlated changes of the nucleosome occupancy and from 2) VASA-Seq (N=15 samples), a single cell total RNA-seq method. Analysing the individual tumour evolution from initial disease to relapse, we identified a cluster in all patients with an outstanding strong cluster marker (log2FC >2), called AHNAK, which expanded significantly (padj <0.05) in the relapse of all patients. As other studies also identified a role of AHNAK in treatment resistance of various cancer types, this indicates a more general function of this factor for the outcome of the patient. Following up on this, we use computational tools, like SCENIC, to identify underlying gene regulatory networks to understand its specific role in the context of cancer. Analysis of the Strand-seq data revealed a previously unknown relapse-specific translocation fusing the T cell receptor and the Growth Hormone Receptor resulting in an upregulation of the latter. VASA-Seq confirmed the upregulation in all relapse cells as well as in two low abundant clones (<1%) during the initial disease that later formed the relapse. This underlines the importance of integrating single cell multi-omic datasets for deciphering mechanisms in T-ALL, and shows how the application of novel methods can provide additional genetic insights even in samples that have been extensively studied before.

82

Germline genetics correlates with aberrant signaling pathways in cancer

Davide Dalfovo[1], Riccardo Scandino[1], Marta Paoli[1], Samuel Valentini[1], Alessandro Romanel[1]

[1] University of Trento, Italy

Presenter: Davide Dalfovo

Cancers are complex diseases influenced by a heterogeneous landscape of both germline genetic variants and somatic aberrations. Although a large number of somatic aberrations in oncogenic signaling pathways are now used as hallmarks in many well-known forms of cancer, the interaction landscape of germline variants and aberrant signaling pathways is still largely unknown. A thorough understanding of germline/somatic links and their impact on cancer genesis and progression is thus needed.
We conducted a collection of genome-wide association studies (GWAS) using >8,500 human samples across 33 cancer types characterized by TCGA, considering 40 somatic traits defined using a large collection of somatic aberration types across 10 well-known oncogenic signaling pathways.
Functional links were established between associated variants and the corresponding oncogenic signaling pathways using genomic and transcriptomic data. Then, GWAS summary statistics were used to build polygenic somatic scores (PSS) using a cross-validation approach. The statistical significance were determined using a permutation analysis. Predictive accuracy of PSS was validated using PCAWG data and cancer specific cohorts and tested against TCGA cancer subtypes, cancer-specific clinical variables and patient's survival data.
We identified 58 genome-wide significant ($p<5e-8$) associations between 51 SNPs and 16 somatic traits and an additional 268 suggestive ($p<1e-6$) associations between 240 SNPs and 35 somatic traits. 138 associated SNPs revealed cis-associations with genes that were observed functionally correlated with the corresponding somatic traits oncogenic pathways. PSS of 24 somatic traits showed an FDR<0.25 in the TCGA dataset and 5, among 15 that could be tested, demonstrated consistent statistically significant shifts across PCAWG patients based on the corresponding oncogenic pathway aberration status; two were further confirmed in CCLE cell lines samples.
We showed that germline genetics shapes susceptibility to somatic aberrations in oncogenic signaling pathways and that polygenic scores can describe patients' genetic liability to develop specific cancer molecular profiles.

83

Phylogenetic analysis shows genetic bottlenecks in embryogenesis

Manas Dave[2], Tim Coorens[1], Sam Behjati[2]

[1] Broad Institute of MIT and Harvard, United States of America
[2] Wellcome Sanger Institute, United Kingdom

Presenter: Manas Dave

Investigating human embryonic development is challenging with much research undertaken on model organisms and the findings extrapolated to humans. Recently, we have been able to retrace development using mutations that serve as evolutionary barcodes detailing the history of a particular cell. By studying normal development, we gain insights into abnormal development including the origins of childhood cancers. However, to fully study early development from the zygote, tissue samples are needed from the embryo proper, extra-embryonic tissues and the placenta.
We used laser capture microdissection to sample and whole-genome sequence over 1200 tissue components within every human organ in a 11-week post-conception foetus. Additionally, we cultured and whole-genome sequenced organoids from all three germ layers and constructed a high-resolution phylogenetic tree demonstrating normal development.
We reconstructed phylogenetic relationships between all organs and the placenta, and showed the latter has a disproportionally high mutational burden compared with all organs. For somatic organs, we traced cell divisions to the point of commitment of differentiation into organs, highlighting genetic bottlenecks and biological processes that positively select clones leading to organ formation. We have high granularity data enriched for certain organs such as the heart and kidneys which has allowed novel insights into relationships between spatially distinct tissue regions that are histologically identical. This has informed our understanding of normal development, a necessity to truly understanding developmental disorders and the origins of childhood cancers.

84

LIFR-AS1 promoter methylation as a potential diagnostic biomarker differentiating colorectal cancer and colorectal liver metastases from other adenocarcinomas

Tina Draškovic[1], Branislava Rankovic[1], Nina Zidar[1], Nina Hauptman[1]

[1] University of Ljubljana, Slovenia

Presenter: Tina Draškovic

Background & objective: The diagnosis of colorectal cancer (CRC) and colorectal liver metastases (CRLM) and its differentiation from primary liver adenocarcinomas and other liver metastases is sometimes challenging. The aim of our study was to investigate the methylation status of the promoter region of Leukemia inhibitory factor receptor antisense RNA1 (LIFR-AS1) and its significance as a potential novel DNA methylation biomarker for CRC and CRLM.
Methods: Bioinformatics analysis revealed the LIFR-AS1 promoter region as a candidate for differentiation among adenocarcinomas. DNA was isolated from formalin-fixed, paraffin-embedded tissue samples from 15 CRCs, its 9 paired CRLMs, 15 hepatocellular carcinomas, 15 cholangiocarcinomas, 15 pancreatic adenocarcinomas, 14 stomach adenocarcinomas and their paired adjacent non-tumor tissue samples. After DNA isolation and DNA-bisulfite conversion, the methylation status of the LIFR-AS1 promoter region was determined by methylation-sensitive high-resolution melting analysis. Hyper- or hypomethylation of each sample was used to determine sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).
Results: The methylation status of the LIFR -AS1 promoter region successfully differentiates between CRC and adjacent normal tissue, with a sensitivity of 86.7%, specificity of 100%, PPV of 1 and NPV of 0.88. This methylation biomarker can differentiate between CRC and all included primary tumor and adjacent normal tissue samples with a sensitivity of 86.7%, specificity of 97%, PPV of 0.76, and NPV of 0.98 for CRC. Moreover, the methylation status of the LIFR-AS1 promoter region from CRC samples is preserved in all paired CRLM samples, of which 7 were hypermethylated and 2 were hypomethylated. It shows a sensitivity of 77.8%, a specificity of 97%, a PPV of 0.64, and an NPV of 0.97 for CRLM.
Conclusion: Our results show that the DNA methylation status of the promoter region of the LIFR-AS1 is preserved from primary CRC to paired CRLM. It shows high sensitivity, specificity, PV and NPV for differentiating CRC and CRLM from other adenocarcinomas and adjacent non-tumor tissues. We propose a novel DNA methylation biomarker, the promoter region of LIFR-AS1, for diagnosis and differentiation of CRC and CRLM from other adenocarcinomas.

85

SAVANA: a computational method to characterise structural variation in human cancer genomes using long-read sequencing

Hillary Elrick[1], Jose Espejo Valle-Inclan[1], Katherine Trevers[2], Francesc Muyas[1], Solange de Noon[2], Adrienne Flanagan[2], Isidro Cortes Ciriano[1]

[1] EMBL-EBI, United Kingdom
[2] University College London Cancer Institute, Royal National Orthopaedic Hospital, United Kingdom

Presenter: Hillary Elrick

Whole-genome sequencing (WGS) of human cancers has revealed that structural variation, which refers to the rearrangement of the genome leading to the deletion, amplification of reshuffling of DNA segments ranging from a few hundred bp to entire chromosomes, is a key mutational process in cancer evolution. Notably, pan-cancer analyses have revealed that both simple and complex forms of structural variation are pervasive across diverse human cancers, and often underpin drug resistance and metastasis. To date, the study of cancer genomes has relied on the analysis of short-read WGS on the dominant Illumina platform, which generates short, highly-accurate reads of 100-300bp that allow the study of point mutations at high resolution. However, detection of structural variants (SVs) using short reads is limited, as breakpoints falling in repetitive regions cannot be reliably mapped to the human genome. As a result, our understanding of the patterns and mechanisms underpinning structural variation in cancer genomes remains incomplete.Here, we present SAVANA, a novel structural variant caller for long-read sequencing data specifically designed for the analysis of cancer genomes. To identify both somatic and germline SVs, SAVANA takes as input long-read WGS data from a tumour and normal sample pair. SAVANA scans sequencing reads to detect split reads and gapped alignments, which are then clustered to define putative SVs. Next, SAVANA applies a machine learning-informed set of heuristics to remove false positives arising from mapping errors and sequencing artifacts.Extensively validated against a multi-platform truthset, we show that SAVANA identifies a range of somatic rearrangements with high recall and precision, outperforming existing tools while maintaining a lower execution time than competing methods. In patient samples, SAVANA identifies clinically relevant alterations, such as oncogenic gene fusions, with high accuracy. Additionally, SAVANA permits the reconstruction of double minutes, multi-chromosomal chromothripsis events, and SVs mapping to highly repetitive regions, including centromeres. In sum, SAVANA permits the characterization of complex structural variants and can uncover clinically relevant mutations across diverse cancer types with high accuracy.

86

Evolutionary trajectories of complex genome rearrangements in cancer

Jose Espejo Valle-Inclan[1], Solange de Noon[2], Katherine Trevers[2], Hillary Elrick[1], Adrienne Flanagan[2], Isidro Cortes Ciriano[1]

1 EMBL-EBI, United Kingdom
2 University College London Cancer Institute, Royal National Orthopaedic Hospital, United Kingdom

Presenter: Jose Espejo Valle-Inclan

Whole-genome sequencing studies of human tumors have revealed that cancer genomes are riddled by remarkably intricate forms of structural variants (SV), collectively known as complex genomic rearrangements (CGR). CGR occur at high frequency in some of the most aggressive cancers and are associated with drug resistance and poor prognosis. Yet, most CGR remain unexplained, hinting at the possibility of undiscovered mechanisms that might lead to novel therapeutic strategies. To elucidate the mechanisms underpinning the formation of CGR and their downstream consequences during tumor evolution, we performed high-depth (>120x) multi-regional short- and long-read whole-genome sequencing of hundreds of osteosarcomas, which exhibit some of the highest rates of CGR across human cancers. Our analysis revealed that whole-genome doubling and diverse types of CGR, such as chromothripsis and double minutes, are frequent clonal events that occur in most tumors. Through the integration of multi-regional WGS data we show that subclonal CGRs are also frequent events, triggering clonal diversification and rapid tumor growth suggestive of punctuated evolution. Indeed, clonal expansions triggered by the acquisition of subclonal CGR often colonize distant tumor regions, indicating that CGR act as subclonal driver events. Notably, metastatic clones often arise very early during tumor evolution (years before diagnosis) and often expand after the occurrence of CGR. Finally, we find that derivative chromosomes generated by chromothripsis events in the early stages of tumor evolution, including double minutes, acquire hundreds of subclonal SV, which increases intra-tumor heterogeneity and cancer cell plasticity. Thus, our results indicate that chromothripsis is a process that occurs throughout tumor evolution and primes the cancer genome for ongoing genomic instability, leading to SV accumulation, rapid karyotype evolution and clonal diversification. These results have implications for intra-tumor heterogeneity and drug resistance development in cancer types driven by genomic instability.

87

The translation of non-canonical ORFs is associated with the generation of neoantigens in hepatocellular carcinoma

Marta Espinosa[2], Patrick Theunissen[1], Pablo Sarobe[1], Puri Fortes[1], Júlia Perera-Bel[2], M.Mar Albà[3]

[1] Cima Universidad de Navarra, Spain
[2] Hospital del Mar Medical Research Institute, Spain
[3] Hospital del Mar Medical Research Institute, Catalan Institute for Research and Advanced Studies, Spain

Presenter: Marta Espinosa

Non-canonical peptides derived from the translation of open reading frames (ORFs) in tumor-specific long non-coding RNAs and novel transcripts are a potentially important class of cancer neoantigens, but they remain poorly characterized. Here we present a novel computational pipeline to identify tumor-specific translated ORFs using tumor/healthy tissue matched RNA-Seq and Ribo-seq data. We have applied this pipeline to RNA-seq data from a meta-cohort of 117 hepatocellular carcinoma (HCC) patients. In order to quantify the number of different types of neoantigens, we have estimated the affinity between cancer-specific peptides and MHC I receptors and investigated cancer immuno-peptidomics data. Some of these predictions were validated by in vitro MHC-I binding assays and a number of non-canonical peptides were demonstrated to elicit an immune reaction when injected in mice. The results indicate that tumor-specific peptides make a larger contribution to the landscape of neoantigens in HCC than proteins carrying single mutations. In addition, these peptides are often shared across different patients. One example is AC079466.1, a tumor-specific long non-coding RNA that generates a peptide with strong affinity for MHC I which is capable of eliciting an immune response in mice. This lncRNA is expressed in 23% of the patient tumors, but it is not found in healthy tissue from the same patients. This study suggest that non-canonical peptides are likely to have a more important role in mediating cancer immunogenicity than initially anticipated, opening new avenues for novel anti-cancer treatments such as vaccines.

88

The evolution of chromosomal instability in prostate cancer

Ángel Fernández-Sanromán[1], Geoff Macintyre[1], Barbara Hernando[1], Atef Sahli[2], Christopher Hovens[3], Joachim Weischenfeldt[4], David Wedge[2]

1 Spanish National Cancer Research Centre (CNIO), Spain
2 The University of Manchester, United Kingdom
3 The University of Melbourne, Australia
4 University of Copenhagen, Denmark

Presenter: Ángel Fernández-Sanromán

Prostate cancer shows substantial variation in clinical outcomes, ranging from indolent tumours to aggressive metastatic forms resistant to standard of care treatment. Unraveling the mechanisms driving these varied phenotypes is crucial for improved patient care. While past studies have emphasized the role of chromosomal instability (CIN) in advanced prostate cancer, it is unknown when CIN is induced and how it evolves during cancer progression. To address this gap, here we reconstructed the point mutation (SBS) and copy-number alteration (CNA) history of 337 samples (10 benign prostate hyperplasia, 29 normal, 175 primary and 123 metastatic samples) from 127 prostate cancer patients. We found that the ratio of CNAs to SBS was significantly higher in the alterations private to metastatic samples than in truncal alterations and alterations private to primary samples, suggesting an acceleration of CIN linked to metastasis. Metastases also exhibit enrichment of copy-number signatures previously linked to whole-genome doubling (CX4), homologous repair deficiency (CX2, CX3, CX5) and replication stress (CX8). These observations imply the induction of specific CIN-generative mechanisms contributing to increased CNA burden in metastatic tumors. Taken together, our results highlight that CIN-generative mechanisms change their activity during the evolution of prostate cancer and that some of these mechanisms might become predominantly active in metastatic cells. The identification of these mechanisms at the latest disease stages might be fundamental to guide treatment in the future, given that recent treatment modalities, such as PARPi and docetaxel, seem to be particularly effective in tumours with specific CIN-generative mechanisms.

Ángel Fernández-Sanromán is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

89

Prevalence, causes and impact of TP53-loss phenocopying events in human tumors

Bruno Fito[1], Marina Salvadores Ferreiro[1], Miguel M Alvarez[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Fran Supek

TP53 is a master tumor suppressor gene, altered in approximately half of all human cancers. Given the many regulatory roles of the corresponding p53 protein, it is possible to infer loss of p53 activity – which may occur due to genetic alterations in trans – from gene expression patterns. Several such alterations that phenocopy p53 loss are known, however additional ones may exist, but their identity and prevalence among human tumors are not well characterized.

We performed a large-scale statistical analysis on transcriptomes of ~ 7,000 tumors and ~ 1,000 cell lines, estimating that 12% and 8% of tumors and cancer cell lines, respectively, phenocopy TP53 loss: they are deficient in the activity of the p53 pathway, while not bearing obvious TP53 inactivating mutations. While some of these cases are explained by amplifications in the known phenocopying oncogenes MDM2, MDM4 and PPM1D, about half of the phenocopying tumors do not have an obvious causal alteration.

An association analysis of cancer genomic scores jointly with CRISPR and RNAi genetic screening data identified an additional candidate p53-loss phenocopying gene, USP28. We estimate that deletions in USP28 are associated with a p53 functional impairment in 2.9–7.6% of breast, bladder, lung, liver and stomach tumors, and have comparable effect size to MDM4 amplifications. Additionally, in the known copy number alteration (CNA) segment harboring MDM2, we identify an additional co-amplified gene (CNOT2) that may boost the TP53 functional inactivation effect of MDM2.

A reanalysis of cancer cell line drug screening data, in a manner aware of phenocopy scores, suggests that p53 pathway (in)activity commonly modulates associations between drug effects and various genetic markers, such as PIK3CA and PTEN mutations. As a resource, we provide the drug-genetic marker associations that differ depending on p53 functional status.

In conclusion, human tumors that do not bear obvious TP53 genetic alterations but that phenocopy p53 loss-of-function are common, and the USP28 gene deletions are one likely cause thereof.

This study was recently published as Fito et al. BMC Biology, 2023. https://doi.org/10.1186/s12915-023-01595-1

90

Defining the Involvement of the Perivascular Niche in Brain Tumour Metastases

Vernon Fong[2], Namal Abeysundara[1], Bryn Livingston[1], Anders Erickson[1], Cory Richman[1], Olga Sirbu[1], Vijay Ramaswamy[1], Michael Taylor[1]

[1] The Hospital for Sick Children, Canada
[2] University of Toronto, Canada

Presenter: Vernon Fong

Medulloblastomas (MB) is the most prevalent malignant pediatric brain tumours and arises within either the cerebellum or the dorsal brainstem. MBs are comprised of four subgroups (WNT, SHH, Group 3, and Group 4) that have distinct gene expression profiles and differential patterns of metastasis and recurrence. Poor prognosis is often accompanied with MB metastatic progression, which is estimated to be present in 30% of MB patients at diagnosis. Unfortunately, most research focuses on the primary tumour and there are no current specific treatments approved for MB leptomeningeal metastases. Therefore, further insight is needed in order to develop treatments that target metastatic MB. The leptomeningeal niche, where metastatic MB tumour cells reside, contains a vast network of blood vessels that is supported by perivascular cells such as pericytes, smooth muscle cells, and fibroblasts. However, the specific adhesion molecules and signalling pathways from the leptomeningeal niche that contribute to tumour cell colonization remain unclear. I hypothesize that the perivascular niche is important for metastatic cell colonization and survival in the leptomeninges. I have utilized a cutting-edge sLP-mCherry niche labelling system and single-cell RNA sequencing to spatially identify and investigate the perivascular cells neighbouring the metastatic MB tumour cells. Single-cell RNA sequencing analysis were performed on mCherry+ perivascular cells in close proximity to tumour cells and distant mCherry- perivascular cells. Genes associated with cholesterol export were identified to be upregulated in tumour associated perivascular cells, suggesting a role in supporting metastatic cell colonization. Lentiviral overexpression and knockdown constructs were performed to interfere with cholesterol uptake in the MB cell lines, and changes in metastatic burden were observed. This investigation into the mechanism of lipid transport that allows for metastatic tumour cells to survive in the vascular microenvironment will uncover potential targets for future treatment for MB metastases.

91

Preclinical exploration of the DNA Damage Response pathway using the interactive neuroblastoma cell line explorer CLEaN

Jonatan Gabre[1], Jimmy Van den Eynden[1], Ruth Palmer[2], Peter Merseburger[1], Joachim Siaw[2], Arne Claeys[1], Bengt Hallberg[2], Frank Speleman[1], Sarah-Lee Bekaert[1]

1 Ghent University, Belgium
2 University of Gothenburg, Sweden

Presenter: Jonatan Gabre

Neuroblastoma is the most common cancer in infancy with an urgent need for more efficient targeted therapies. The development of novel (combinatorial) treatment strategies relies on extensive explorations of signaling perturbations in neuroblastoma cell lines, using RNA-Seq or other high throughput technologies (e.g., proteomics, phosphoproteomics). This typically requires dedicated bioinformatics support, which is not always available. Additionally, while data from published studies are highly valuable and raw data (e.g., fastq files) are nowadays released in public repositories, data processing is time-consuming and again difficult without bioinformatics support. To facilitate NB research, more user-friendly and immediately accessible platforms are needed to explore newly generated as well as existing high throughput data. To make this possible, we developed an interactive data centralization and visualization web application, called CLEaN (Cell Line Explorer of Neuroblastoma data; https://ccgg.ugent.be/shiny/clean/). By focusing on the regulation of the DNA Damage Response, a therapeutic target of major interest in neuroblastoma, we demonstrate how CLEaN can be used to gain novel mechanistic insights and identify putative drug targets in neuroblastoma.

92

Epigenetic and transcriptomic signatures to predict response to immune checkpoint inhibition in NSCLC

María Gallardo[2], Artur Carreras-Soldevilla[2], Irene Alonso-Álvarez[2], Martín E Lázaro-Quintela[2], Carme García-Benito[2], Aitor Rodríguez-Casanova[1], Ángel Díaz-Lagares[1], Joaquin Casal[2], Mónica Martínez-Fernández[2]

[1] Cancer Epigenomics, ONCOMET, IDIS, CHUS, CIBERONC, Roche-Chus Joint Unit, Santiago de Compostela, Spain
[2] Health Research Institute Galicia Sur (IIS Galicia Sur), SERGAS-UVIGO, Hospital Álvaro Cunqueiro, Spain

Presenter: María Gallardo

Immunotherapy based on Immune Checkpoint Inhibitors (ICI) changed the management of advanced Non-Small Cell Lung Cancer (NSCLC). However, due to low response rates, a deeper understanding of ICI response mechanisms is still an urgent need. A close relation between DNA methylation and the immunogenic status of the tumor microenvironment (TME) has been recently proposed for several solid tumors, yet needs to be further explored.
In this work we evaluated the epigenetic and transcriptomic status of NSCLC tumors before immunotherapy to disclose the possible role of DNA methylation in ICI response, and to identify immune-related signatures that predict ICI outcome. We included 24 advanced NSCLC patients treated with anti-PD1 at first line. Genome-wide DNA methylation was quantified with the MethylationEPICv2 array (930k) and RNA-seq was performed with a NovaSeq6000 system. Data was annotated to the GRCh38 genome assembly and R/Bioconductor was used for downstream bioinformatic analyses. Differential methylation and expression analysis were performed between patients showing radiological response at 6 months (n=10) and non-responder patients (n=14). Gene set enrichment, as well as GO/KEGG over representation analyses were performed for differential methylation and expression signatures.
We identified a DNA methylation signature composed of 197 CpG sites that predicts a 6-months response to anti-PD1 therapy. Also, we reported a hypermethylated profile associated with longer progression-free survival (PFS) ($p<0.001$). Functional consequences of differential DNA methylation were confirmed by RNA-seq, identifying dysregulated immune pathways significantly enriched for both methylation and transcriptomic signatures. Pathways associated to ICI response related to T-cell and toll-like receptors signaling pathways, MHC-II complex, and interleukins. We also found aberrant methylation of the PI3K-Akt signaling pathway, which plays an essential role in the TME, regulating immune checkpoints and the sensitivity to ICI.
Overall, our results confirm that DNA methylation plays a role in the response to ICI in advanced NSCLC. These immune-related methylation profiles can predict response to anti-PD1 therapy and correlate with longer PFS, representing a source of predictive biomarkers for ICI outcome.

93

Evolutionary trajectories and mechanisms of colorectal peritoneal metastasis formation

Miaomiao Gao[1]

[1] University of Manchester, United Kingdom

Presenter: Miaomiao Gao

Colorectal cancer (CRC) is the 4th most common cancer in the UK, accounting for 11.4% of all new cancer cases and 10% of all cancer deaths. There are three common sites of metastases in colorectal cancer, of which the peritoneum is the most deadly. Patients with peritoneal metastasis have significantly shorter median overall survival (OS) (16.3 months) than those with liver (19.1 months) or lung (24.6 months) metastases. There is an urgent need to improve the treatment and prognosis of this disease.

Cancer metastasis is a process in which cancer cells undergo genetic and epigenetic changes due to the instability of their own genomes and are dynamically reshaped through evolution and adaptation to environmental changes. Analysis of the characteristics and genomic factors driving metastasis is key to understanding the occurrence and progression of colorectal peritoneal metastasis(CRPM).

Thanks to the rapid development of sequencing technology and computational methods, the analysis of multi-cancer samples using whole genome sequencing(WGS) has become more rapid and affordable. By analyzing genomic changes in primary and metastatic samples, we can understand the evolutionary trajectories of CRPM, further providing some useful insight for clinical translation.

94

Global analysis uncovers frequent loss of tumor suppressor proteins by stop-loss mutations

Avantika Ghosh[3], Jagriti Pal[2], Sonam Dhamija[1], Marisa Riester[5], Sven Diederichs[4]

1 CSIR-Institute of Genomics and Integrative Biology New Delhi, India
2 University Hospital Freiburg, Germany
3 University Hospital Freiburg and German Cancer Consortium (DKTK), Germany
4 University of Freiburg, Germany
5 University of Freiburg and German Cancer Consortium (DKTK), Germany

Presenter: Sven Diederichs

Nonstop mutations, a.k.a. stop-loss mutations, convert a stop codon into a sense codon resulting in translation into the subsequent 3' untranslated region (3'UTR). Thereby, the protein is extended at its C-terminus until translation terminates at the next in-frame stop codon (extension) or into the polyA-tail (readthrough). In a recent pan-cancer analysis, we found and curated nonstop mutations in 3412 tumors from 62 entities, generating a comprehensive database of nonstop mutations in cancer at http://NonStopDB.dkfz.de (Nat Cell Biol 2020).
For the tumor suppressor SMAD4, we identified six different nonstop extension mutations, present in pancreatic and colon cancers, the cancer entities known to be driven by the loss-of-function of SMAD4. All six mutations extended the C-terminus of the protein by 40 amino acids which led to a complete loss of SMAD4 protein expression without affecting mRNA levels. We identified a hydrophobic signal sequence of ten amino acids within the C-terminal extension necessary and sufficient to direct rapid degradation of SMAD4 via the ubiquitin-proteasome system (Nat Cell Biol 2020).
Now, we prepared a library of all somatic nonstop extensions in human cancer according to NonstopDB and tested their impact on protein expression in a high throughput screening format. Of the 2335 individual extension sequences, 966 (41.4%) caused a robust decrease in protein expression compared to only 0.1% of the wildtype sequences which served as negative controls. On average, effective extensions showed distinct patterns of amino acid enrichments as well as stronger hydrophobicity in line with our discovery of the hydrophobic degron in SMAD4. Notably, our screen identified two distinct effects of nonstop mutations on tumor suppressor genes. One, in which all nonstop mutations on the tumor suppressor gene showed the same effect, indicating that the effect was dependent on the extension sequence after the mutated stop codon. Second, extensions in tumor suppressor genes that showed distinct patterns for each mutation indicating that the impact was dependent on the individual amino acid change at the stop codon position.
In summary, we provide evidence that nonstop mutations can be functionally important in cancer and systematically characterize their impact on human proteins altered in cancer.

95

Clonal expansion and adaptive DNA methylation-based epigenetic plasticity underlie resistance to oncogenic pathway inhibition in pancreatic cancer

Laura Godfrey[11], Jan Forster[4], Sven-Thorsten Liffers[11], Christopher Schröder[4], Johannes Köster[4], Leonie Henschel[5], Kerstin U. Ludwig[5], Marija Trajkovic-Arsic[11], Diana Behrens[3], Aldo Scarpa[8], Rita T. Lawlor[1], Kathrin E. Witzke[6], Barbara Sitek[6], Steven Johnsen[7], Sven Rahmann[2], Bernhard Horsthemke[9], Michael Zeschnigk[9], Jens Siveke[10]

1 ARC-Net Cancer Research Centre, University and Hospital Trust of Verona, Italy
2 Center for Bioinformatics and Department of Computer Science, Saarland University, Germany
3 EPO Experimental Pharmacology and Oncology GmbH, Germany
4 Genome Informatics, Institute of Human Genetics, University Duisburg-Essen, Germany
5 Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Germany
6 Medizinisches Proteom-Center/Zentrum für Protein-Diagnostik, Ruhr-Universität Bochum, Germany
7 Robert Bosch Center for Tumor Diseases, Germany
8 University and Hospital Trust of Verona and ARC-Net Cancer Research Centre, Italy
9 University Hospital Essen, University of Duisburg-Essen, Germany
10 West German Cancer Center, Germany
11 West German Cancer Center, University Hospital Essen, Germany

Presenter: Laura Godfrey

Pancreatic ductal adenocarcinoma (PDAC) is an aggressive tumor entity marked by extraordinary resistance to conventional therapies including chemotherapy and radiation, as well as to essentially all targeted therapies evaluated so far. More than 90% of PDAC cases harbor an activating KRAS mutation. As the most common KRAS variants in PDAC remain undruggable so far, it seems promising to inhibit a downstream target e.g. MEK1/2, but up to now (pre-)clinical evaluation of MEK inhibitors (MEKi) failed due to emergence of resistance. Here, we study resistance mechanisms to MEKi focusing on genome-wide DNA methylation dynamics. We have characterized ten short-term passaged cell lines from genetic mouse models that recapitulate resistance phenomena such as tumor subtype switching and cell death escape. We use a combination of whole-genome bisulfite sequencing (WGBS), whole-genome sequencing (WGS), RNA sequencing (RNA-Seq) and mass spectrometry to integrate genetic aberrations, global methylation patterns and transcriptional dynamics with functional analyses to reveal differences between parental cells and their resistant counterparts.
We find the resistant cell populations that arise upon MEKi exposure to be based on clonal expansions of individual cells, but not as a consequence of known resistance-conferring mutations. Rather, resistant cells showed adaptive DNA hypermethylation of 209 and hypomethylation of 8 genomic sites, most of which overlap with regulatory elements known to be active in murine PDAC cells. Both DNA methylation changes and MEKi resistance were transient and reversible upon drug withdrawal. The effector caspase CASP3 is one of the 114 genes for which transcriptional downregulation inversely correlated with the methylation status of associated DNA regions. CASP3 inactivation in resistant cells led to attenuation of drug induced apoptosis which could be reversed by DNA methyltransferase inhibition with remarkable sensitivity exclusively in the resistant cells. Thus, the concept of MEKi resistance as a result of the expansion of a single cell clone with epigenetic plasticity sheds light on (epi-)genetic and phenotypic patterns during evolvement of treatment resistance in a tumor with high adaptive capability and provides potential for reversion through epigenetic targeting.

96

Landscape of differentiation induced oncogenesis regulated by pseudogenes: a study of gastrointestinal tract

Pravallika Govada[1], Rajasekaran Ramalingam[1]

[1] Vellore Institute of Technology, India

Presenter: Pravallika Govada

Non-coding RNAs, especially pseudogenes have been well-studied for their role in regulating organismal development as well as cellular differentiation. However, the role of pseudogenes in promoting cellular differentiation in the context of cancer initiation, progression and metastasis remains elusive. Tumours of the gastro-intestinal (GI) tract often indicate presence of metaplastic epithelium with diverse cell signatures as a precursor to tumour development or as a part of tumorigenesis. Hence, our current study explores the tumour landscape of the GI tract using convolutional neural networks (CNN) to identify pseudogenes potentially regulating the events of differentiation that ultimately promotes oncogenic transformation. Our analysis of oesophagal, gastric, colon and rectal carcinomas indicate distinct stage-wise expression pattern of pseudogenes as well as their direct and indirect interacting partners. In fact, tissue-specific homeostatic expression of pseudogenes prompts the effective use of de-regulated pseudogene expression identified from our analysis as a valid diagnostic marker. Previously, we defined three unique metrics that characterized the extent of differentiation independent of histopathological tumour grading. Along with gene regulatory network and topology analysis, these metrics helped us identify the relationship between the extent of differntiation and pseudogene expression pattern. Of note, we identified combinatorial de-regulation of pseudogenes as a potential driving factor of increased differentiation and metaplastic epithelium, especially for patient stratified Stage II oesophageal carcinoma. Further analysis of transcription factors helped us reveal SOX2 as a key regulatory factor upstream as well as potentially downstream of pseudogenes, thus, driving differentiation induced oncogenic transformation along with pseudogenes and other transcription factors including FEV, PRRX1 and TFAP2A. Finally, CNN-based survival analysis of the GI tract tumours reveals distinct pseudogenes as valuable prognostic markers, similar to our previous observations where, ARSDP1 and GYG2P1 were promoters of poor prognosis across a distinct landscape of pseudogene expression and differentiation within oesophageal carcinoma.

97

Whole-genome CRISPR-Cas9 screens reveal genetic dependencies in NRAS-mutant melanoma cell lines

Andrea Gu[3], Tet-Woo Lee[2], Aziza Khan[3], Francis Hunter[1], Dean Singleton[2], Stephen Jamieson[2]

1 Janssen Research & Development, United States of America
2 University of Auckland, New Zealand
3 University of Auckland, Auckland Cancer Society Research Centre, New Zealand

Presenter: Andrea Gu

Melanoma is the most lethal form of skin cancer, with over 300,000 new cases worldwide every year. Specifically, melanoma with NRAS mutations are of particular clinical concern due to their association with a poor prognosis and lack of specific treatment options. Therefore, there is a pressing need for novel approaches to address the treatment of NRAS-mutant melanoma. A current approach to identify novel drug targets is based on the genetic concept of induced essentiality, where functional interactions that occur in response to oncogene addiction create a dependency on another gene.
To identify genetic dependencies in NRAS-mutant melanoma, whole-genome CRISPR-Cas9 knockout screens were conducted in 6 NRAS-mutant and 7 NRAS-wildtype melanoma cell lines that were established from New Zealand melanoma (NZM) patients. The NZM cell lines were stably transduced with the whole-genome Brunello lentiviral single guide (sg) RNA library and screened for up to 35 days. BAGEL (Bayesian Analysis of Gene Essentiality) analyses of the NZM whole-genome knockout screens, alongside CRISPR-Cas9 screening data using the Avana sgRNA library from an additional 28 melanoma cell lines, available on the Cancer Cell Line Encyclopaedia (CCLE) database, revealed 45 prospective candidates that exhibit greater detrimental effects on the fitness of NRAS-mutant cell lines compared to the NRAS-wildtype lines. These genes are being further validated as essential genes for NRAS-mutant melanoma cells through custom sgRNA library knockout screens and in vitro individual gene knockout studies. In particular, we demonstrate that knockout of SHOC2, a scaffold protein essential for activation of the MAPK signalling pathway, results in the prevention of ERK phosphorylation and a more substantial reduction in cell proliferation in NRAS-mutant NZM cell lines when compared to NRAS-wildtype lines. Currently, our research involves further investigating SHOC2 as a therapeutic target for the treatment of NRAS-mutant melanoma.

98

High clonal diversity and spatial genetic admixture in early prostate cancer and surrounding normal tissue

Luuk Harbers[1], Luuk Harbers[1], Nicola Crosetto[2], Michele Simonetti[1], Ning Zhang[1]

1 Karolinska Institute, Sweden
2 Karolinska Institute / Human Technopole, Sweden

Presenter: Luuk Harbers

Somatic copy number alterations (SCNAs) are pervasive in advanced human cancers, but their prevalence and spatial distribution in early-stage, localized tumors and their surrounding normal tissues are poorly characterized. Here, we performed multi-region, single-cell DNA sequencing to characterize the SCNA landscape across multiple tumor-rich and normal tissue regions (~125 mm3 tissue cubes) obtained from prostatectomy performed in two patients with localized prostate cancer. We identified two distinct populations of cells with abnormal karyotypes, one marked by sparse deletions or amplifications ('pseudo-diploid' cells) and the second characterized by genome-wide copy number changes reminiscent of 'monster' cells previously described in colorectal cancer. Pseudo-diploid cells formed numerous small-sized subclones ranging from highly spatially localized to broadly spread subclones mainly featuring (sub-)chromosomal arm deletions. In contrast, monster cells harbored whole-chromosome gains and losses and were mostly singular events detected throughout the prostate, including normal tissue regions. Targeted deep sequencing of cancer-associated genes revealed a more confined pattern of mutations overlapping with tumor-rich regions, although we also detected mutations in regions deemed normal based on morphological assessment and bulk RNA-seq. Highly localized pseudo-diploid subclones were confined within tumor-rich regions and typically carried deletions involving chromosome (chr) 6 and 13, resulting in simultaneous loss of multiple tumor-suppressor genes, including FOXO1 and FOXO3 encoding two transcription factors belonging to the Forkhead family previously implicated in prostate carcinogenesis. Tumor-rich regions also contained mutations in genes frequently mutated in prostate cancer, including FOXA1, LRP1B, SPOP, and SPTA1. Our study reveals that SCNAs are widespread in both normal and tumor regions across the prostate gland of patients with localized prostate cancer and suggests that a subset of pseudo-diploid cells harboring chromosomal deletions that result in the loss of specific tumor-suppressor genes drive tumorigenesis in the aging prostate.

Luuk Harbers is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

99
Cancelled

100

Integrating large-scale genomic and transcriptomic datasets to identify colorectal cancer susceptibility genes with therapeutic potential

Emma Hazelwood[8], Pik Fang Kho[6], Xuemin Wang[6], D. Timothy Bishop[11], Andrew T Chan[3], Stephen B Gruber[2], Jochen Hampe[7], Loic Le Marchand[10], Michael O Woods[5], Rish K Pai[4], Stephanie Schmit[9], Jane C Figueiredo[1], Emma E Vincent[8], Dylan M Glubb[6], Tracy O'Mara[6]

1 Cedars-Sinai Medical Center, United States of America
2 City of Hope National Medical Center, United States of America
3 Harvard Medical School, United States of America
4 Mayo Clinic, United States of America
5 Memorial University of Newfoundland, Canada
6 QIMR Berghofer Medical Research Institute, Australia
7 Technische Universität Dresden (TU Dresden), Germany
8 University of Bristol, United Kingdom
9 University of California, United States of America
10 University of Hawaii Cancer Center, United States of America
11 University of Leeds, United Kingdom

Presenter: Emma Hazelwood

Colorectal cancer (CRC) is the third most common cancer worldwide. However, the biological mechanisms underlying disease development are unclear, meaning the development of therapeutic interventions remains challenging. Previous transcriptome-wide association studies (TWAS) have revealed several genes potentially implicated in CRC development. However, these analyses have not stratified by anatomical subsite or sex, nor applied robust causal inference analysis. We conducted a comprehensive analysis to identify genes with robust evidence for a causal role in CRC risk. We conducted a thorough investigation using two multi-tissue TWAS methods (S-MultiXcan and Joint Tissue Imputation) to maximise power to detect associations. We integrated data from a large CRC genome-wide association study (52,775 cases; 45,940 controls) and gene expression data from six relevant tissues from GTEx v8 (comprising lymphocytes and colon and adipose tissues). Analyses were repeated stratifying by sex and anatomical subsite. Identified genes were further assessed for a causal role in CRC development using Mendelian randomization and colocalization analyses. We also assessed the causal relationship between colorectal cancer risk and genes encoding 1,263 actionable proteins that are targeted by drugs which are approved or in clinical development (the “druggable genome”). Prioritised genes were then evaluated for therapeutic potential by exploring drug repurposing opportunities using the Open Target and Connectivity Map platforms. Our analysis identified 29 genes with robust evidence for a causal role in CRC development across all anatomical subsites and sexes, two of which encode proteins that are targetable with available therapeutics. These genes highlight novel biological pathways which may have a role in CRC development. In summary, our study provides valuable insights into the molecular underpinnings of sex- and subsite-specific CRC and identifies promising candidate genes with a causal role in disease aetiology. Furthermore, our findings offer potential avenues for therapeutic intervention, including the repurposing of existing drugs and the development of novel treatments.

101

Chromatin accessibility and pioneer factor FOXA1 shape glucocorticoid receptor action in prostate cancer

Laura Helminen[1], Jasmin Huttunen[1], Niina Aaltonen[1], Einari Niskanen[1], Ville Paakinaho[1], Jorma Palvimo[1]

[1] University of Eastern Finland, Finland

Presenter: Laura Helminen

Treatment of prostate cancer relies predominantly on the inhibition of androgen receptor (AR) signaling. Despite the initial effectiveness of the antiandrogen therapies, the cancer often develops resistance to the AR blockade. One mechanism of the resistance is glucocorticoid receptor (GR)-mediated replacement of AR function. Nevertheless, the mechanistic ways and means how the GR-mediated antiandrogen resistance occurs have remained elusive. Here, we have discovered several crucial features of GR action in prostate cancer cells through genome-wide techniques. We detected that the replacement of AR by GR in enzalutamide-exposed prostate cancer cells occurs almost exclusively at pre-accessible chromatin sites displaying FOXA1 occupancy. Counterintuitively to the classical pioneer factor model, silencing of FOXA1 potentiated the chromatin binding and transcriptional activity of GR. This was attributed to FOXA1-mediated repression of the NR3C1 (gene encoding GR) expression via the corepressor TLE3. Moreover, the small-molecule inhibition of coactivator p300's enzymatic activity efficiently restricted GR-mediated gene regulation and cell proliferation. Overall, we identified chromatin pre-accessibility and FOXA1-mediated repression as important regulators of GR action in prostate cancer, pointing out new avenues to oppose steroid receptor-mediated antiandrogen resistance.

102

Optimal hematopoietic stem cell dynamics suppress the selection of leukemic mutations

Thomas Höfer[1], Verena Körber[1], Hans-Reimer Rodewald[1], Csilla Kongsaysak-Lengyel[1]

[1] German Cancer Research Center (DKFZ), Germany

Presenter: Thomas Höfer

The continuous production of new cells in renewing tissues poses a risk for malignant transformation. Here we develop a theory for the selection of oncogenic drivers in hematopoiesis, a prototypical stem-progenitor cell hierarchy. The theory predicts that intermediate stem cell activity minimizes cancer risk. Higher activity promotes stem cell transformation, whereas paucity of stem cell output allows drivers to accumulate in, and eventually transform, progenitors. We confirm key predictions of the theory in a mouse model of T-cell acute lymphocytic leukemia that recapitulates the genetics of the human disease. Abrogating hematopoietic stem cell output to the lymphoid lineage causes multiple premalignant T-cell progenitor clones to emerge. In a subset of these clones, the rapid tunnelling of Notch1 mutations, and sometimes further drivers, leads to the explosive onset of the disease. In sum, our findings suggest: (i) that leukemic driver mutations are abundant and not rate-limiting for the development of acute leukemia; rather (ii) selection presents the rate-limiting step; and (iii) selection is shaped by the dynamics of the stem-progenitor cell hierarchy.

103

Identification of genes diving metastasis in the intestinal type of gastric cancer

Maximilian Hohenleitner[4], Therese Seidlitz[3], Moustafa Abohawya[1], Jovan Mircetic[2], Daniel Stange[3]

1 Dresden University of Technology, Germany
2 German Cancer Consortium (DKTK), Germany
3 University Hospital Carl Gustav Carus and Dresden University of Technology, Germany
4 University Hospital Carl Gustav Carus Dresden, Germany

Presenter: Maximilian Hohenleitner

Gastric cancer (GC) ranks the fifth most common and the third leading cause of cancer-related deaths worldwide. Due to the lack of early clinical symptoms, 75% of patients show an advanced cancer stage at time of diagnosis. The widely used Lauren classification divided GC based on the morphological appearance into the intestinal, diffuse and mixed type. The intestinal subtype has a well-differentiated architecture with a tendency to develop liver and lung metastasis. In contrast, the diffuse subtype shows an undifferentiated mass of tumor cells preferentially metastasizing to the peritoneum.
Genetically engineered mouse models have led to a tremendous increase in knowledge regarding tumor initiation, progression and metastasis. With the help of our stomach-specific inducible Anxa10-CreER^{T2} mouse line we could model the intestinal type of GC. The intestinal subtype is characterized by a high percentage of TP53 mutations, RTK-RAS amplifications and TGFβ pathway alterations. Interestingly, the murine tumor model with TGFβ pathway alterations (besides Tp53 mutations and RTK-RAS pathway alterations) consistently and rapidly developed liver and lung metastases. Tumor models with only Tp53 mutations and RTK-RAS alterations metastasized inconsistently and only after a long time. Using the sequencing data from the primary tumors of the different intestinal mouse models, we now want to answer the question which TGFβ signaling pathway-independent alterations may be relevant for metastasis.
For this purpose, a CRISPR/Cas9 screen will be performed in gastric organoids. Potentially relevant genes found in the sequencing data will be depending on the mutation knocked out using a CRISPR/Cas9 knock out screen or up-regulated in expression using a catalytically deactivated Cas9. Organoids will be further orthotopically transplanted, tumor growth via in vivo imaging observed and metastases for gRNA appearance sequenced. This allows the identification of metastasis relevant genes.
Such CRISPR screens allow the identification of the relationships between genotype and phenotype and has been established in recent years as a powerful laboratory tool for the study of cancer genomics. Finally, genetic alterations leading to tumor progression and metastasis can be found which are potential targets for future GC therapy.

104

The potential of whole-body donors in studying mutant clones in normal tissues

Sofie Hoogstoel[1], Tom Luijts[1], Isabelle Hoorens[2], Anne Vral[1], Wouter Willaert[1], Jimmy Van den Eynden[1]

[1] Ghent University, Belgium
[2] Ghent University Hospital, Belgium

Presenter: Sofie Hoogstoel

Somatic driver mutations accumulate in the genome of ageing cells, occasionally resulting in a cellular fitness advantage, positive selection and clonal expansion. The progressive accumulation of these driver mutations can lead to malignant tumor formation. It was recently demonstrated that these mutations also drive small clonal alterations in histologically normal tissues from many organs, likely underlying the foundation of human carcinogenesis.
An important bottleneck for these studies is the availability of normal human tissue. Additionally, commonly used tissue resources have limitations. Surgical remnant tissues, residual tissues from transplant organ donors, peritumoral samples or clinical biopsies are restricted to a specific location and/or influenced by a patient's clinical history. Tissues derived from autopsy subjects could be an alternative, but the lack of previous patient consent poses severe limitations on the availability of the required clinical background information (e.g., iatrogenic mutagen exposure).
To overcome these limitations, we propose using post-mortem tissues derived from whole-body donors, which are routinely used in human anatomy units for educational purposes. These donors give previous consent to use their organs and clinical information for educational and/or research purposes. The high number of available donors and previous screening of clinical records by the responsible physician allow including or excluding donors for well-defined studies on normal tissue mutant clonality.
We have demonstrated that mutant clones can be reliably identified in mutagen-exposed (e.g., cigarette smoke, UV light) and unexposed oral and skin epithelial tissues, sampled from whole-body donors. Positive selection signals were detected in the expected driver genes (e.g., TP53, NOTCH1-3, FAT1) and the identified mutational signatures corresponded to the exposed mutagens. Most donors could be sampled within a 72h post-mortem interval (PMI). Variation of the PMI between 0 and 72h did not affect epithelial isolations and DNA remained stable, both on fresh and snap-frozen samples.
Our results demonstrate that post-mortem tissues sampled from whole body donors could provide a nearly unlimited source of normal tissue to study mutant clones.

105

Molecular adaption to radio-chemotherapy in esophageal adenocarcinoma and influence of BRCA2 function

Sascha Hoppe[3], Oscar Velazquez Camacho[3], Ali Yazbeck[3], Jan Grossbach[3], Alexander Quaas[2], Stefan Müller[1], Reinhard Büttner[2], Axel Hillmer[3]

1 CECAD Proteomics Facility, Germany
2 University Hospital Cologne, Germany
3 University of Cologne, Germany

Presenter: Axel Hillmer

Introduction: Esophageal carcinoma is the sixth most common cause of cancer-related death world-wide. The most prevalent subtype in western countries is esophageal adenocarcinoma (EAC) with highly increasing rates in the past decades. Survival time of less than a year and a mortality rate of >85 % outline the poor response to current treatment regimens. We aim to predict common mechanisms that evolve after treatment and lead to resistance against neoadjuvant chemo-radiotherapy.
Methods: EAC cell lines were genetically engineered to obtain a stable BRCA2 deficiency, which increases the genomic instability and mimics the BRCAness that has been reported for ~20 % of EAC patients. Parental and BRCA2-deficient cell lines were used to generate mouse xenografts, which were treated with a combination of radiation and chemotherapy (RCT). Dissected tumors were analyzed via 3'-mRNA sequencing, whole- and phospho-proteomics.
Results: A partial deletion, resulting in BRCA2 knockdown (BRCA2kd), leads to enhanced migration of cells in vitro. In vivo, BRCA2kd slows down tumor growth. RNA and proteomic analysis revealed that BRCA2kd leaves a stronger expression phenotype than RCT with enrichment of pathways involved in migration, angiogenesis, and metabolic processes. RCT treated tumors are enriched for pathways involved in resistance mechanisms that partly were not reported before, i.e. epithelial differentiation towards a basal cell type, a process similar to keratinization. Furthermore, phospho-proteomics revealed an increased phosphorylation of Erk1/2 targets and hippo pathway related proteins after treatment. Interestingly, RCT treatment also led to alterations in the tumor microenvironment of the human xenografts, emphasizing the importance of cancer associated fibroblasts and immune cells on tumor progression and resistance. On the genomic level, we observed an altered somatic copy number landscape after BRCA2kd and characteristic changes after RCT. Overall, our model system provides insight into the development of EAC resistance to RCT and the role of BRCA2 in this process.

106

Clonal hematopoiesis through dysfunction of the Fanconi anemia DNA repair pathway

Ashley Kamimae-Lanning[3], Franziska Hörsch[1], Thomas Höfer[1], KJ Patel[2], Jill Brown[2]

1 German Cancer Research Center (DKFZ), Germany
2 University of Oxford, United Kingdom
3 WIMM Oxford, United Kingdom

Presenter: Franziska Hörsch

Patients with Fanconi anemia suffer from aberrant hematopoiesis, including bone marrow failure and leukemia. Mice with a parallel deletion of the aldehyde detoxifying enzyme Adh5 and the DNA repair enzyme Fancd2 in hematopoietic tissues recapitulate this phenotype. Here we analyzed somatic variants in hematopoietic cells of these mice. Based on mathematical theory of mutation accumulation and clonal dynamics, we found granulocytes but not B cells descended from a single hematopoietic stem or progenitor cell in 30 week old mice. Mutational signature analysis of clonally expanded samples revealed great similarity to the human aging signature. Interestingly, we do not detect known oncogenic drivers in the selected clones and subclones, suggesting that the selective advantage is caused by a mechanism distinct from oncogenesis. Our findings indicate that Fanconi anemia patients may suffer from very early clonal hematopoiesis, similar to many aged individuals.

107

RNA-sequencing reveals a subset of uterine leiomyomas with FGFR1 and FGFR2 mutations

Vilja Jokinen[1], Niko Välimäki[1], Auli Karhu[1], Isa Ahlgren[1], Oskari Heikinheimo[2], Ralf Bützow[2], Annukka Pasanen[2], Lauri Aaltonen[1]

1 University of Helsinki, Finland
2 University of Helsinki and Helsinki University Hospital, Finland

Presenter: Vilja Jokinen

Uterine leiomyomas (ULs), benign tumors arising from the myometrium, are the most common tumors in women. Approximately one in four patients experience symptoms, such as pelvic pain, excessive uterine bleeding, and infertility. In addition, ULs cause a massive economic burden, and all curative treatment options are currently invasive.
ULs can be divided into five subtypes based on mutually exclusive genetic alterations in MED12, HMGA2, FH, SRCAP complex genes, and defects in Cullin 3-RING E3 ligase neddylation. However, the primary genetic driver is unknown for 10% of ULs. Furthermore, secondary drivers and their effects on myomagenesis are not very well known.
Our sample collection consists of 2898 ULs and 936 corresponding normal myometrium samples from 936 patients. Sanger sequencing, immunohistochemistry, and qPCR are used to screen well-known driver alterations. We have produced RNA-sequencing (RNAseq) data of 410 ULs and 151 myometrium samples, in part for mutation characterization. We are also utilising SNP array, whole-genome sequencing, whole-exome sequencing, and Nanopore long-read sequencing data to analyse copy number alterations, structural variants, and methylation.
Preliminary results of variant calling from RNAseq data revealed three hotspot mutations in FGFR1 and four mutations in FGFR2 in ULs, both well-known cancer genes. The same mutation hotspot in FGFR1 has been reported in many cancer types, including breast cancer and gliomas. We are currently screening mutations in FGFR1/2 hotspots in our UL collection to estimate the frequency of these mutations in ULs. To understand how FGFR1/2 mutations drive tumorigenesis and dysregulate cellular pathways in ULs, pathway analysis will be utilised for RNAseq data.
Our methodology combining multi-omics data is a valuable approach to construct a more comprehensive view on mutations, molecular subtypes, and dysregulated pathways in ULs, which is necessary for developing non-invasive and personalized treatments as well as diagnostics. More research is needed to understand whether ULs with FGFR1/2 mutations have increased potential to progress to malignancy. Research on ULs expands our knowledge on benign tumors in general, facilitating our understanding on what distinguishes malignancies from benign tumors.

108

Exploring changes in protein levels in plasma before and after cystectomy in patients with bladder cancer

Randi Istrup Juul[1], Asbjørn Kjær[1], Nanna Kristjánsdóttir[1], Iver Nordentoft[1], Karin Birkenkamp-Demtröder[1], Mads Agerbæk[1], Jørgen Bjerggaard Jensen[1], Lars Dyrskjøt[1], Nicolai Birkbak[1]

[1] Aarhus University Hospital, Denmark

Presenter: Randi Istrup Juul

The immune system state and composition affects the outcome for patients with cancer. Here we investigated if immune-related proteins were affected by treatment in muscle-invasive bladder cancer (MIBC).
For 22 patients with MIBC we compared the plasma protein level in pre-treatment samples to paired samples taken 1) three weeks after neoadjuvant chemotherapy and radical cystectomy, and 2) one year after treatment or at relapse. We used the Olink Explore proteomics assay to measure the level of 1536 proteins at each timepoint. Significantly different proteins were identified with a Wilcoxon Signed Rank test, and were further investigated using gene set enrichment analysis (GSEA). We used logistic regression and cox proportional hazard regression to identify proteins that affect recurrence/recurrence free survival.
We found 389 proteins with significant differences between pre- and three weeks post-treatment. Of these, 293 were present at higher levels after treatment, and GSEA showed that multiple pathways related to the immune system were enriched among these proteins. Only two proteins were significantly different between pre-treatment and at relapse, and none were significantly different between pre-treatment and one year post-treatment. Using the ratio of pre- to post-treatment protein levels for both post-treatment timepoints, we found a total of 215 proteins that potentially affect recurrence using logistic regression. Of these 66 were immune-related proteins. Using these ratios in a cox proportional hazard regression model, we identified 147 proteins that may affect recurrence free survival, of which 46 were immune-related proteins.
We find enrichments of immune-related pathways among proteins that are present at higher levels three weeks after treatment compared to baseline. Furthermore, approximately one third of the proteins that potentially affect recurrence are immune-related proteins. Our results underline that treatment (neoadjuvant chemotherapy and surgery) may affect immune system composition and activation.

109

Detection, characterization, and prevention of MMEJ deletions

Aditee Kadam[2], Shay Shilo[2], Hadas Naor[2], Alexander Wainstein[1]

[1] Weizmann Institute of India, Israel
[2] Weizmann Institute of Science, Israel

Presenter: Aditee Kadam

Background: Detecting medium-sized deletions in short-reads is highly challenging due to reference biases and mapping issues. We developed an algorithm that enables de novo detection of medium-sized deletions: DelRead. The algorithm focuses on a specific type of deterministic deletion with a well-defined genetic mechanism - Micro-Homology mediated End Joining deletions (MMEJ-del). Using prior knowledge of the MMEJ mechanism, our algorithm compiles a complete set of potential deletions in the exome. Subsequently, it maps these deletions to sequencing reads, thereby reducing reliance on mapping differences to a reference genome.
Aims: To explore the somatic and germline MMEJ-del landscape using Del-Read, and provide insights into preventing somatic deletions through genome editing.
Methods: The Del-Read algorithm was applied to two datasets - Beat AML and TCGA-breast - which comprised of tumor-control paired exomes (N=359 and 225, respectively). A subset of these mutations underwent deep targeted sequencing in a cohort of 672 healthy individuals.
Results: The Del-Read algorithm identified novel germline (N=486), somatic (N=20), and reported (N=82) MMEJ-del in the datasets. A subset of these mutations (N=37) was validated with comparable population frequencies using targeted sequencing of healthy individuals (N=672) in ethnicity-matched controls.
The magnitude of novel MMEJ-del discovered allowed us to associate them with genomic features of replication stress such as G-quadruplexes and minisatellites. Interestingly, we also observed a new class of MMEJ-del characterized by mismatches in the homologies, although not all mismatches were equally tolerated. Further, we demonstrated that a specific single-base substitution can restrict the occurrence of pre-leukemic MMEJ-del in the ASXL1 gene.
Summary: Our findings highlight Del-Read's potential to uncover previously undetected deletions and provide insights into preventing somatic deletions through genome editing.

110

Elucidating the clonal evolutionary dynamics of hypermutated tumors using single-cell whole-genome sequencing

Maria Kalyva[2], Freddy Gibson[3], Francesc Muyas[1], Carolin Sauer[1], Jose Espejo Valle-Inclan[1], Gabriele Picco[3], Carmen Herranz-Ors[3], Shriram Bhosle[3], Syd Barthorpe[3], Hayley Francies[3], Mathew Garnett[3], Isidro Cortes Ciriano[1]

1 EMBL-EBI, United Kingdom
2 EMBL-EBI and University of Cambridge, United Kingdom
3 Wellcome Sanger Institute, United Kingdom

Presenter: Maria Kalyva

Immune checkpoint blockade (ICB) has revolutionized the cancer treatment landscape. Yet, durable responses are only observed in a subset of patients. The highest response rates to ICB have been observed in mismatch repair-deficient (MMRD) tumors, which are caused by the inactivation of the MMR pathway. Tumors with MMRD are characterized by a high burden of neoantigens. However, only ~50% of MMRD tumors respond to ICB. Previous clinical and preclinical studies of MMRD tumors indicate that the clonality rather than the burden of mutations determines the variable responses to ICB observed in the clinic. Given that the clonality of mutations is determined by the evolutionary dynamics and patterns of intra-tumour heterogeneity, it is paramount to elucidate the evolutionary trajectories of MMRD tumors to understand the variable immunogenicity of MMRD tumors. Prior genomics analysis of MMRD tumor evolution used bulk whole-genome sequencing (WGS), which is limited to infer the clonal structure and dynamics of tumors due to the impossibility to assign mutations to subclones and limited sensitivity for mutation detection.
Here, to study the evolution of MMRD tumors, we performed single-cell WGS on ~200 single cells from eight patient-derived colorectal MMRD organoids. Mutational burden, though variable across models, was comparable across single cells from the same tumor. Phylogenetic analysis uncovered multiple clones per tumor, with most pathogenic mutations mapping to the trunk of the phylogenetic tree, demonstrating that these mutations are shared across all cancer cells. Notably, few copy number aberrations were observed, indicating that tumor heterogeneity is driven by point mutations and indels. Furthermore, we developed a phylodynamic analysis software, called ClonalSim, to analyze the growth dynamics of tumors using the phylogenetic trees constructed for each tumor, and to infer the fitness of cancer cells and the timing of clonal expansions.
In sum, we present a novel experimental and computational framework that has allowed us to study the patterns of genetic heterogeneity and clonal dynamics underpinning the evolution of MMRD colorectal cancer at single-cell resolution. This framework is broadly applicable to study genetic heterogeneity and clonal evolution in other malignancies and cancer types.

111

Inferring copy number signatures from distinct evolutionary events

Tom Kaufmann[1], Roland Schwarz[2]

[1] Max Delbrück Center for Molecular Medicine, Germany
[2] University Hospital and University of Cologne, Germany

Presenter: Tom Kaufmann

Somatic copy number alterations include large-scale events, such as chromosome arm-level gains and losses as well as focal amplifications and deletions and play a key role in the evolutionary processes that shape cancer genomes. In the case of small-scale events such as point mutations and indels, there exists a list of established mutational signatures that can be linked to distinct exogenous or endogenous exposures such as tobacco use. Despite previous efforts, accurate and meaningful copy number signatures are still elusive. The biggest obstacle in creating copy number signatures is that due to their cascading nature, traditional segment-based representations of copy numbers do not reveal individual evolutionary events.

Here we introduce a new method for deriving copy number signatures that explicitly models evolutionary copy number events. We derive these events using a minimum evolution framework based on our whole-genome-doubling-aware phylogenetic copy number model MEDICC2 (Kaufmann 2022, Genome Biology). In contrast to traditional segment-based approaches, we can infer whether a segment was gained or lost and recover the exact event length even when multiple events are overlapping. In the case where the precise chain of events cannot be resolved by our evolutionary model, we leverage orthogonal data from structural variant calls as well as the relative timing of copy-number changes as inferred from the presence of single nucleotide variants. We benchmark our approach on an independent simulation of mutational processes and demonstrate that we can recover many events that are missed in segment-based approaches, including whole-chromosome gains as well as gains and losses relative to whole-genome duplications. We verify these findings on real world data from 2,778 tumors from the Pan Cancer Analysis of Whole Genomes.

112

Evolutionary analysis reveals independent clonal populations in single Barrett's esophagus biopsies

Phoebe He[1], Mariya Kazachkova[1], Ludmil Alexandrov[1]

[1] University of California San Diego, United States of America

Presenter: Mariya Kazachkova

Barrett's esophagus (BE) is a non-obligate precursor of esophageal adenocarcinoma (EAC), often measuring multiple centimeters in length and characterized by columnar epithelium replacing normal squamous epithelium in the distal esophagus. Early studies in BE suggested that all cells within a patient's Barrett's segment are progenies of a single founder cell. However, in recent years this previously accepted model of BE as a monoclonal disease has been retired in favor of a multiclonal model with new evidence suggesting that a patient's Barrett's segment can be composed of multiple independent clones with independent founder cells. Nonetheless, single BE samples are still routinely assumed to be monoclonal, leading to independent sets of genomic aberrations erroneously being analyzed as a single set in the case of multiclonal samples. Here, we utilize a previously generated whole-genome sequenced BE dataset and repurpose an established pipeline for detecting subclonal populations to identify BE samples that are comprised of multiple independent clones. Additionally, we present a novel approach for validating multiclonality from sequencing data by phasing pair-end sequencing reads using germline variants. We demonstrate that 20/80 patients have at least one multiclonal sample and that the independent clones contained in these single biopsies can have drastically different genomic profiles. Further, we speculate that these multiclonal samples capture a snapshot of clonal competition and evolution within BE. These findings highlight the importance of accounting for multiclonality when studying BE and provide the first systematic analysis of multiclonal BE samples.

113

Signatures of conditional selection in cancer whole-genome sequences identify chemotherapy resistance genes

Ahmed Khalil[1], Daniel Naro[1], Elizaveta Besedina[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Ahmed Khalil

Tumors often show an initial response to chemotherapy, but then develop resistance, leading to relapse and poor prognosis. Therefore, it is essential to reveal the genetic mechanisms of tumor progression and chemotherapy resistance, and to suggest genes to be targeted to overcome resistance. We hypothesized that a genomic comparison of mutations in pre-treated versus treatment-naïve tumors would serve to identify genes that confer resistance. A challenge in such an analysis is that chemotherapy exposure alters mutation burdens and signatures, confounding association studies and complicating identifying causal selected mutations. Therefore, establishing a mutation rate baseline that accounts for these shifts in neutral mutagenesis during cancer evolution is imperative.

Given the increasing availability of whole-genome sequencing (WGS) data, we developed DiffInvex (Differential Introns Versus Exons), based on a Poisson regression model regularized by a weakly-informative prior, appropriate for sparse mutation data, to quantify conditional selection in cancer. DiffInvex determines the differential excess of point mutations over a baseline by comparing two conditions or time points. Importantly, DiffInvex rigorously estimates the baseline local mutation rate in each condition based on intronic and intergenic mutations; this approach is similar to or more accurate than the covariate-based state-of-the-art tools for identifying somatic selection. Confounding by trinucleotide and pentanucleotide composition and DNA sequence mappability is stringently accounted for via a locus sampling approach.

We applied DiffInvex to >8000 cancer whole-genomes from >30 cancer types from 8 studies, containing both naïve tumors and tumors pre-treated by various drugs, grouped by mechanism-of-action. In a pan-cancer analysis, DiffInVEx confirmed known drug-associated driver genes as being under conditional selection, such as ESR1 with antiestrogen therapy and EGFR with EGFR inhibitor drugs. Additionally, DiffInvex identified >10 putative treatment-associated genes for different classes of therapeutics. Notably, controlling for confounding between conjoint therapies was crucial for eliminating spurious associations.

This work was supported by Horizon2020 project DECIDER (965193).

114

Enhanced single cell DNA methylation analysis using combinatorial sequencing

Sanika Khare[1], Hosu Sin[1], Ashley Woodfin[1], Eric Pu[1], Phuong Dang[1], Dominic Skinner[1], Aimee Beck[1], Margaret Nakamoto[1], Beth Walczak[1]

[1] Scale Biosciences, United States of America

Presenter: Katharina Danielski

The epigenetic landscape of the human brain undergoes many changes during natural development and in malignant transformation (1,2). In recent years, DNA methylation-based epigenetic modifications have been widely studied using techniques like bisulfite sequencing and enzymatic methyl sequencing (EM-seq). However, these methods have been restricted to bulk cell populations. Although advances in single-cell analysis have revolutionized our understanding of cellular heterogeneity, these approaches are still costly, low throughput, and laborious.

To address these challenges, ScaleBio has pioneered combinatorial indexing technology, enabling a significant increase in cell throughput. This novel method utilizes the cell itself as a compartment to perform 2-3 rounds of sequential barcoding in a plate-based workflow, eliminating the need for complex and expensive instrumentation. This robust technology has now been successfully adapted to assess DNA methylation at the single-cell level.

In this study, we used ScaleBio's single-cell methylation kit to investigate DNA methylation patterns in both fetal and adult brain cells, along with cancer cells containing widespread DNA methylation changes, such as isocitrate dehydrogenase (IDH) mutant glioma cells. The IDH gene family, comprising IDH1, IDH2, and IDH3, encodes enzymes involved in the tricarboxylic acid (TCA) cycle. Mutations in IDH1/2 genes are frequently found in tumors (more than 80% of high-grade gliomas) and exhibit a distinct hyper-methylation pattern (4,5).

We achieved high cell recovery and robust cytosine coverage throughout our analysis of single-cell methylomes isolated from brain tissue. Using this data we generated a ranked list of the top hypo- and hypermethylated genomic regions and identified cell type-specific clusters seen in different developmental or pathological states by looking at Differentially Methylated Regions (DMR). We then compared these to known bulk methylation profiles, uncovering unique single-cell methylation profiles that may be obscured in bulk or pseudo-bulk analysis.

Taken together these data show that the ScaleBio single-cell methylation workflow offers increased sensitivity, specificity, and accuracy in identifying DNA methylation sites, and provides insights into cellular heterogeneity.

115

An atlas of CMS in colorectal cancers at spatial single-cell resolution

Batuhan Kisakol[3], Anna Matveeva[3], Sanghee Cho[1], Mohammadreza Azimi[3], Andreas Lindner[3], Elizabeth McDonough[1], Simon McDade[2], Daniel Longley[2], Fiona Ginty[1], Jochen Prehn[3]

1 GE Research, United States of America
2 Queen's University Belfast, United Kingdom
3 Royal College of Surgeons in Ireland, Ireland

Presenter: Batuhan Kisakol

Identification of the consensus molecular subtypes (CMS) opened an immense potential for understanding the prognosis, tumour biology and intertumour heterogeneity in colorectal cancer (CRC). Molecular subtyping in CRC often relies on bulk transcriptomics data. However, it was shown in single-cell studies that CRC tumours may be composed of tumour cells displaying different CMS traits. Therefore, we investigated the intratumour heterogeneity of CRC tumours using spatially resolved single-cell datasets and compared different CMSs as classified by bulk transcriptomics. We analysed >2 million cells in 581 tumour microarrays (TMA) from 219 patients. TMAs were stained and imaged with Cell DIVE technology using selected 56 protein markers, ranging from the markers of the extrinsic and intrinsic apoptosis pathways to the metabolic and immune cell markers. Through multiplexed immunofluorescence imaging analysis, we revealed an atlas illustrating the cell states, spatial heterogeneity, cellular neighbourhoods, cellular network and protein profile of different CMS tumours. Our findings suggest that, at cellular composition, CMS1 tumours have more CD3+, CD8+ and PD1+ cells. Moreover, immune cells are seen in the epithelial layer more frequently in CMS1 than in the other subtypes. KI67, BCLXL and SR2B levels were found to be higher in epithelial cells and CDX2 to be lower in CMS1. We observed higher spatial autocorrelation (Moran's I) scores of many proteins in CMS2 which suggests expression of the proteins tend to be more clustered in CMS2 tumours. Epithelial cells in the CMS3 tumours were clustered together and closely surrounded by stromal cells. Finally, in CMS4 tumours, regulatory T cells and helper T-cells were found to be far away from cancer cells and the overall epithelial cell content was lower. In conclusion, combining cutting-edge multiplexed imaging technology with novel spatial single-cell analysis, our study provides the first atlas of CRC tumours with regard to their molecular subtypes at single-cell protein resolution and demonstrates a new spatial aspect of tumour structures.

116

T cell receptor repertoire diversity and blood T cell fraction is associated with outcome in bladder cancer

Asbjørn Kjær[1], Iver Nordentoft[1], Nanna Kristjánsdóttir[1], Randi Istrup Juul[1], Karin Birkenkamp-Demtröder[1], Jørgen Bjerggaard Jensen[1], Mads Agerbæk[1], Nicolai Birkbak[1], Lars Dyrskjøt[1]

[1] Aarhus University Hospital, Denmark

Presenter: Asbjørn Kjær

As cancer develops, cancer neoantigens are recognized by T cells from the adaptive immune system. The clinical importance of this interaction is highlighted by the success of immunotherapies that primarily work by activating cancer-specific T cells. Expansion of these cancer-specific T cell clones is an early response to malignancy. We hypothesized that characterization of the T cell receptor (TCR) repertoire based on blood samples could be used to evaluate immune competency and serve as a biomarker for cancer progression.
To explore this potential, we characterized the peripheral TCR repertoire of 119 patients with muscle invasive bladder cancer (MIBC), using ultra-deep amplicon-based sequencing of the TCR beta chain. We found that lower TCR diversity was associated with worse outcome in bladder cancer, particularly when combined with a lower amount of circulating T cells relative to other blood cells. Analysis of the individual TCR sequences revealed that expanded TCR clones disproportionately targeted Epstein-Barr and cytomegalovirus antigens. We detected DNA from latent viral infection by analysis of plasma circulating tumor DNA (ctDNA) whole genome sequencing data, and found lower TCR diversity to be strongly associated with latent cytomegalovirus infection. Using longitudinal TCR sampling, we found that highly expanded viral-targeting clones were persistent through time. However, sequencing of tumor TCRs revealed that these expanded clones were rarely found in the tumor, indicating that they are not tumor-specific. Taken together, this indicates that T cell clones found expanded in circulation are often not targeting cancer neoantigens, but may reflect previous or persistent viral infections.
We suggest that high TCR diversity and high T cell fraction represent improved immune competency. This may be associated with reduced risk of progression due to a greater ability to control cancer growth and limit systemic cancer cell dissemination. These findings underline that immune health is an important factor during malignant disease, even when patients are not treated with immunotherapies, and thus improving immune health could be a potential goal of cancer treatment and prevention in the future.

117

Regulation of Brachyury gene expression in breast cancer

Akechai Kongprajug[1], Claire Wells[1], Fiona Wardle[1]

[1] King's College London, United Kingdom

Presenter: Akechai Kongprajug

Brachyury, the founding member of the T-box family of transcription factors, plays a critical role in early embryonic development; however, it is typically absent from most normal adult tissues. Recent studies have revealed that Brachyuryexpression is reactivated in various cancers, where it actively contributes to cancer progression and metastasis through the activation of the epithelial-to-mesenchymal transition (EMT) process. In breast cancer, elevated levels of Brachyury have been correlated with the late stages of the disease and are associated with poor prognosis. Consequently, the objective of this research is to identify the cis-regulatory regions (CREs) and signalling pathways responsible for controlling Brachyury re-expression in late-stage breast cancer, with the ultimate aim of developing novel therapeutic strategies. To identify potential CREs, several publicly available ChIP-seq, ATAC-seq, and GRO-seq datasets were analyzed, comparing a Brachyury-expressing breast cancer cell line (MDA-MB-436) with a non-Brachyury-expressing cell line (MDA-MB-231 cells). However, none of the candidate CREs displayed any discernible enhancer activity when assessed by reporter assay. It is important to note that the regulatory mechanisms of certain genes within their native chromatin context may require multiple CREs to effectively enhance gene transcription. Additionally, the native chromatin conformation itself is crucial for gene activation. Therefore, to identify the Brachyury enhancer within an endogenous chromatin context, an enhancer inhibition (SID4x-dCas9-KRAB) and enhancer activation (dCas9-P300) will be employed. Furthermore, it has been established that the FGF signaling pathways play a significant role in activating Brachyury expression during embryogenesis. To investigate this pathway in breast cancer, inhibitors of the FGF receptor (Pemigatinib) and downstream effector ERK (Ravoxertinib) were employed to inhibit FGF signaling in the Brachyury-expressing cell line. These inhibitors effectively suppressed FGF signalling and also Brachyury expression, providing further evidence of the involvement of FGF signaling in Brachyury regulation.

118

Genome segmentation by DNA damage repair dynamics

Hanna Kranas[2], Ferran Muiños[2], Abel Gonzalez-Perez[2], Joan Frigola Rissech[2], Núria López-Bigas[1]

1 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
2 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Hanna Kranas

Somatic mutagenesis results from the interplay between DNA damage and repair, both influenced by chromatin features. Previous attempts to characterize somatic mutagenesis have studied the efficiency of repair related to specific chromatin features. With the availability of DNA damage and repair genomic maps at different time points after exposure to a mutagen (i.e. UV light (Adar et al. 2016, Hu et al. 2017)), we are now in a position to characterize this interplay along the entire human genome.
Thus, instead of segmenting the genome by chromatin features, we propose an alternative approach to partition the genome according to, solely, the dynamics of the DNA repair. Inspired by the idea of partitioning the genome into chromatin states (chromHMM (Ernst & Kellis 2012, Ernst & Kellis 2017, Kundaje et al. 2015)), we devised a similar strategy to infer “repair states”, representing genomic chunks with differing repair activity. We take an autoregressive modeling perspective, exploiting Bayesian nonparametric Hidden Markov Models which allow for inference of a set of the repair states from time-resolved damage and repair data.
We have thus obtained UV-damage genomic repair states encoding differences in overall repair speed, acceleration of repair at different times, and dynamics of repair of different photoproducts. With this repair-based, data-driven partitioning, we could undertake systematic surveys to uncover links to various genomic features. The states range from slow-and-late repair in repressed, packed, inaccessible regions to fast-and-early repair in expressed, transcribed, active regions of the genome. This spectrum of UV-damage repair states reflects a continuum of chromatin feature changes.
This work sets a framework for a comprehensive analysis of the interplay of DNA damage and repair with the chromatin and understanding how those processes lead to the deposition of mutations. While we focused on UV damage, Nucleotide Excision Repair, and UV-induced mutations, in the future, this approach will allow us to explore other mutagens, their subsequent repair dynamics, and interplays with genomic features.

119

Identification of neoantigen candidates from splicing in human tumor cell lines

Franziska Lang[2], Patrick Sorn[2], David Gomez-Zepeda[1], Yannic Chen[1], Alina Henrich[2], Anne Kölsch[2], Barbara Schrörs[2], David Weber[2], Martin Löwer[2], Ugur Sahin[3], Stefan Tenzer[4], Jonas Ibn-Salem[2]

[1] Helmholtz Institute for Translational Oncology Mainz (HI-TRON Mainz), Germany
[2] TRON, Germany
[3] TRON; University Medical Center of the Johannes-Gutenberg University Mainz; BioNTech SE, Germany
[4] University Medical Center of the Johannes-Gutenberg University Mainz; Helmholtz Institute for Translational Oncology Mainz (HI-TRON Mainz), Germany

Presenter: Franziska Lang

Personalized cancer immunotherapies are effective treatments against cancer by targeting individual and tumor-specific neoantigens to trigger de novo T-cell responses. Detecting tumor-specific mRNA splicing events that could encode neoantigens for personalized cancer immunotherapies is challenging because splice junctions identified in tumor transcriptomes can also appear in healthy tissues. However, somatic mutations can disrupt canonical splicing, leading to individual tumor-specific targets. Previously, we developed the R-package splice2neo to integrate the predicted splice effects from somatic mutations with splice junctions detected in tumor RNA-seq and established a stringent detection rule to predict tumor-specific splice junctions in melanoma patient cohorts. To experimentally confirm such predictions and characterize the resulting effects on the proteome and the human leukocyte antigen (HLA)- ligandome, we here applied splice2neo and the detection rule to five tumor cell lines. We identified multiple tumor-specific targets of which a large subset encode promising neoantigen candidates. Notably, we found further experimental evidence for the tumor-specificity of the predicted target junctions by showing that those targets are absent in the RNA-seq of matched normal cell lines. While there was evidence from RNA-seq that the targets are expressed in the tumor cell lines on the transcript level, it still remained unclear if they are translated into proteins and are presented in the form of epitopes by the HLA molecules on the cell surface of tumor cells. To address this question, we also performed mass spectrometry-based immunopeptidomics in these tumor cell lines to directly identifiy HLA-bound peptides from splicing and other mutation types. Taken together, we predict a valuable set of neoantigen candidates from splicing in human tumor cell lines to better understand how these candidates expand the target repertoire for personalized cancer immunotherapies.

120

Longitudinal tracking of acute myeloid leukemia clonal evolution after allogeneic hematopoietic cell transplantation

Nathan Lee[3], Elizabeth Krakow[2], Isaac Jenkins[1], Bryce Fukuda[5], Olga Sala-Torra[1], Lan Beppu[1], Jerry Radich[2], Brenda M. Sandmaier[2], Cecilia Yeung[2], Ivana Bozic[4]

1 Fred Hutchinson Cancer Center, United States of America
2 Fred Hutchinson Cancer Center & University of Washington, United States of America
3 University of Washington, United States of America
4 University of Washington & Fred Hutchinson Cancer Center, United States of America
5 University of Washington Tacoma, United States of America

Presenter: Nathan Lee

Acute myeloid leukemia (AML) is an aggressive neoplasm affecting hematopoietic tissues, with a poor overall survival rate. Allogeneic hematopoietic cell transplantation (HCT) provides a potentially curative treatment for AML. However, relapse rates for HCT remain high, motivating further study on relapse monitoring and pre-emptive therapy. We present the clonal evolution and patterns of relapse for a cohort of AML patients receiving HCT, with bulk clinical next generation DNA sequencing performed pre-HCT and multiple times post-HCT. From the longitudinal sequencing data, we infer the clonal structure and phylogeny for each patient. We developed an interactive visualization tool for exploring the effect of treatment on biomarkers, clinical events, and inferred cancer evolution. We find a remarkably conserved mode of clonal evolution across the cohort, where clones arising from the donor become dominant post-HCT. We show that clones can be clearly identified as donor or patient-derived from sequencing patient samples alone, without the need for sequencing donor samples. We confirmed these predictions by sequencing the donor samples. Additionally, we investigate which features of clonal dynamics can be used to better predict relapse. This work presents a tool for comprehensive characterization of the clonal evolution of AML before and after HCT. By combining the clonal structure and evolution inferred from longitudinal sequencing with biomarkers and clinical events, we present a way to monitor for AML relapse that is likely more robust than conventional methods that rely on flow cytometry data and/or single mutations, because leukemic immunophenotype may change over time and single mutations may be lost at the time of relapse.

121

Studying differences in mutability between parental sets of chromosomes

Maxime Lefèbvre[1], Maxime Tarabichi[1]

[1] Université Libre de Bruxelles, Belgium

Presenter: Maxime Lefèbvre

Each cell contains two copies of the human genome, one inherited from the mother and one from the father. These parental genomes exhibit variations in genomic structure and differ at millions of positions in the DNA sequence along homologous chromosomes. Various factors, such as DNA methylation, histone modifications, gene expression, and chromatin structure, known as mutability factors, have been found to exhibit allele-specific characteristics in some genomic regions. As a result, it is probable to observe variations in somatic mutations at specific genomic locations between alleles. So far, no one has quantified this effect in cancer genomes.

To fill this knowledge gap, the goal of this project is to quantify the potential differences in mutation rates of parental allele sets separately, both locally and globally, with an original methodology.

To accomplish this, the mutation rate is computed utilizing the whole-genome sequencing data of more than 2,600 cancers available in the Pan-Cancer Analysis of Whole Genomes (PCAWG) database. Considering the absence of trio data and the lack of genome phasing precision, two probabilistic assignments have been developed by exploiting SNP information. The first identifies mixed-ethnicity patients of PCAWG cohort and the second accurately attributes the parental origin of each copy of the genome without having access to the genotype of the parents. Finally, a method has been developed to assign each somatic mutations to a parental copy with the goal to quantify the local and global mutation rate per parental copy and compare them.

By conducting this research, we hope to contribute to a better understanding of the mutability patterns and shedding light on the influence of parental genomes on mutational processes.

122

Single-cell dissection of CK-AMLs characterizes targetable disease-driving leukemic stem cell clones

Aino-Maija Leppä[4], Karen Grimes[1], Tobias Boch[7], Hyobin Jeong[3], Alexander Waclawiczek[4], Darja Karpova[8], Florian Grünschläger[4], Anna Jauch[6], Vera Thiel[4], Alwin Krämer[2], Ashley Sanders[5], Jan Korbel[1], Andreas Trumpp[4]

1 EMBL Heidelberg, Germany
2 German Cancer Research Center (DKFZ), Germany
3 Hanyang University, Republic of Korea
4 Heidelberg Institute for Stem Cell Technology and Experimental Medicine; German Cancer Research Center (DKFZ), Germany
5 MDC/BIH/Charité, Germany
6 University Hospital Heidelberg, Germany
7 University Hospital Mannheim, Germany
8 Washington University School of Medicine in St. Louis, United States of America

Presenter: Aino-Maija Leppä

Chromosomal instability is a major driver of intra-tumor heterogeneity (ITH) that promotes tumor progression and treatment resistance. While single-cell technologies have emerged as potent tools to study ITH, the complexity and dynamics of structural variants (SVs) that shape cell phenotypes and determine disease progression as well as clinical outcome remain largely unexplored. To address this, we devised an integrated single cell multi-omics framework, combining SV discovery and nucleosome occupancy profiling (scNOVA), with concurrent immunophenotypic and transcriptomic profiling (CITE-seq), to characterize the ITH of Acute Myeloid Leukemia with complex karyotype (CK-AML). Applying the scNOVA-CITE framework to primary cells from four CK-AML patients and the respective patient-derived xenografts unveiled the complex spectrum of SV landscapes in CK-AML and revealed intricate subclonal growth patterns. We identified linear and circular breakage-fusion-bridge cycles along with chromothripsis, and uncovered subclones exhibiting extensive cell-to-cell karyotype heterogeneity, mediating continued karyotypic remodeling. Genetically distinctive and patient specific subclones exhibited distinct nucleosome occupancy, transcriptomic, immunophenotypic and functional features. Transplantation of primary CK-AML cells into immunocompromised mice demonstrated predominantly monoclonal expansion of specific subclones characterized by highest level of genomic instability and stemness-associated phenotypes. This recapitulated the ITH evolution during disease progression after therapy, offering a promising model to study relapse and enabled the identification of drugs targeting the disease-driving leukemic stem cell clones ex vivo. Together, our data provide unique insights into the dynamic genetic and functional complexity of CK-AML and offer clinically relevant avenues to target relapse-driving leukemic stem cell subclones.

123

Longitudinal assessment of NPM1-mutated acute myeloid leukemia patient samples reveals novel insights into targeting therapy-resistant leukemic stem cells

Aino-Maija Leppä[3], Darja Karpova[6], Patrick Stelmach[2], Nesrine Aroua[2], Anne-Kathrin Merbach[4], Sergi Beneyto[1], Elisa Donato[2], Konstanze Döhner[5], Andreas Trumpp[2]

1 Centre for Genomic Regulation (CRG), Spain
2 German Cancer Research Center (DKFZ), Germany
3 Heidelberg Institute for Stem Cell Technology and Experimental Medicine; German Cancer Research Center, Germany
4 University Clinics Heidelberg, Germany
5 University Hospital of Ulm, Germany
6 Washington University School of Medicine in St. Louis, United States of America

Presenter: Aino-Maija Leppä

Despite high rates of initial therapy response, majority of Acute Myeloid Leukemia (AML) patients relapse with a hard to control disease. This relapse is considered to be driven by a population of disease propagating leukemic stem cells (LSCs) that are inherently resistant or acquire secondary resistance during treatment. Due to the low number of therapy resistant leukemic cells present among healthy blood cells at complete remission, the molecular and cellular features of these cells remain poorly understood, hindering the development of preemptive relapse treatment options. Using a unique set of matched samples collected longitudinally at diagnosis, remission, and relapse, we assessed mechanisms driving disease progression in four NPM1-mutated AML patients. We combined clonal single-cell multi-omics and functional assays to dissect and target the leukemia-specific cellular programs of therapy-resistant LSCs. Using CloneTracer, an optimized 10x Genomics-based single-cell analysis pipeline, we were able to measure RNA expression, surface antigen expression, nuclear single nucleotide variants and mitochondrial mutations from the same cell. Longitudinal comparison of NPM1-mutated leukemic cells revealed enrichment for genes encoding for ribosomal proteins and translation regulators in therapy-resistant leukemic cells at remission and relapse, suggesting excessive ribosome biogenesis upon disease progression. Assessment of leukemic cells amongst healthy cells, consistently showed an upregulation of HOX genes in NPM1-mutated cells compared to myeloid NPM1-WT cells, confirming that mutant NPM1 maintains the leukemic state through HOX expression. Intriguingly, NPM1-mutated cells at remission also showed significantly higher expression of the chemokine-encoding gene IL8 compared to the matching NPM1-WT cells. To validate resistance pathways identified at remission, we subsequently established an in vivo patient-derived xenograft model for minimal residual disease using diagnostic samples. Targeting the IL8 receptor CXCR2 resulted in significant reduction in tumor burden, exemplifying the possibilities of our approach. Together, our study showcases how combined longitudinal single-cell multi-omics and functional assays can be used to study and target therapy-resistant LSCs.

124

Investigating the function of the recurrent TBR1 G275C mutation in group 4 medulloblastoma

Bryn Livingston[2], Anders Erickson[2], Ludivine Coudière-Morrison[5], Parthiv Haldipur[1], Liam Hendrikse[4], Jane McGlade[2], Michael Taylor[3]

[1] Seattle Children's Hospital Research Institute, United States of America
[2] The Hospital for Sick Children, Canada
[3] The Hospital for Sick Children, University of Toronto, Canada
[4] University Health Network, Canada
[5] University of Manitoba, Canada

Presenter: Bryn Livingston

Medulloblastoma (MB) is the most common malignant pediatric brain tumour. Standard of care for MB involves radiation to the brain, which has long-term consequences on the health and development of children. Targeted treatments are needed to improve prognoses and avoid the developmental sequelae associated with traditional therapies. Of the four molecularly defined subgroups of MB, group 4 is the least well characterised, despite being the most common. A highly specific single-nucleotide mutation causing a glycine to cysteine amino acid change (G275C) in the DNA-binding domain of T-box Brain 1 (TBR1), an important transcription factor in neuronal development, occurs in a subset of group 4 MB cases, but the function of this mutation remains unknown. RNA sequencing of a large cohort of patient tumour tissues shows upregulation of TBR1 in group 3 and 4 MB, and significantly higher TBR1 expression in G275C tumours than in TBR1 wild-type tumours, suggesting a gain of function. In addition, upregulated genes in G275C tumours map to the most stem-like compartment of the developing human glutamatergic cerebellum. In situ hybridisation shows TBR1 expression in the human cerebellar rhombic lip, the transient germinal zone from which group 4 MB arises. In addition, TurboID proximity labelling experiments on the interactome of CBFA2T2, a core scaffold which binds many group 4 MB driver proteins, indicate that TBR1 interacts with this complex. Together, this suggests an oncogenic role of mutant TBR1 in group 4 MB tumours. Ongoing experiments, including ChIP-seq of TBR1 in patient tissues, and TurboID proximity labelling experiments, will seek to reveal differences in DNA binding sites and protein-protein interactions of wild-type and G275C TBR1. As a transcription factor, TBR1 presents a challenge as a therapeutic target. To utilize the knowledge of this recurrent mutation it will be necessary to uncover its downstream functions and find more effective targets.

125

Multi-dimensional DPClust is a valuable strategy for analysing the clonal/subclonal relationship between multiple WGS LCM samples.

Marian Love[2], Jamie Weaver[1], Rebecca Fitzgerald[3], David Wedge[2]

1 The Christie NHS Foundation Trust, United Kingdom
2 The University of Manchester, United Kingdom
3 University of Cambridge, United Kingdom

Presenter: Marian Love

WGS from laser capture microdissected (LCM) areas of tissue provide a powerful strategy for investigating the mutational architecture of normal and pre-malignant tissues. In many tissue types with clearly defined clonal structures, such as glandular epithelia of the intestinal tract, mutations are expected to be present in all epithelial cells within an LCM sample. However, in some tissues clonal structures are not readily apparent and LCM regions may contain subclones or admixtures of clones. The computational pipelines for processing the WGS data from LCM material are optimised to remove artifacts that occur at low allele frequencies, similar to those seen in subclones. Alternative filtering and mutation clustering strategies may help to optimise identification of subclonal mutations.
To develop our methodology, we have re-examined a case from the very inciteful paper on the mutational landscape of normal endometrial epithelial tissue (Moore et al 2020, PMID: 32350471). The case contained nineteen WGS LCM endometrial glands. Unlike the published method, which uses a variant allele frequency (VAF) hard cut off of 0.1, and up to 0.3 was ambiguous, we applied no VAF filter. Instead, we removed additional known potential artifacts by filtering against the panel-of-normals from BROAD institute's Mutect2, before applying the multi-dimensional implementation of DPClust. This algorithm uses a Dirichlet process, a powerful tool in Bayesian nonparametric statistics, which enables flexible clustering of single-nucleotide variants of calculated allele frequency, into a non-predetermined number of clusters. We found that this revealed a more complex clonal/subclonal architecture.

126

Single-cell lineage mapping of IBD or FAP-associated colorectal tumorigenesis

Zhaolian Lu[1], Kantian Zhou[1], Duo Xiie[1], Kun Wang[3], Xionglei He[2], Zheng Hu[1]

1 Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China
2 Sun Yat-Sen University, China
3 Xiamen University, China

Presenter: Zheng Hu

Pre-cancerous lineage architecture and evolution of colorecal cancer remains elusive, impeding effective prevention of the maligant progression. Here we used a base editor-enabled DNA barcoding system to map high-resoltuion single-cell phylogenies in mouse intestinal neoplasms related to inflammatory bowel disease (IBD) or familial adenomatous polyposis (FAP). Analysis of 249,665 cells from various tissues (normal, inflammatory, polyps, cancers and organoids) revealed that both IBD and FAP-related neoplasms consisted of tens to hundreds independent clonal lineages undergoing parallel expansions. Computational modeling suggested pro-cancer interactions rather than simple spatial collisions among the polyclonal lineages. Moreover, we found chronic inflammation spurred clonal expansions even in healthy colon epithelium. Single-cell RNA sequencing revealed shifting immune landscape during the polyclonal-to-monoclonal transitions. Finally, quantitative timing analysis indicated that intestinal adenomas from Apc germline-mutated mice initiate growth around 42–145 days post-birth, akin to 4.6–15.9 years old (childhood to adolescent age) for human FAP patients. Together, our comprehensive lineage mapping unveils intricate polyclonal interactions in early-stage colorectal tumorigenesis, thus illuminating preventive strategies.

127

Genetic alterations as independent prognostic factors to predict organ-specific metastases of lung cancer

Laura Luhari[2], Ann Valter[1], Olli-Pekka Smolander[2], Anu Planken[1], Kersti Oselin[1]

1 North Estonia Medical Centre, Estonia
2 Tallinn University of Technology, Estonia

Presenter: Laura Luhari

Introduction: Despite improved treatment options the prognosis of LC is still poor with a 5-year overall survival around 15%. Recurrence, especially distant recurrence to other organs is the primary causes of death and metastatic sites may predict prognosis. It remains unclear why some LC patients develop brain metastases, others liver and/or bone metastases. Previous studies have focused on analyzing genomic differences between primary tumor and metastases and no profound conclusions have been made. We aimed to analyze primary tumor genomic features in LC patients and compare them with the clinical information of distant metastases to brain, liver and/or bone.
Methods: We retrospectively selected from the North Estonia Medical Centre (NEMC) Thoracic Tumors' Database previously curatively treated patients who were diagnosed with LC distant recurrence from 2015 to 2017. All the histological specimens (formalin-fixed paraffin-embedded tumor resection or biopsy samples) were sent for whole exome sequencing (WES). Genomic data was analyzed for small genetic alterations, namely single nucleotide polymorphisms (SNPs) and insertion-deletion mutations (INDELs).
Results: From distantly recurred patients 18 had brain, 13 liver and 16 bone metastases. In brain metastases group the two most significantly mutated genes were TMEM64 and ZDHHC9 co-occuring in 33% patients. In liver metastases group the most significant gene mutations were RUNX1T1 and PKMYT1, both present in 43% patients. Bone metastases group exhibited exclusive mutations of PCGF5, ACTN2, ARFGAP1, ELL, KRIT1, KRT32, MPP5, NBN, TRPM7, ZNF621. RUNX1T1 has been described as a promoter of liver metastases in pancreatic cancer. PCGF5 is related to prostate cancer bone metastases. Other found gene alterations' associations with cancer, especially LC and distant metastases have not been previously studied, nor published.
Conclusions: Our study highlights that there are distinct genomic features of cancer cells that may confer the distant metastatic ability to metastasize to specific organ in LC. Our findings are in small number of patients, but important and warrant further analysis to improve patients' management. Understanding the mechanisms of site-specific metastases could potentially improve treatment of patients with LC.

128

Genetic alterations as independent prognostic factors to predict the type of recurrence of lung cancer

Laura Luhari[2], Ann Valter[1], Olli-Pekka Smolander[2], Anu Planken[1], Kersti Oselin[1], Heti Pisarev[3], Birgit Truumees[1]

1 North Estonia Medical Centre, Estonia
2 Tallinn University of Technology, Estonia
3 University of Tartu, Estonia

Presenter: Laura Luhari

Background: Around 60% of lung cancers (LC) are radically treated with surgery or chemoradiation (CRT). Still 33-70% of patients develop recurrence (R) with a median time to relapse 11-16.8 months and approximately 80% of R [locally (LR) or distantly (DR)] occur within the first 2 years. Patients with LR usually do better. Previous studies have focused mainly on the role of clinico-pathological characteristics for the risk of R. The role of molecular mechanisms remains unclear. We aimed to analyze genomic features in LC patients with LR versus (vs) DR to predict the type and risk of R.
Methods: From the North Estonia Medical Centre (NEMC) Thoracic Oncology Database we retrospectively enrolled patients diagnosed R LC from 2015 to 2017, who were previously treated with curative intent. All histological specimens (formalin-fixed paraffin-embedded tumor resection or biopsy samples) were sent for whole exome sequencing (WES). Genomic data was analyzed for small genetic alterations, namely single nucleotide polymorphisms (SNPs) and insertion-deletion mutations (INDELs).
Results: 191 patients were included. 33% of patients had LR and 67% DR, with median R-free survival 15.4 vs 11.2 months (m) (p=0.20) and overall survival after R 12.9 vs 8.5 m (p=0.007), respectively. We identified significant INDEL mutations in 38 and 98 genes and SNP mutations in 63 and 179 genes in DR and LR groups, respectively. DMXL2 mutations were specific only for samples in the DR group. Also, in DR group mutated genes, like STIM1, ITPR3 and RYR3, were significantly enriched in cytosolic Ca2+ related GO terms and pathways, whereas in LR group enrichment of terms related to endoplasmic/sarcoplasmic reticulum Ca2+ was observed. Furthermore, ABCC9 gene mutations caused by INDELs were only prominent in the DR group. Association between those gene alterations and R in LC has not been published previously.
Conclusions: The addition of genomic markers to clinico-pathological characteristics may predict the type of R and prognosis for patients with LC. Our study highlights genetic alterations that warrant further analysis to improve patients' management.

129

Predicting TP53 mutations from spatial transcriptomics data using graph neural networks

Tom Luijts[1], Sofie Hoogstoel[1], Jimmy Van den Eynden[1]

[1] Ghent University, Belgium

Presenter: Tom Luijts

Spatial transcriptomics (ST) is a novel approach to study transcriptomic alterations in a spatial tissue context. While some genomic alterations can already be derived from ST data (e.g., copy number variants), currently available ST technologies do not allow for the direct detection of single nucleotide variants (SNVs), which are the main drivers of cancer.
We developed a machine learning model to predict the spatial organisation of mutant TP53 clones from ST data. The approach employs a graph neural network (GNN) capable of inferring non-synonymous mutations in the TP53 gene, based on downstream gene expression signals and the proximity between surrounding spots on the ST slide. The GNN was trained on in silico simulations of ST experiments using RNA-sequencing and mutation data from 5568 cancer samples obtained from The Cancer Genome Atlas (TCGA).
The model demonstrates a high level of accuracy, boasting an average precision and recall of respectively 83% and 92% on TCGA data. Nevertheless, variable performance was observed across different cancer-type, indicating that the specific characteristics inherent to each cancer type might be challenging to generalize. In the application of the model on ST data obtained from human squamous cell carcinoma (SCC) from a study by Andrew L. Ji (Cell, 2020), the model predicts the tumor-specific keratinocyte regions to carry TP53 mutations. Notably, some SCC tumours that lack TP53 mutations in corresponding bulk whole-exome sequencing data, also have TP53 mutated regions as identified by the model. It remains to be determined whether these predictions are false positive classifications, subclonal mutations or other types of genomic TP53 alterations (e.g., deletions) resulting in similar downstream gene expression patterns.
Our ongoing work demonstrates that GNNs can integrate RNA expression signals and the spatial proximity between different spots on an ST slide to gain new insights into the organization of mutated TP53 clones in a tumor.

130

Genomic and transcriptomic analyses of small intestinal neuroendocrine tumors

Netta Mäkinen[1], Meng Zhou[1], Zhouwei Zhang[1], Yosuke Kasai[2], Grace E Kim[2], Erik Nakakura[2], Chrissie Thirlwell[3], Matthew Meyerson[1]

1 Dana-Farber Cancer Institute, United States of America
2 University of California, San Francisco, United States of America
3 University of Exeter, United Kingdom

Presenter: Netta Mäkinen

Background: Small intestinal neuroendocrine tumor (SI-NET) is one of the major cancer subtypes of the small bowel. Most SI-NETs locate in the terminal ileum with a high incidence of multiple synchronous primary tumors. Exome and genome sequencing studies have reported low somatic mutation rates in SI-NETs. Loss of heterozygosity (LOH) at chr18 is the most frequent genomic event identified (~60% of tumors), and the only known recurrently mutated gene is CDKN1B (~8% of tumors). Recently, we showed that synchronous primary tumors from the same SI-NET patient have distinct somatic mutational profiles, suggesting that these tumors originate independently. A better understanding of the causes of SI-NETs is essential for the optimal treatment of the patients.
Methods: Our sample cohort included 144 fresh-frozen tissue specimens from 23 SI-NET patients, including 85 primary tumors, 21 metastases, and their matched normal ileum and/or whole blood samples. Thirteen SI-NET patients had been diagnosed with multiple synchronous primary tumors. Whole-genome sequencing was used to study the roles of field cancerization and germline variation in SI-NET tumorigenesis, while RNA-sequencing was carried out to analyze differential gene expression between primary tumors, metastases, and normal ileum samples.
Results: We identified chr18 LOH and CDKN1B mutations in multiple tumor samples, as previously described. Otherwise, there was little overlap between the somatic mutational profiles of SI-NETs, whether between patients or among multifocal tumors within patients. No acquired genomic alterations were present in normal ileum samples of more than one SI-NET patient at a time. We also observed recurrent germline variants and will assess if these variants are enriched in our patient cohort. We identified differentially expressed genes between primary tumors and normal ileum samples, including several genes previously associated with neuroendocrine cell biology, and will assess these differences further.
Conclusions: Our results indicate major genomic diversity among uni- and multifocal SI-NETs despite few detectable driver alterations. We see no evidence for mutation-driven field cancerization, and differential gene expression analyses provide interesting new candidates for in vitro studies.

131

The importance of somatic reference materials for precision medicine.

Gras Margaux[1], Violette Turon[1], Mélanie Letexier[3], Jean-Francois Deleuze[2]

[1] CEA, France
[2] Centre National de Recherche en Génomique Humaine, France
[3] CNRGH/CRefIX, France

Presenter: Gras Margaux

The use of next-generation sequencing (NGS) to establish a clinical diagnosis for cancer patients is spreading. Clinical applications of precision oncology require accurate tests that can distinguish tumor specific mutations from technical artefacts. Reference materials (RM) are important to calibrate the global workflow NGS (DNA extraction, library preparation, sequencers, etc.).
Whole Genome Sequencing (WGS) requires both tumour and unaffected tissue, usually blood, to be sequenced. A good RM for WGS should include both type of sample-like standards, the tumour and paired normal sample. The best-known RMs for WGS are coming from the Genome In a Bottle (GiaB) like the CEPH trios (NA12878) essentially use in for germinal diseases workflows. There is no such RM available for somatic WGS workflow.
The CRefIX is the center of reference, innovation and expertise for the French Genomic Medicine Plan 2025, which aim is to implement WGS directly into the healthcare pathway. To reach such a goal, RMs for all applications including oncology need to be investigated to ensure WGS accuracy and specificity for a use in a clinical context.
A pair of cell lines, CRL2324/2325, from the American Type Culture Collection (ATCC) was identified as a good candidate to become a RM for somatic WGS workflow. This pair is derived from a patient with triple negative breast cancer, with both the tumour and paired normal tissue available (blood PBMC) . CRL2324/2325 has been extensively studied by the Sequencing quality control phase II Consortium (SEQC2). The SEQC2, initiated in 2016, with more than 200 scientists from 60 institutes investigated CRL2324/2325 and established reference data sets based on multiple WGS replicates obtained with different technologies including long read sequencing. This cancer cell line is highly heterogeneous, with an aneuploid genome, and is enriched in variety somatic alterations. There are roughly ~40,000 SNVs, ~2,000 small insertions and deletions (indels), CNAs in ~56% of the genome and >256 complex genomic rearrangements.
The CRefIX and the European consortium GenomeMET started a study on reference materials, such as the candidate CRL2324/2325, dedicated to precision medicine and compatible with NGS technology including WGS.

132

Genomic characterisation of mismatch repair deficient colorectal cancer: tumours from Lynch syndrome patients show extreme resemblance with sporadic cases

Samantha Martin[3], Riku Katainen[3], Ari Ristimäki[1], Toni Seppälä[1], Selja Koskensalo[1], Laura Renkonen-Sinisalo[1], Anna Lepistö[1], Jukka-Pekka Mecklin[2], Saija Ahonen[3], Kristiina Rajamäki[3], Kimmo Palin[3], Lauri Aaltonen[3]

1 Helsinki University Central Hospital, University of Helsinki, Finland
2 Jyväskylä Central Hospital, University of Jyväskylä, Finland
3 University of Helsinki, Finland

Presenter: Samantha Martin

Microsatellite unstable colorectal cancer (MSI-CRC) makes up approximately 15% of CRC cases and is defined by a defect in mismatch repair (MMR). The underlying MMR defect can occur sporadically, most commonly through MLH1 promoter methylation, or a germline mutation in one of the MMR genes can be inherited (Lynch syndrome; LS). Despite the differing origins of sporadic and hereditary MSI-CRCs, they have not been extensively compared, partly due to the added complexity from the high numbers of passenger mutations in these tumours that result from having defective MMR.
The aim of this study is to compare sporadic and hereditary MSI-CRCs, all of which are defective in the MMR protein MLH1. Whole genome sequencing (WGS) and RNA-sequencing were performed for 29 sporadic and 14 hereditary MSI-CRCs and, where possible, paired normal colon tissue from the same patients.
Somatic variants were extracted with the standard GATK pipeline, and the GangSTR tool was used to look more closely at the repeat mutation landscape. Overall, sporadic and hereditary MSI-CRCs did not have a significantly different number of SNVs or indels. Sporadic MSI-CRCs did show a trend towards a higher number of insertions, particularly in dinucleotide repeats. LS tumours, however, tended to have more recurrently mutated loci.
Three mutational signatures were identified in the tumours, corresponding to COSMIC single base substitution (SBS) signatures SBS1, SBS15/SBS6 and SBS26/SBS12. These represent a clock-like signature that correlates with the age of the individual and two MSI-related signatures, respectively. There was no significant difference between the exposures to these signatures in sporadic as compared to hereditary MSI-CRCs.
A differential expression analysis identified 136 genes upregulated in sporadic MSI-CRCs compared to 64 in LS CRCs. Sporadic and hereditary MSI-CRCs were not separated in hierarchical clustering indicating similar global differential expression in both tumour groups.
Overall, our results suggest that MLH1-defective MSI-CRCs do not have substantial differences in their molecular landscapes in sporadic and hereditary tumours despite their differing pathways of origin.

133

Expanding the utility of transcriptome analysis in high-risk childhood precision oncology

Chelsea Mayoh[1], Paulette Barahona[1], Angela Lin[1], Lujing Cui[1], Pamela Ajuyah[1], Ann Altekoester[1], Loretta Lau[3], Dong Anh Khuong Quang[2], Vanessa Tyrrell[1], Michelle Haber[1], Marie Wong-Erasmus[1], Paul Ekert[1], Mark Cowley[1]

[1] Children's Cancer Institute, Australia
[2] Royal Children's Hospital, Australia
[3] Sydney Children's Hospital, Australia

Presenter: Chelsea Mayoh

Utilising whole genome (WGS) and transcriptome sequencing (RNA-seq) to identify the molecular features of individual cancers sits at the heart of paediatric cancer precision medicine. Whilst the utility of WGS for mutation detection is well established, most precision medicine programs utilise RNA-seq for fusion detection. In ZERO (Australia's national precision medicine program for children with high-risk cancer), we showed that combining WGS with RNA-seq identified the molecular driver of their cancer in 94% of patients (n=247). We have sought to measure the added utility integrating a comprehensive RNA-seq pipeline provides in the interpretability of aberrations identified in a precision medicine program (n=477).

Our RNA-seq pipeline identified 93% of SVs beyond chimeric gene fusions, including intragenic deletions, duplications, inversions, and insertion events, previously only identifiable through WGS. In addition, RNA-seq resolves the expressed chimeric transcript from DNA chromosome shattering events, multi-hop complex rearrangements and fusion events occurring in homologous and difficult to sequence regions that WGS failed to identify in 33 patients. This information also aides the functional interpretation of SVs through accurate frame calling of potential activating oncogenic SV events, and disruption of key tumour suppressor genes.

We extended our RNA-seq pipeline to identify expressed single nucleotide variants (SNVs) and small insertion and deletions (InDels), identifying 90% of the driver SNVs/InDels called by WGS. Those events not identified are due to either non-sense mediated decay, no gene expression or upstream gene mutations. For 23% of the identified variants, RNA-seq provided functional interpretation by changing the pathogenicity assessment, principally through resolving allele specific expression and confirmation of the transcriptional consequence of splice site mutations.

Our results show that the enhancements to our RNA-seq pipeline has clinical applicability in precision medicine beyond fusion detection, providing functional insights derived from the transcriptional consequences of SVs and SNVs/InDels identified in WGS. We propose that RNA-seq is an essential tool in interpretation of the pathogenicity prediction of molecular drivers of cancer.

134

Polymerase Error Rate sequencing (PER-seq); a novel method for detection of DNA polymerase errors in single molecules

Michael McClellan[2], Marketa Tomkova[2], Skirmantas Kriaucionis[2], Benjamin Schuster-Boeckler[2], Sue Cotterill[1], Gilles Crevel[1], Jakub Tomek[2]

1 St Georges University of London, United Kingdom
2 University of Oxford, United Kingdom

Presenter: Michael McClellan

DNA polymerases are involved in most mutational processes by copying DNA damage or mismatches into mutations, creating mismatches by incorporating incorrect bases, and through involvement in DNA repair. Understanding the mechanisms of mutational processes is important for our understanding of cancer, aging, evolution, and genetic diseases. Despite the importance of DNA polymerases in mutagenesis, the mechanisms whereby individual polymerases contribute to specific mutational processes are largely unknown. A key challenge in this research area has been a lack of reliable bias-free methods to measure error signatures of individual DNA polymerases. Error rates of DNA polymerases have previously been measured using reversion assays, however, these methods introduce considerable biases as only certain mutations produce a measurable phenotype, leading to poor representation of sequence contexts. To address this need, we developed Polymerase Error Rate Sequencing (PER-seq), a novel sequencing-based method that can detect mismatches introduced by DNA polymerases in vitro at single-molecule resolution. PER-seq relies on the addition of Illumina adapters directly to molecules synthesised by the polymerase of interest by PCR, allowing subsequent linear rounds of PCR to use the original molecule as a template. Using only three such linear products, PER-seq is able to measure errors as rare as 1 in a million replicated bases, as validated by its use on serially diluted artificial mutants and on bacterial polymerases with known error signatures. As an in vitro system, PER-seq also allows controlled interrogation of factors involved in replication, such as the effect of template DNA damage or epigenetic modification and perturbation of dNTP pools. We therefore also applied PER-seq on P286R mutant human Pol ε to copy methylated and unmethylated template DNA. Our in vitro Pol ε P286R measurements show a striking similarity with mutational patterns observed in human cancers with deficiency in Pol ε proofreading and mismatch repair and shed novel light on the aetiology of mutational processes in human cancers, as detailed in the associated back-to-back abstract by Tomkova et al. In conclusion, PER-seq is a valuable tool for exploring rare mutations introduced by polymerases in a highly adaptable in vitro setting.

135

Combinatorial gene editing to model genetic interactions in DNA repair deficiencies

Marcel McCullough[1], Maia Munteanu[1], Fran Supek[1]

[1] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Marcel McCullough

Various DNA repair pathways are often deficient in human tumors, resulting in increased mutation rates and altered mutation spectra. These deficiencies have strong implications for predicting inherited cancer risk, modelling tumor evolution and for determining therapy by DNA damaging drugs or by immunotherapy, and should thus be studied better. Cultured human cells represent tractable experimental systems for inducing DNA repair-deficient states by gene editing and observing mutational patterns, which can eludicate the causes of mutational signatures seen in cancer genomes. However, DNA repair systems are characterized by extensive redundancy, and abolishing activity of one gene may be masked to some extent by various backup systems. In other words genetic interactions between DNA repair genes are common, and so gene editing on one gene at a time may not reveal the full diversity of mutational spectra resulting from DNA repair deficiencies. To address this, we employ a Cas12-based system for dual k.o. and apply it to the model human tumor cell line K562, knocking out various pairwise combinations of 8 genes in the mismatch repair (MMR) and 7 genes in the base excision repair (BER) pathway. Validation of edits is performed by Nanopore sequencing, cells grown for ~6 months, and clones expanded from single cells for WGS. Analysis of global SNV and indel mutational patterns in 54 genome sequences reveal that inactivity of individual MMR genes can be distinguished by mutational footprints particulary in the indel spectrum, and also that there are genetic interactions between BER and MMR gene deficiencies. Further applications to cell lines from different tissues (ovary, lung) and to other pathways potentially relevant to mutagenesis (chromatin modifiers, cell cycle control genes) will elucidate roles of genetic interactions between tumor suppressive genes in shaping the mutational landscape of cancer, as well as suggesting tumor vulnerabilities via combinatorial genetic screens. Work funded by ERC StG „HYPER-INSIGHT“ (757700), Horizon2020 project „DECIDER“ (965193) and Spanish MICINN „REPAIRSCAPE“.

136

Circulating cell-free methylated DNA reveals tissue-specific, cellular damage from radiation treatment

Megan McNamara[1], Netanel Loyfer[4], Amber Kiliti[1], Marcel O Schmidt[1], Sapir Shabi-Porat[4], Sidharth Jain[1], Sarah Martinez Roth[1], Arthur McDeed[1], Nesreen Shahrour[1], Elizabeth Ballew[2], Yun-Tien Lin[1], Heng-Hong Li[1], Anne Deslattes Mays[3], Sonali Rudra[2], Keith Unger[2], Anna T Riegel[1], Tommy Kaplan[4], Anton Wellstein[1]

1 Georgetown University Medical Center, United States of America
2 Medstar Georgetown University Hospital, United States of America
3 Science and Technology Consulting, United States of America
4 The Hebrew University of Jerusalem, Israel

Presenter: Megan McNamara

Radiation therapy is an effective cancer treatment, although damage to healthy tissues is common. Here we analyzed cell-free, methylated DNA released from dying cells into the circulation to evaluate radiation-induced cellular damage in different tissues. To map the circulating DNA fragments to human and mouse tissues, we established sequencing-based, cell-type-specific reference DNA methylation atlases. We found that cell-type-specific DNA blocks were mostly hypomethylated and located within signature genes of cellular identity. Cell-free DNA fragments were captured from serum samples by hybridization to CpG-rich DNA panels and mapped to the DNA methylation atlases. In a mouse model, thoracic radiation-induced tissue damage was reflected by dose-dependent increases in lung endothelial and cardiomyocyte methylated DNA in serum. The analysis of serum samples from patients with breast cancer undergoing radiation treatment revealed distinct dose-dependent and tissue-specific epithelial and endothelial responses to radiation across multiple organs. Strikingly, patients treated for right-sided breast cancers also showed increased hepatocyte and liver endothelial DNA in the circulation, indicating the impact on liver tissues. Thus, changes in cell-free methylated DNA can uncover cell-type-specific effects of radiation and provide a readout of the biologically effective radiation dose received by healthy tissues.

137

Revealing aberrant DNA methylation in SINETs through pure-methylome analysis with an EC-like cell line

Carla Castignani[3], Nana Mensah[3], Netta Mäkinen[1], Amy Webster[7], Elizabeth Larose Cadieux[3], Stephen Henderson[5], Javier Herrero[5], Pratik Singh[2], Ramesh Shivdasani[2], Jonas Demeulemeester[8], Stephan Beck[5], Erik Nakakura[6], Matthew Meyerson[1], Chrissie Thirlwell[7], Peter Van Loo[4]

1 Dana-Farber Cancer Institute, United States of America
2 Harvard Medical School, United States of America
3 The Francis Crick Institute, United Kingdom
4 The University of Texas MD Anderson Cancer Center, United States of America
5 University College London, United Kingdom
6 University of California, San Francisco, United States of America
7 University of Exeter, United Kingdom
8 VIB - KU Leuven Center for Cancer Biology, Laboratory of Integrative Cancer Genomics, Belgium

Presenter: Nana Mensah

Small intestinal neuroendocrine tumours (SINETs) are slow-progressing cancers. Patients often present multifocal primary tumours with independent genetic origins. Few recurrent genomic variants have been observed beyond chromosome 18 loss of heterozygosity, indicating the potential for epigenetic influence. The epigenetic progression of SINETs is not well understood owing to the lack of DNA methylation profiles for enterochromaffin (EC) cells, the putative SINET cell of origin.
We introduce the DNA methylome of an EC cell culture derived from human stem cells and profiled with reduced representation bisulfite sequencing (RRBS). Reference-based deconvolution confirms the EC cell line as a more accurate cell of origin proxy than matched normal tissue. We investigate the epigenetic landscape of SINETs by performing RRBS on 27 primary tumours from 11 patients, and 11 metastases from a subset of seven patients. Pure tumour methylomes were derived using CAMDAC to perform in silico purification of tumour methylomes, addressing the confounding effects of tumour purity and aneuploidy.
Our analysis identifies 94,799 significant aberrant CpG methylation events, the majority of which were hypomethylated (65.5%). Importantly, 83.8% fewer CpGs were revealed than using matched normal samples, highlighting the reduction in noise from pure-methylome analysis. DMRCate analysis aggregated these events through 566 DMRs, 25 of which affected promoters and 285 across enhancers. We link these results with aberrant gene expression and transcription factor programmes impacted by enhancer dysregulation. DNA methylation differences were not observed between 18-LOH subtypes. Integrating a cohort of multi-region SINETs as a null model, we observe that intra-tumour DNA methylation heterogeneity in multifocal SINETs supports a model of independent epigenetic evolution.
This work reveals new features of SINET epigenetics and provides a novel analysis approach for future research in larger cohorts.

138

Clinical application of tumour in normal contamination assessment from whole genome sequencing data

Jack Bartram[3], Salvatore Milite[1], Jonathan Mitchell[2], Susan Walker[2], Nadezda Volkova[2], Jane Chalker[5], Luca Vago[4], Alona Sosinsky[2], Giulio Caravagna[6]

1 Centre for Computational Biology, Human Technopole, Italy
2 Genomics England, United Kingdom
3 Great Ormond Street Hospital for Children, United Kingdom
4 IRCCS Hospital San Raffaele, Italy
5 SIHMDS-AG, Great Ormond Street Hospital for Children, United Kingdom
6 University of Trieste, Italy

Presenter: Nadezda Volkova

The unexpected contamination of normal samples with tumour cells reduces variant detection sensitivity, compromising downstream analyses in canonical tumour-normal analyses. Leveraging whole-genome sequencing data available at Genomics England, we developed a tool for normal sample contamination assessment, which we validated in silico and against minimal residual disease testing. From a systematic review of 771 patients with haematological malignancies and sarcomas, we found contamination across a range of cancer clinical indications and DNA sources, with highest prevalence in saliva samples from acute myeloid leukaemia patients, and sorted CD3+ T-cells from myeloproliferative neoplasms. Further exploration revealed 108 hotspot mutations in genes associated with haematological cancers at risk of being subtracted by standard variant calling pipelines. Our work highlights the importance of contamination assessment for accurate somatic variants detection in research and clinical settings, especially with large-scale sequencing projects being utilised to deliver accurate data from which to make clinical decisions for patient care.

139

Factors shaping biallelic mutation frequencies of tumour suppressor genes

Nivedita Mukherjee[1], Sabarinathan Radhakrishnan[1]

[1] National Centre for Biological Sciences, India

Presenter: Nivedita Mukherjee

A tumour is a microcosm of Darwinian evolution where natural selection fosters cells with 'advantageous' or 'driver' genetic alterations. These mainly comprise mutations in oncogenes and tumour suppressor genes (TSGs). Unlike the former, mutations in the latter are usually recessive and need to affect both alleles of the gene (Knudson's Two-hit model of TSG alterations). The universality of this hypothesis has however been refuted as several exceptions to it have become well-known over the years. Yet, a comprehensive understanding of the extent to which these deviations prevail across TSGs and tumour types remains elusive. To find out whether such exceptional occurrences are reflected in the broader context of cohort-wide two-hit mutation frequencies of TSGs, we sought evidence for the roles of evolutionary processes like natural selection and genetic drift in shaping their distribution. We define the two-hit mutation frequency as the probability of deletion of the wild-type allele given the presence of a protein-affecting mutation (point or insertion-deletion, somatic or germline) in the other allele. We compute this for 127 TSGs recurrently mutated across 31 tumour types in the TCGA cohort (comprising ~10,000 samples) by using somatic mutation and allele-specific copy number data. Despite observing a heterogenous distribution of these probabilities as expected given previous reportings, the two-hit hypothesis of TSGs is evident from comparisons with oncogenes and randomly sampled non-cancer-associated genes. We find these probabilities to be dependent on the consequence of the point/indel mutation present, with truncating or loss-of-function mutations co-occurring with deletion of the other allele at significantly higher frequencies than missense mutations, indicating differences in selection pressures for the same. This finding is not confounded by factors such as genome-wide copy number alteration and point mutation levels. We also validate this finding in two additional independent cohorts that add up to 20,000 samples in strength. Together, these results make the effect of selection discernible over that of drift, pointing toward the possibility of being able to identify dominant mutations in TSGs from cohort-wide two-hit frequencies.

Nivedita Mukherjee is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

140

Using machine learning to tackle tumor heterogeneity

Jennifer Neumaier[1], Luisa Bresadola[1], Jonas Ibn-Salem[1], Gudimella Ranganath[1], Pablo Riesgo Ferreiro[1], Barbara Schrörs[1], Ugur Sahin[2]

[1] TRON, Germany
[2] TRON; University Medical Center of the Johannes-Gutenberg University Mainz; BioNTech SE, Germany

Presenter: Jennifer Neumaier

Cancer develops through an evolutionary process, which creates highly diverse primary tumors and metastases. This intra- and intertumor heterogeneity is a very challenging aspect for clinical practice. Selecting therapeutic targets that exist only in a subpopulation of tumor cells can lead to ineffective treatments and in worst cases even to more resistant tumors and relapses. Therefore, the characterization of cancer heterogeneity and the ability to identify targets that are present in all cancer cells of a primary tumor and related metastases are vital to an effective therapy.
Truncal mutations are somatic mutations that appear very early in tumor development and are therefore carried by all cells of a tumor throughout subsequent generations. Currently, multiple tumor samples from a patient are necessary to identify truncal mutations, an approach that is not feasible in a clinical setting. However, the ability to classify mutations as truncal or non-truncal can give key information on a tumor's characteristics and provide more effective targets for individualized therapies. The aim of this project is thus to develop a computational tool to infer this information from a single tumor sample of a patient.
To this purpose, the data generated by the Pan-Cancer Analysis of Whole Genomes (PCAWG) study represent a very useful resource, as they comprise a high number of variants from multiple cancers and patients, labeled according to their evolutionary timing. Additionally, we are analyzing several published whole-exome sequencing datasets of matched primary-metastasis samples to identify somatic mutations and label them as truncal or non-truncal. Mutations from both sources are annotated with several biological and technical features, among which the ratios of non-synonymous to synonymous substitutions showed promising results in an XGBoost model trained on the PCAWG data. We intend to combine this feature with additional ones to improve the performance of the model.
A model able to successfully classify mutations as truncal or non-truncal will bring information on intra- and intertumor heterogeneity into clinical practice, thereby improving therapies for patients.

141

Quantifying the mechanics of transcription coupled repair

Michael Nicholson[1]

[1] University of Edinburgh, United Kingdom

Presenter: Michael Nicholson

DNA base damage is a major source of oncogenic mutations. Such damage can produce strand-phased mutation patterns and multiallelic variation through the process of lesion segregation. Here, by combining mathematical modelling with an experimental system of liver oncogenesis, we exploited these properties to show that DNA damage tolerance is common during mammalian transcription. RNA-polymerases frequently bypass lesions without triggering repair, indicating that small lesions are unlikely to be an efficient barrier to gene expression. The efficiency of transcription coupled repair gradually decays through gene bodies, modelling of this places upper bounds on the efficiency of lesion detection and RNA-polymerase re-initiation following the triggering of nucleotide excision repair.

142

Genomic profiling for predicting ICI response in lung and bladder tumors

Ana Oitaben[4], Jenifer Brea[2], Martin E Lazaro-Quintela[2], Luis Leon-Mateos[3], Carme García-Benito[2], María Gallardo[2], Joaquin Casal[2], Jesus M Paramio[1], Jose M. C. Tubio[5], Mónica Martínez-Fernández[2]

[1] Biomedical Research Institute I+12, University Hospital “12 de Octubre”, Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Molecular Onco, Spain
[2] Health Research Institute Galicia Sur (IIS Galicia Sur). SERGAS-UVIGO. Hospital Álvaro Cunqueiro, Spain
[3] Hospital Clínico Universitario & Instituto de Investigación Sanitaria de Santiago, Spain
[4] Mobile Genomes Lab. CiMUS. University of Santiago de Compostela, Spain
[5] Mobile Genomes. Genomes and Disease. CiMUS. Universidad de Santiago de Compostela., Spain

Presenter: Ana Oitaben

Therapy with Immune Checkpoint Inhibitors (ICI) has become a promising therapeutic strategy for cancer treatment. However, its success depends on predictive biomarkers to select responsive patients.
We analyzed potential predictive genomic biomarkers evaluated at baseline in lung cancer (LC) and bladder cancer (BC). For this purpose, we sequenced tumors from 17 LC and 10 BC patients treated with ICIs at first line by Whole Genome Sequencing to evaluate: a) Tumor Mutational Burden (TMB), b) the presence of SNVs on specific genes related to response, c) somatic Mobile Element Insertions (MEIs) as source of tumor-specific neoantigens and d) Tumor Microenvironment (TME) by studying HLA. We performed variant calling analysis to calculate TMB and identify predictive SNVs using Mutect2; MEIs were evaluated with ERVCaller and TraFiC, and HLA and TCR repertoire were inferred using LiLAC and MiXCR. We have designed the RetroTest, a method to evaluate LINE-1 (L1) activation in tumor biopsies, which can be easily translated to clinical practice. We applied our RetroTest to 31 LC and 26 BC patients to investigate the correlation between L1 activation and response.
Our study found that responder (R) patients in BC presented higher TMB than non-responder (NR), while no differences were found in LC. We identified SNVs specific for each group finding APOBEC mutational signature in BC R patients, one of the modulators of mobile elements. RetroTest detected a significant association between high L1 activity rates and ICI response for BC, and the same positive tendency was observed in LC. In LC, enrichment analyses found pathways related to immune response, CTCF and ERBB affected in R patients, while genes potentially involved in neoantigen presentation were affected in NR patients. In BC, Sema3A and again ERBB pathways appeared mutated among R patients. For the TME study, specific haplotypes were identified, able of discriminating between R and NR.
In conclusion, our data identified the potential of applying different genomic approaches to predict ICI response in BC and LC, affecting both tumour genomes and tumour microenvironment, offering new potential biomarkers which could be included in a predictive signature (conformed by TMB, LINE-1 activation, specific SNVs and HLA haplotypes).

143

SMARCB1 loss activates patient-specific distal MYC enhancers that drive malignant rhabdoid tumor growth

Irene Paassen[4], Lars Custers[4], Hans Teunissen[6], Ning Qing Liu[1], Dilara Ayyildiz[4], Irene Paassen[4], Jiayou He[4], Juliane Buhl[4], Eelco Hoving[4], Peter Zeller[2], Alexander van Oudenaarden[3], Elzo de Wit[6], Jarno Drost[5]

1 Erasmus Medical Center (MC) Cancer Institute, The Netherlands
2 Hubrecht Institute, The Netherlands
3 Hubrecht Institute-KNAW (Royal Netherlands Academy of Arts and Sciences), Oncode Institute, and University Medical Center Utrecht, The Netherlands
4 Princess Máxima Center for Pediatric Oncology, The Netherlands
5 Princess Maxima Center for Pediatric Oncology; Oncode Institute, The Netherlands
6 The Netherlands Cancer Institute, The Netherlands

Presenter: Irene Paassen

Malignant rhabdoid tumor (MRT) is a highly malignant and often lethal childhood cancer. MRTs are genetically only defined by bi-allelic inactivating mutations in SMARCB1, a member of the BRG1/BRM-associated factors (BAF) chromatin remodeling complex. Mutations in BAF complex members are common in human cancer, yet in many cases their contribution to tumorigenesis remains poorly understood. Here, we studied derailed regulatory landscapes as a consequence of SMARCB1 loss in the context of MRTs. Previous work from our lab suggests that loss of SMARCB1 is required, but not sufficient for MRT development. We hypothesize that additional epigenetic events are required for oncogenic transformation. Using a multi-omics approach, we found that reconstitution of SMARCB1 in patient-derived MRT organoid models dramatically reshaped the regulatory landscape of these cells. Chromosome conformation capture experiments subsequently revealed patient-specific looping of distal super enhancer regions with the promoter of the MYC oncogene. Interestingly, the MYC distal enhancers display a high degree of patient specificity, suggesting that intertumoral epigenetic heterogeneity in the MYC enhancer repertoire underlies oncogenic MYC activation. This intertumoral heterogeneity in MYC super enhancer utilization was also present in patient MRT tissues as shown by combined single-cell RNA-seq and ATAC-seq. Pharmacological targeting of the non-canonical BAF complex in MRT organoids phenocopied the effects of SMARCB1 reconstitution. In conclusion, our study reveals that loss of SMARCB1 drives patient-specific epigenetic reprogramming underlying MRT tumorigenesis. We identified a novel epigenetic mechanism of MYC overexpression in MRTs via distal super enhancers, under control of the non-canonical BAF complex. Targeting this complex might give rise to new therapeutic avenues for MRT treatment.

Irene Paassen is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

144

Chromosome instability and aneuploidy tolerance influence single-cell fate upon replication stress

Alba Cardona-Jiménez[2], Veronica Pablo-Fontecha[2], Carolina Parra[2], Sara Lahoz[2], Juan José Lozano[1], Antoni Castells[2], Jordi Camps[2]

1 Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CiberEHD), Spain
2 Institut D'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Spain

Presenter: Veronica Pablo-Fontecha

Chromosome instability (CIN), the ongoing chromosome missegregation phenotype, leads to a progeny with variable aneuploid genomes. To what extent single cells benefit from high levels of CIN and copy number alterations (CNAs) to progress towards a more aggressive phenotype is still under debate. Tetraploid cells resulting from whole-genome doubling (WGD) showing CIN and DNA damage as a consequence of replication stress, have been proposed to cause highly abnormal karyotypes and favor tumorigenesis. Here, we explore the tolerance to CNAs and mitotic errors at the single-cell level in isogenic diploid (2N) and near-tetraploid (4N) clones originated from the DLD1 parental cell line. Single-cell proliferation assays revealed that near-tetraploid clones carrying different CNAs did not show differentiated growth patterns. However, single-cell proliferation of diploid clones was impaired after transient inhibition of DNA replication, while remaining unaltered in near-tetraploid clones. We show that proliferation capacity is not deficient during the first few divisions after single-cell plating, while mid-term colony formation is diminished. Low-pass whole-genome sequencing revealed heterogeneous aneuploid profiles both in 2N and 4N five-day-cultured colonies resulting from a single cell, after exposure to excess of thymidine. In contrast, neither 2N nor 4N clones have shown novel CNAs when maintained long term in culture without exposure to replication inhibitors. This work suggests that replication stress induced by DNA replication inhibition contributes to genetic instability and thus generates heterogeneous karyotypes at the single-cell level. Finally, we postulate that near-tetraploid cells have acquired an elevated tolerance to CIN and aneuploidy that confers a proliferative advantage under stressful conditions.

Veronica Pablo-Fontecha is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

145

Subtyping Xeroderma Pigmentosum in Tanzania through blood whole-exome sequencing

Mavura Daudi[1], Mathieu Parmentier[1], Maxime Tarabichi[1], Ievgenia Pastushenko[1]

[1] Université Libre de Bruxelles, Belgium

Presenter: Mathieu Parmentier

Xeroderma Pigmentosum (XP) is a rare autosomal recessive disease characterized by severe skin, ocular and neurological manifestations. This disease is caused by mutations in nucleotide excision repair genes. There are seven known subgroups XP-A through XP-G corresponding mutations in 8 different genes (DDB2, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, POLH, XPA) and one variant form with slight differences in the clinical presentation and mutated in XPC. The incidence of XP in Tanzania is higher than in Europe due to consanguineous marriages (Hafidh et al, 2023). As genetic profiling is not performed in clinical routine in Tanzania, knowing the XP subgroup landscape in Tanzania would help both assess the risk of complications and improve the management of these patients.
However, so far the genetic landscape of XP patients in Tanzania remains unknown.
This study aims to determine the genetic subtypes of XP in Tanzania and its correlation with clinical presentation leveraging blood whole-exome sequencing of patients with XP and their parents if available.
Our results show a higher prevalence of mutations linked to the XPD (homozygous mutations) and XPB (heterozygous “double-hit” mutations) subgroups. XPE, XPD and XPC are not carried by any of the patients.

146

Evolution on graphs and the transition to cancer

Chay Paterson[1]

[1] University of Manchester, United Kingdom

Presenter: Chay Paterson

Multi-stage mathematical models have contributed a great deal to our understanding of cancer, from Armitage and Doll's original work in the 1950s, to Knudson's studies of retinoblastoma in the 1970s. My ongoing research extends these models using evolutionary graph theory, with edges on the graph representing distinct molecular mechanisms. This approach links population incidence to microscopic mechanisms, connecting epidemiology to genomics.

My latest advance is a new semi-analytical numerical method to compute hazard curves, which is a much more efficient way to compute probabilities and likelihoods than random sampling, or approximate Bayesian computation. By combining this method with maximum likelihood estimation, we can make new, precise estimates of mutation rates from data on age and copy number alterations. In particular, when applied to vestibular schwannoma, this approach enables new estimates of mutation rates in Schwann precursor cells.

147

Structure-based models to identify driver genes and mutations

Stefano Pellegrini[2], Olivia Dove[2], Ferran Muiños[2], Abel Gonzalez-Perez[2], Núria López-Bigas[1]

[1] ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
[2] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Stefano Pellegrini

In cancer development, certain mutations in specific genes, called cancer drivers, confer cells a selective advantage and are thus under positive selection. Identifying these genes and the specific mutations capable of driving tumorigenesis is a key aim of cancer genomics. In recent decades, bioinformatics methods that exploit the traces that positive selection leaves in mutational patterns across tumours have been developed to accomplish this objective. One such signal of positive selection is the uneven mutation distribution in the 3D structure of the protein with respect to the expectation under neutrality. Methods that exploit it have been hampered by the paucity of high-quality 3D structures of proteins. A possibility to close this gap has been opened recently by the availability of AlphaFold2 models for the entire human proteome.
Here, we present Oncrodrive3D, a fast and accurate novel 3D-clustering algorithm for driver genes discovery. Our approach involves analysing patterns of observed mutations in cancer to identify volumes that exhibit a higher-than-expected (under neutrality) frequency of mutations. Oncodrive3D leverages AlphaFold's structure predictions and predicted aligned error (PAE) to construct contact probability maps. Moreover, it uses the observed tri-nucleotide mutation profiles of cohorts to simulate neutral mutagenesis and rank-based statistics to compute empirical p-values for each volume of mutated residues. In a survey across 28,000 tumours from 44 cancer types, Oncodrive3D outperforms a state-of-the-art 3D-clustering-based driver discovery method and complements other driver discovery methods based on different signals of positive selection.
In addition to the aforementioned advances, over the past decade, researchers have devised machine learning-based methods that leverage signals of positive selection as features of observed mutations to comprehensively identify all potential driver mutations within cancer genes across different types of cancer. In this context, we are exploiting structural mutational features to build structurally-aware models for the identification of driver mutations. Our initial findings indicate that incorporating structural features helps shed light on the mutational mechanisms of tumorigenesis across genes and tissues.

148

Comparative analysis of non-coding constraint mutations in canine and human osteosarcoma

Raphaela Pensch[3], Sophie Agger[1], Suvi Mäkeläinen[3], Sharadha Sakthikumar[3], Jaime Modiano[2], Lauren E. Burt[2], Karin Forsberg Nilsson[3], Maja Louise Arendt[1], Kerstin Lindblad-Toh[3]

1 University of Copenhagen, Denmark
2 University of Minnesota, United States of America
3 Uppsala University, Sweden

Presenter: Raphaela Pensch

Osteosarcoma (OSA) is a common primary bone cancer in dogs and serves as a valuable model for the corresponding human cancer, which is rare and hence harder to study. Protein-coding mutations and their role in driving cancer have been extensively analyzed. Non-coding mutations have also been recognized as cancer drivers, however, their involvement in canine OSA and their translational relevance remain unexplored. In this study, we applied an evolutionary constraint approach that combined tumor/normal whole-genome sequencing data from 116 canine patients with constraint scores derived from the 240 mammals in the Zoonomia project to investigate the role of non-coding mutations in OSA. Furthermore, our findings in the canine cohort were compared to somatic variants from 38 human OSA patients obtained from the International Cancer Genome Consortium. We identified non-coding constraint mutations (NCCMs) in evolutionarily conserved positions and extracted genes with an enrichment of NCCMs in their regulatory regions. Our candidate gene set exhibited a significant overrepresentation of cancer-related genes as well as genes that are transcriptionally regulated by RUNX2, a protein critical for skeletal development and osteoblast differentiation. Several genes showed a strong NCCM enrichment in both the human and canine dataset and their NCCMs emerged as potential novel non-coding drivers in OSA. Our findings highlight the power of evolutionary constraint in predicting the functional impact of somatic non-coding mutations and identifying novel candidate drivers. This could lead to novel therapeutic targets and improved patient stratification in the future. Moreover, our study contributes to a deeper understanding of OSA and emphasizes the significance of canine OSA as a reliable model for studying the human disease.

Raphaela Pensch is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

149

A comprehensive survey of somatic mutations in normal human cells

My Pham[1], Thomas R. W. Oliver[1], Ellie Dunstone[1], Rashesh Sanghvi[1], Pantelis Nicola[1], Andrew Lawson[1], Yvette Hooks[1], Federico Abascal[1], Emily Mitchell[1], Tim Coorens[1], Matthew Neville[1], Luke Harvey[1], Inigo Martincorena[1], Raheleh Rahbari[1], Michael Stratton[1]

[1] Wellcome Sanger Institute, United Kingdom

Presenter: My Pham

A ubiquitous feature of normal human tissues is their accumulation of somatic mutations with age. Recent studies have shown that the somatic mutation burden varies between cell types, tissues, organs, and individuals. However, for technical reasons, the numbers of cell types/tissues studied has previously been limited and the causes of the variation in mutation rates between cell or tissue types are yet to be uncovered.
Our present study catalogues somatic mutations from 52 distinct microscopic tissue structures, encompassing both somatic and germline cells, from individuals with an age range of 9-90 years. We used NanoSeq, which is based on duplex-sequencing of individual DNA molecules with < 5x 10-9 errors/bp, to detect somatic mutations in all cell types, irrespective of their ability to form localised clonal structures.
We observed significant variation in mutation burden and rates across different cell types. The skin demonstrated the highest mutation load, mainly due to UV light exposure. The liver exhibited the next highest mutation burden, with ~7000 mutations detected in normal hepatocytes from an 89-year-old. In people in their fifties, the mutation load in epithelial, mesenchymal and haematological cells generally ranged from 1000-5000. In contrast, the lowest mutation burden was found in testes from which sperm originate, with only 300 mutations at a similar age. Most cell types displayed linear mutation accumulation with age, including neurons, smooth, cardiac and skeletal muscle which are known to be post-mitotic indicating that many somatic mutations are not generated during cell division.
Our analysis revealed the presence of mutational signatures SBS1 and SBS5 across all tissue structures, albeit to different extents. Also, we observed distinct tissue-specific signatures related to local exposures. For instance, SBS7, the UV-radiation signature, was present in skin, while SBS4, associated with tobacco smoking, was extracted from lung tissues and SBS9, due to somatic hypermutation, was found in B lymphocytes.
The use of NanoSeq allowed a more comprehensive investigation of the breadth of cell types in the healthy human body, compared to previous pan-tissue studies. The results provide a better understanding of the cell-type-specific contributions to age-related disease development.

150

Clonal evolution preceding cancer revealed using single-cell DNA sequencing and computational modelling

Jamie Blundell[1], Yeuk Pin Gladys Poon[4], Aditi Vedi[3], Elisa Laurenti[4], Mathijs Sanders[2], Peter Valk[2]

1 Early Cancer Institute, United Kingdom
2 Erasmus University Medical Center, The Netherlands
3 Stem Cell Institute, University of Cambridge, United Kingdom
4 University of Cambridge, United Kingdom

Presenter: Yeuk Pin Gladys Poon

The acquisition and expansion of cancer driver mutations can predate diagnosis by decades. However, our understanding of the evolution that occurs in these early stages of cancer evolution remains limited because most studies focus only on sequencing the malignant clone that “wins” the evolutionary race. The representation of driver mutations in the stem cells which do not transform into the malignant clone provides a window into this precancerous evolution. Here, we isolate pre-leukaemic haematopoietic stem cells (pHSCs) from the marrow of 17 Acute Myeloid Leukaemia (AML) patients and perform single-cell DNA sequencing on tens of thousands of pHSCS to reconstruct the evolution of AML from healthy HSCs. In the majority of cases AML required the acquisition of 3 or 4 driver mutations in the same cell lineage, however the phylogenetic trees which led to this end-point are highly variable. The structure of the trees and the sizes of intermediate pre-leukaemic clones encode important information about the roles of selection, mutation and chance during precancerous evolution. Using stochastic simulations and machine learning approaches, we show that the observed trees are consistent with a model of somatic evolution in which acquisition of each driver mutation is stochastic but provides a fitness advantage in a relatively narrow range providing only a modest growth advantage to the clone relative to its ancestor. Individuals destined to develop AML late in life are likely to habour multiple-mutant clones (those with two or three drivers) that have arisen anomalously early in life. This points towards more rational approaches for improving cancer risk prediction based on the landscape of somatic clones in pre-cancerous tissues.

151

DNA damage in alveolar stem cells mirrors long-term lung cancer risk in smokers

Moritz Przybilla[2], Amany Ammar[1], Andrew Lawson[2], Pantelis Nicola[2], Inigo Martincorena[2], Peter Campbell[2], hyunchul jung[2], Chuen Ryan Khaw[1], Zoe Frazer[1], Sam Janes[1], Kate Gowers[1], Kate Davies[1]

1 University College London, United Kingdom
2 Wellcome Sanger Institute, United Kingdom

Presenter: Moritz Przybilla

Extensive epidemiological studies have illuminated the intricate relationship between tobacco smoking and lung cancer (LC). Squamous cell carcinoma (SCC), which is clinically tied to smoking, is formed in the proximal airway of the lung, whereas adenocarcinomas (LUAD), arising in the distal airway, have a weaker connection to tobacco and often occur in never-smokers. The cessation of smoking promotes lung health and gradually lowers cancer risk. However, dissecting the impact of smoking cessation on different LC types, comprehensive studies have established that the risk of SCC diminishes rapidly after quitting, while LUAD risk endures. We previously found that normal bronchial epithelium in the proximal airway accumulates DNA damage from smoking. However, our analysis identified a population of cells with near-normal mutation burden in ex-smokers, evading tobacco-induced damage and potentially explaining a reduction in SCC risk with time.
The genomics of the distal airway have not been extensively explored. Here, we established organoids from single-cell-derived alveolar stem cells (ASCs), which are believed to be the origin of LUAD. Analysing >800 organoids from nine patients, we found that tobacco smoking is the major driver of increased mutation burden in ever- compared to never-smokers. Despite that, we did not detect significant differences in the burden of genomic damage between ex- and current-smokers, suggesting that smoking-induced damage in the distal airway is both persistent and long-lasting. Moreover, we observed a more uniform pattern of genomic changes across ASCs, potentially accounting for the ongoing risk of LUAD even after smoking cessation.
Lastly, we used single-molecule sequencing to investigate driver mutations in a cohort of 43 individuals. While we detected positive selection of TP53, we also noted recurring mutations in surfactant proteins, a mechanism previously associated with lineage definition and an imprint of ASCs as the origin of LUAD. Although we found mutations at common LUAD hotspots, ASCs generally exhibited driver paucity. Overall, contextualising our findings within the carcinogenic process promises to enhance our understanding of LC development and the association between genomic changes and cancer risk.

152

YAP/TAZ activation predicts clinical outcomes in mesothelioma and is conserved in in vitro model of driver mutations

Krishna Purohit[1], Richard Cunningham[1], Carsten Gram Hansen[1], Siyang Jia[1], Ning Sze Hui[1], Yue Lin[1], Neil Carragher[3], Richard Cunningham[1], Adriano Rossi[1], Janne Lehtiö[2]

1 Institute for Regeneration and Repair, The University of Edinburgh, United Kingdom
2 Karolinska Institute and SciLifeLab, Sweden
3 MRC Institute of Genetics & Cancer, The University of Edinburgh, United Kingdom

Presenter: Krishna Purohit

The Hippo signalling pathway is dysregulated across a wide range of cancer types and, although driver mutations that directly affect the core Hippo components are rare, a handful is found within pleural mesothelioma (PM). PM is a deadly disease of the lining of the lung caused by asbestos exposure. By pooling the largest-scale clinical datasets publicly available, we interrogated associations between the most prevalent driver mutations within PM and Hippo pathway disruption in patients, while assessing correlations with a variety of clinical markers. This analysis reveals a consistent worse outcome in patients exhibiting transcriptional markers of YAP/TAZ activation, pointing to the potential of leveraging Hippo pathway transcriptional activation status as a metric by which patients may be meaningfully stratified. Preclinical models recapitulating disease are transformative in order to develop new therapeutic strategies. We here establish an isogenic cell-line model of PM, which represents the most frequently mutated genes and which faithfully recapitulates the molecular features of clinical PM. This preclinical model is developed to probe the molecular basis by which the Hippo pathway and key driver mutations affect cancer initiation and progression. Implementing this approach, we reveal the role of NF2 as a mechanosensory component of the Hippo pathway in mesothelial cells. Cellular NF2 loss upon physiological stiffnesses analogous to the tumour niche drive YAP/TAZ-dependent anchorage-independent growth. Consequently, the development and characterisation of this cellular model provide a unique resource to obtain molecular insights into the disease and progress new drug discovery programs together with future stratification of PM patients. Proteomics and mRNA level investigations reveal potential therapeutic targets that drive cellular dysregulation upon NF2 and BAP1 loss. Targeting these selected dysregulated proteins, and generating knock out of YAP, but not TAZ in the established NF2 deficient cellular model exhibits a change in the molecular phenotype associated with NF2-loss linked dysregulation. This strategy allows for the evaluation of Hippo pathway downstream mediators such as YAP/TAZ and identified dysregulated targets as potential drug targets for future stratification.

153

Deciphering drug-induced transcriptional responses in cancer cells

Adelina Rabenius[1], Anam Minhas[1], Anniina Vihervaara[1]

[1] KTH Royal Institute of Technology, Sweden

Presenter: Adelina Rabenius

Understanding how cancer cells respond to drug treatments on a transcriptional level is crucial for developing and advancing future cancer diagnostics and treatments. A deeper understanding of the precise mechanisms that coordinate RNA Polymerase II (Pol II) molecules across genes and enhancers during and after drug treatments are expected to enhance efforts in targeted medicine and counter drug resistance and relapses. In our on-going study, we have investigated how leukemia K562 and lymphoma L428 cell lines respond and adapt to the classical chemotherapy drugs doxorubicin and vinblastine over a time course of treatment and recovery. By studying nascent RNA synthesis with Precision Run-On sequencing (PRO-seq), and levels of mature RNAs with RNA-seq, we aim to decipher the mechanistic basis of drug response on a transcriptional level. The tracking of Pol II complexes and production of mature RNAs will be conducted in combination with live-cell microscopy to analyze cell morphology, proliferation, and adaptation on an organelle level. Our preliminary results identify genes associated with drug-resistance and metastasis in cells that survive drug treatments. Thus, we highlight the necessity of comprehending the detailed cancer cell responses stemming from drug treatment.

154

Chromatin state annotation to unravel epigenetic changes in tumorigenesis with uterine leiomyoma as a model

Maritta Räisänen[1], Eevi Kaasinen[1], Niko Välimäki[1], Lauri Aaltonen[1]

[1] University of Helsinki, Finland

Presenter: Maritta Räisänen

Epigenetic changes in chromatin are often seen in tumorigenesis of both benign and malignant tumors. Uterine leiomyomas (ULs) are benign tumors originating from the uterine wall and affect ~70% of women during their reproductive years. ULs can be divided into five mutually exclusive molecular subclasses characterized by mutations in MED12, overexpression of HMGA2, biallelic inactivation of FH, deficient H2A.Z. loading by mutations on SRCAP complex genes and defects in Cullin 3-RING E3 ligase neddylation. In our previous study, we reported changes in chromatin accessibility in all UL subclasses compared to myometrium (Berta et al. 2021), suggesting that epigenetic changes play a role in tumorigenesis of ULs. As a very common tumor, ULs provide an interesting platform to study epigenetic changes occurring during tumorigenesis, which might be challenging with primary tissues from malignant tumors with only a small amount of starting material available.
To study the overall changes in chromatin states, we generated ChIP-seq data against five histone markers (H3K4me1, H3K4me3, H3K27me3, H3K36me3 and H3K9me3) from fresh frozen tissues (three tumors per subclass and four myometrium). With these data we created a 15-state chromatin segmentation for the most common molecular subclasses of ULs (MED12, HMGA2 and FH), as well as normal myometrium using chromHMM (Ernst & Kellis, 2012).
We used a Roadmap epigenomics trained model for 15 chromatin states to annotate the segmentations for each myoma subclass and normal myometrium, as well as separately for each individual tumor. Integration with other data from the same molecular subclasses of ULs will provide a more thorough understanding of the chromatin state changes causing and/or caused by the respective driver changes in these benign tumors. These findings provide an important framework to understand genome regulation in tumorigenesis. As inactivation of FH causes cancer and overexpression of HMGA2 occurs in malignant tumors, these results from benign ULs can be applied in the context of cancer and broaden our understanding of genome regulation in tumorigenesis.

155

Deciphering whole genome doubling and temporal mutational dynamics in mouse models of pancreatic ductal adenocarcinoma (mPDAC)

Anantharamanan Rajamani[2], Chiara Falcomatà[2], Deelaka Wellappili[2], Michael Menden[1], Dieter Saur[2]

1 Helmholtz Munich, Germany
2 Technical University of Munich, Germany

Presenter: Anantharamanan Rajamani

The influence of whole-genome doubling (WGD) and chromosome arm aneuploidies (CAAs) on the progression of pancreatic ductal adenocarcinoma (PDAC) remains largely unexplored. Here, we analyze ten whole-genome, 149 whole-exome, and 375 low-coverage whole-genome sequencing datasets of primary low-passaged PDAC cell cultures originating from genetically engineered mouse models driven by Kras^{G12D} or Pik3ca^{H1047R} mutations. Our investigation unveiled the occurrence of whole-genome doubling (WGD) in more than 40% of the cell cultures. We determined the molecular timing of 22 crucial somatic mutations, such as Trp53, Map2k4, Fbln2, Jak3, and Grin2a, based on their emergence during tumor evolution. Furthermore, we detected recurrent CAAs on chromosomes 4, 6, and 19 in our cohort. The systematic co-occurrence and mutual exclusivity (COME) analysis revealed a co-occurrence between CAAs on chromosomes 4 and 19. In contrast, CAAs on chromosomes 4 and 6 displayed a pattern of mutual exclusivity. By employing a multivariable logistic regression model, we identified an arm-level deletion of chromosome 4 as a potential early event, showing an association with WGD. Mutational timing and cancer progression model analyses further confirmed the early acquisition of chromosome 4 deletion in genome-doubled tumors. Additionally, distinct patterns of copy number signatures and arm-level aneuploidies were observed between Kras^{G12D} and Pik3ca^{H1047R} driven tumors. Moreover, the elevated ploidy levels in Pik3ca^{H1047R}, Kras^{G12D}/Pik3ca^{H1047R}, and Kras^{G12D} tumors with Trp53-altered genotypes suggest a connection between Pik3ca mutations and Trp53 in WGD-positive tumors. In summary, our study describes a detailed characterization of copy number signatures in mouse-derived primary PDAC cell cultures, providing insights into the potential role of WGD and CAAs in PDAC evolution.

Anantharamanan Rajamani is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

156

In silico saturation mutagenesis to identify clonal hematopoiesis driver mutations

Santiago Demajo[2], Ferran Muiños[2], Joan Enric Ramis Zaldivar[2], Miguel Grau[2], Maria Andrianova[2], Abel Gonzalez-Perez[2], Núria López-Bigas[1]

1 ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
2 Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Joan Enric Ramis Zaldivar

Clonal hematopoiesis (CH) is a common condition characterized by a clonal expansion in the blood caused by somatic mutations in hematopoietic stem cells that confer them a selective advantage. CH is usually linked to aging or to chemotherapy exposure, and is associated with increased risk of hematological cancer and other pathologies such as cardiovascular diseases and certain solid tumors. Although in recent years the main CH driver genes have been characterized, identifying which specific mutations in those genes are capable of driving CH is still an unsolved problem. Here, we repurposed a machine learning method originally devised to identify cancer driver mutations to identify blood somatic mutations across more than 36,000 individuals. We obtained reliable BoostDM-CH models for twelve genes, including the most common CH drivers DNMT3A, TET2, and ASXL1. These models provide a thorough picture of potential driver mutations in each gene, defining the specific features that characterize them. BoostDM-CH models show an accuracy comparable to currently used manually curated rules in identifying CH driver mutations with the advantage of being automatic and unbiased. Furthermore, the availability of BoostDM models trained on myeloid cancer mutations for 4 of these CH driver genes allows to highlight similarities and differences in the configuration of CH and cancer driver mutations. Using a large cohort of close to 470,000 individuals from the UK Biobank we show that CH driver mutations identified by BoostDM-CH models (unlike non-driver mutations) are highly correlated with age and are associated with an increased risk of hematological cancer (especially myeloid neoplasms), and some solid tumors such as lung and breast cancers. Additionally, by estimating the fitness of specific mutations in this large cohort we observed that CH driver mutations were associated with a higher fitness than non-drivers. In summary, we developed and validated gene-specific machine learning models for in silico saturation mutagenesis to identify CH driver mutations. These comprehensive models may support the identification and clinical interpretation of CH mutations in newly sequenced individuals.

157

Experimental and statistical evidence for fitness-altering tumour mutations acting via long non-protein-coding RNAs

Sunandini Ramnarayanan[1], Roofiya Koya[1], Jack Roban[1], Rory Johnson[1]

[1] University College Dublin, Ireland

Presenter: Rory Johnson

The extent and nature of genetic mutations driving the pathogenic fitness of tumor cells remain unclear. Statistical analyses of whole-genome sequencing projects have revealed a high mutational burden in diverse non-protein-coding 'driver' elements. Particularly interesting amongst these are long non-coding RNAs (lncRNAs), due to their widespread dysregulation in tumours and growing functional evidence implicating them in cancer hallmarks. However, conclusive experimental evidence linking non-coding mutations to cell fitness is lacking.

Thus we are interested in the possibility that tumour single nucleotide variants (SNVs) can enhance cell fitness through lncRNAs. We analyzed cohorts from PCAWG, Hartwig, and most recently, Genomics England, to identify 'driver' lncRNAs that carry significant mutational burdens. These lncRNA SNVs are consistent across cohorts, specific to particular tumour types, and predictive of patient survival. Most important, when lncRNA SNVs are experimentally introduced into living cells, they reproducibly enhance proliferative fitness in vitro and accelerate tumour growth in vivo. Ongoing studies integrating bioinformatic analyses with high-throughput CRISPR-Cas genome editing are shedding light on underlying molecular mechanisms by which lncRNA SNVs promote cellular fitness.

In conclusion, this work presents experimental and statistical evidence for the importance of non-protein-coding RNA mutations in shaping tumour cell phenotypes, and demonstrates that fitness-altering tumour mutations are more diverse than previously thought.

158

Web-based bioinformatic tool LYNX: lymphoid next-generation sequencing data analysis and visualization in hematological malignancies

Jakub Porc[1], Tomáš Reigl[1], Veronika Navrkalova[2], Jakub Hynšt[1], Karol Pal[1], Kamila Stránská[2], Jana Kotaskova[2], Sarka Pospisilova[2], Karla Plevova[2]

1 CEITEC, Masaryk University, Czech Republic
2 CEITEC, Masaryk University & University Hospital Brno, Czech Republic

Presenter: Jakub Porc

Introduction and aims: Next-generation sequencing (NGS) is rapidly advancing technology in cancer genomics. The single NGS experiment can replace several laboratory tests and provide requested information about the tumor samples. However, new requirements for bioinformatic tools and computational resources are emerging in diagnostics and research.
Methods: Following the development of an integrative, targeted NGS panel LYNX, enabling the detection of genetic markers recurring in lymphoid malignancies (PMID: 34082072), we focused on designing a LYNX bioinformatics tool, featuring all necessary analytical steps as well as a user interface (UI) for easy data representation and interpretation. State-of-the art resources were used to achieve our goal, such as React.js, D3.js visualisation library, Integrated Genome Browser and GATK industry standard variant calling pipeline.
Results: We have created an online tool LYNX for targeted NGS data analysis. It utilizes our custom pipelines to provide information about single nucleotide variants in regions of interest, genome-wide copy number variation, antigen receptor rearrangements and lymphoma-specific translocations. Results are available in the form of interactive tables and graphs, allowing fast and simple analysis.
Conclusion and discussion
Our tool's comprehensive features make data processing simple and speed up result interpretation. It supports the identification of genetic markers for individual cases and is tailored for both research and diagnostic objectives. LYNX is appropriate for other applications as it can be simply customized for different diagnostic panels.
Acknowledgments: MH CZ DVO FNBr65269705, AZV NU21-08-00237, NU22-08-00227, NPO-NUVR LX22NPO5102. Computational resources supplied by e-INFRA LM2018140.

159

CLLue: Searching for connections among clinical, biological, and molecular features in the dataset of leukemia patients

Tomas Reigl[1], Veronika Navrkalova[2], Jakub Porc[1], Karol Pal[1], Kamila Stránská[2], Sarka Pospisilova[2], Karla Plevova[2]

[1] CEITEC, Masaryk University, Czech Republic
[2] CEITEC, Masaryk University & University Hospital Brno, Czech Republic

Presenter: Tomas Reigl

Introduction: Today's analytic techniques produce extensive molecular biological and genetic information for every tumor sample. Understanding the significance of each feature and its connections with a patient's clinical history is one of the greatest challenges. Advanced computational approaches can provide such information, leading to the discovery of novel attributes critical for diagnosis and potentially new classifiers for categorizing patients into distinct groups. We investigated such attributes in a local dataset of patients with chronic lymphocytic leukemia (CLL).
Methods: We chose a dataset of 100 CLL patients with two samples acquired at distinct time points: diagnosis and a follow-up after a median of 36 months. The data were examined using CLLue, a web-based tool for multiple statistical and visualization methods. The data underwent visual exploratory data analysis and computational analysis using statistical tests, machine learning, and clustering techniques.
Results: Interactive visualizations in CLLue helped us identify outliers and correlate features to predict the number of patient clusters. We applied different clustering methods using several pre-selected feature sets, revealing various outcomes. These results were inspected to identify key features defining clusters distinctly. The most important features (p-value < 10^-8, XGBoost Importance Gain > 0.1) included absolute lymphocyte count (ALC) rise per month, CLL activity, treatment initiation after time point 2 (TP2), ALC in TP2, and a period between TP1 and TP2. Final clustering focused on these features, resulting in three potentially clinically relevant clusters.
Conclusion and discussion: We've shown that CLLue can improve the users' understanding of a given dataset and establish clusters of data with a similar profile of features. However, it is essential to interpret the resulting data within the clinical context to avoid misinterpretation. Currently, CLLue works in a single-threaded computational mode, and analyzing large datasets (thousands of features per patient) can be time-consuming. Therefore, our focus in the following development will be on parallelization.
Acknowledgments: MHCZ DRO FNBr65269705, AZV NU21-08-00237, NU22-08-00227, and NUVR LX22NPO5102. Comp. resources by e-INFRA LM2018140.

160

Mutational topography reflects clinical neuroblastoma heterogeneity

Elias Rodriguez Fos[6], Mercè Planas Fèlix[7], Martin Burkert[3], Montserrat Puiggròs[1], Joern Toedling[4], Nina Thiessen[2], Eric Blanc[2], Annabell Szymansky[4], Falk Hertwig[4], Naveed Ishaque[2], Dieter Beule[2], David Torrents[8], Angelika Eggert[5], Richard Koche[10], Roland Schwarz[9], Kerstin Haase[4], Johannes H. Schulte[5], Anton Henssen[6]

1 Barcelona Supercomputing Center - Centro Nacional de Supercomputación, Spain
2 Berlin Institute of Health at Charité (BIH), Germany
3 Centre for Molecular Medicine Norway (NCMM), Norway
4 Charité, Germany
5 Charité - University Medicine Berlin, Germany
6 Charité - University Medicine Berlin; Max Delbrück Center for Molecular Medicine, Germany
7 Charité/MDC/DKFZ, Germany
8 ICREA, Barcelona Supercomputing Center, Spain
9 MDC/CCCE/BIFOLD, Germany
10 Memorial Sloan Kettering Cancer Center, United States of America

Presenter: Elias Rodriguez Fos

The activity of different endogenous and/or exogenous mutational processes imprint characteristic patterns of mutations in the genome defined as mutational signatures. Recent analyses in neuroblastoma have extracted mutational signatures associated with simple and specific complex variants such as extrachromosomal circular DNA (ecDNA), chromothripsis, and breakage-fusion-bridge cycles. However, our current understanding of the whole catalogue of complex rearrangement patterns and mutational processes, their co-occurrence, as well as their impact on clinical outcomes in neuroblastoma remains limited.

Applying variant detection algorithms and signature analyses to 114 neuroblastoma tumor-normal pairs, we extracted 10 catalogued mutational signatures associated with SNVs and Indels as well as 13 recently defined CNA and SV signatures. Interestingly, we observed that specific mutational processes were differentially active across neuroblastoma risk groups. Using statistical and graph-based methods, we integrated mutational signatures with complex structural variant pattern classes as well as clinical annotations, identifying previously unrecognized co-occurring mutational footprints, which we termed mutational scenarios.

We demonstrate that clinical neuroblastoma heterogeneity is linked to differences in the processes driving these mutational scenarios. Whereas high-risk MYCN-amplified neuroblastoma genomes were characterized by signs of damage caused by ROS, replication slippage and stress, and ecDNA presence, homologous recombination-associated signatures along with non-circular complex rearrangements defined high-risk non-MYCN-amplified patients. Non-high-risk neuroblastomas, on the other hand, were marked by footprints of chromosome missegregation and TOP1 mutational activity. This analysis provides a systematic perspective on the repertoire of mutational patterns that contribute to clinical neuroblastoma heterogeneity.

161

A mathematical model for pancreatic cancer during intraepithelial neoplasia*

J. Roberto Romero Arias[2], Joshua Briones-Andrade[1], Guillermo Ramirez-Santiago[3]

1 Facultad de Ciencias National Autonomous University of Mexico, Mexico
2 IIMAS National Autonomous University of Mexico (UNAM), Mexico
3 Institute of Mathematics, National Autonomous University of Mexico (UNAM), Mexico

Presenter: Guillermo Ramirez-Santiago

Cancer is the result of complex interactions of intrinsic and extrinsic cell processes, which promote sustained proliferation, resistance to apoptosis, reprogramming and reorganization. The microenvironmental conditions of certain complexes and the understanding of certain signalling pathways as well as some important molecular mechanisms are necessary elements to try to understand the evolution of any type of cancer. As in most cancer quantitative models, the understanding of the early onset of pancreatic adenocarcinoma requires a multiscale representation of the cellular microenvironment through elastic cell interactions and their intercellular communication mechanisms, such as growth factors and cytokines. In this paper we propose a gene regulatory network associated with the processes of proliferation and apoptosis of pancreatic cells and its representation in terms delayed differential equations to mimic cell development. Likewise, we couple the cell cycle with the spatial distribution of cells and the transport of growth factors to show that inflammatory processes trigger the adenocarcinoma evolution. Finally, we show that the oncogene RAS may be an important target to develop anti-inflammatory strategies that limit the emergence of more aggressive adenocarcinomas.

*Supported by DGAPA-UNAM under contract: IN109722

162

Estimating mutation risks conferred by mutational processes in cancer genomes

Axel Rosendahl-Huber[2], Ferran Muiños[2], Abel Gonzalez-Perez[2], Núria López-Bigas[1]

[1] ICREA and Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain
[2] Institute for Research in Biomedicine Barcelona (IRB Barcelona), Spain

Presenter: Axel Rosendahl-Huber

DNA mutations can disrupt vital cellular functions, leading to disease. One of the most striking examples is cancer, which arises as a result of mutations affecting cancer driver genes. Determining the risk of mutations in cancer driver genes across tissues depending on their exposure to different mutational processes and comparing them to their observed prevalence can give insight into the fitness provided by different mutations. The estimation of the risk to acquire specific types of mutations, such as neoantigens, may also be exploited by personalized cancer medicine. The mutation risks of a gene or given region depends on the mode of action of the mutational process. Furthermore, these rates are associated with chromatin organization, sequence context and the local activity of DNA repair. Thus, mutation risks may display considerable variation across cells depending on exposure and genomic determinants. We aim to determine the mutation risk for individual mutational processes across the genome using statistical learning approaches. Currently, we have explored mutation risk using linear regression and decision tree based methods. As input for modeling, we use mutation probabilities from 33 mutational processes, defined as mutational signatures, using WGS-derived mutation data from the PCAWG and HMF cohorts (>7100 genomes, >150M mutations). Mutations are distributed in genomic bins ranging from 1 mb to exons. As determinants for the mutation rates we use the mean scores of epigenetic modifications, replication timing and transcription factor binding site data, which are known predictors of mutation rates. Currently, we are able to predict genomic mutation risks for different mutagenic processes (R2 values ranging between 0.6 and 0.9), with tree-based regression performing better at predicting mutational loads. Decomposing the contribution of covariates is used to gain insights in the factors underlying the variance in mutation risks. The models generated during this study provide the basis to calculate relative risk scores for mutagenic processes for specific mutations across the genome. In future research efforts, these risk scores can be used to determine the role of different endogenous and environmental mutagenic processes in the induction of cancer driver mutations and neoantigens.

163

Exploring the exome of lung adenocarcinoma in Mexican patients and some insights about worldwide discrepancies in its treatment

Bertha Rueda[4], Jorge Melendez-Zajgla[2], Humberto García-Ortiz[2], Lorena Orozco[2], Joaquín Zúñiga-Ramos[1], Federico Ávila-Moreno[3], Gustavo Ramírez-Martínez[1]

1 Instituto Nacional de Enfermedades Respiratorias, Mexico
2 Instituto Nacional de Medicina Genómica, Mexico
3 Universidad Nacional Autónoma de México, Mexico
4 Universidad Nacional Autónoma de México / Instituto Nacional de Medicina Genómica, Mexico

Presenter: Bertha Rueda

Lung cancer has the highest mortality rates worldwide, mainly due to tobacco exposure. The most common type is lung adenocarcinoma, which has various mutations that can be targeted with certain drugs. The prevalence of these mutations and risk factors varies among different populations. New treatment approaches focus on targeted therapies and immunotherapies. However, access to these treatments is limited to developed countries due to the high costs involved and the lack of specific genomic data for local populations. Disparities in lung cancer treatment are explored in the editorial of the journal "Lung Cancer Worldwide" in the context of different countries. For example, some countries like Sudan lack access even to the first generation of EGFR TKIs, while in European countries, certain targeted therapies are only available through patient self-funding. Our study investigates the genomic landscape of Mexican patients with lung adenocarcinoma using whole-exome sequencing. We examined 25 patients from the National Institute of Respiratory Diseases in Mexico. The frequencies of EGFR and KRAS mutations were found to be 24% and 19% respectively. Patients with EGFR mutations were correlated with Native American ancestry and familial history, while KRAS mutations were associated with smoking history and European heritage. The most common mutation signatures in our cohort were SBS1 (resulting from 5-methylcytosine deamination), SBS5 (related to tobacco exposure), and SBS40 (age-related). Within our population, exposure to wood smoke emerged as a significant risk factor (30%, 3 out of 10 cases), along with mutations in less frequently reported genes like KMT2C, which may have potential relevance and require further comprehensive characterization. Data suggests both differences and similarities between Mexican patients with adenocarcinoma and patients from other countries. This observation is important as clinical decisions are primarily based on reports from European populations

Bertha Rueda is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

164

Identification of potential neoantigens in cancer-associated fibroblasts isolated from breast cancer patients

Kevin Ryan[1], Barry Digby[1], Domhnall O'Connor[1], Laura Barkley[1], Pilib Ó Broin[1]

[1] University of Galway, Ireland

Presenter: Kevin Ryan

Background: Cancer-associated fibroblasts (CAFs) are a heterogeneous cell type found in the tumour microenvironment (TME). CAFs support tumour growth and metastasis and contribute to therapeutic resistance. CAFs also have a significant impact on immune infiltration and immune responses in the TME. Therefore, therapeutic targeting of CAFs is a viable strategy in the treatment of cancer. In this study, we aim to identify somatic mutations in CAFs, potentially giving rise to neoantigens. Ultimately, we aim to elucidate the therapeutic potential of targeting CAFs through the exploitation of CAF-specific neoantigens.
Methods: CAFs and corresponding tumour-associated normal fibroblasts (TANs) were cultured from tissue of 12 breast cancer patients (11 Luminal A and one triple-negative). Bulk RNA-sequencing was carried out on all samples. Leveraging publicly available data, CIBERSORTx was used to characterise CAFs and TANs into three fibroblast subpopulations. Whole-exome sequencing (WES) was carried out on CAFs and TANs from six patients. Landscape of Effective Neoantigens Software (LENS) was used to identify CAF-specific neoantigens.
Results: Our studies confirm the heterogeneity of our patient-derived CAFs and TANs, with the immunosuppressive-myofibroblastic subpopulation being the most prevalent in our samples. This is important as for the effective design of CAF-targeting therapies, it is necessary to target pro-tumourigenic CAF subpopulations. WES identified 13 private missense mutations, with five of the six patients exhibiting one or more such variants. Interestingly, genes with these mutations included previously reported CAF markers and genes implicated in tumour metabolism, specifically lipid metabolic pathways. CAFs contribute to lipid metabolism within the TME, thus playing a vital role in cancer progression and tumour immunogenicity. The next step is to validate these CAF neoantigens using T-cell immunogenicity assays. These studies may unravel the potential of targeting CAF neoantigens as a way of enhancing the efficacy of anti-cancer therapy.

165

GROVER: building a language model of the human genome

Melissa Sanabria[1], Anna Poetsch[2], Jonas Hirsch[1]

[1] Dresden University of Technology, Germany

Presenter: Melissa Sanabria

Large Language Models have enabled the generation of coherent and syntactically accurate text in natural language, a significant advancement that can be adapted to (cancer) genomics. DNA sequence follows rules similar to natural language, but a distinguishing factor is the absence of a concept of words. We established byte-pair tokenization on the human genome and trained a language model called GROVER (“Genome Rules Obtained Via Extracted Representations") which selected the optimal vocabulary with a custom fine-tuning task of next-k-mer prediction. This vocabulary, composed of 601 different words/tokens, is optimized for the information content for DNA language models, outperforming fixed-size k-mer models. Analyzing GROVER’s learned representations, we observed that at the word level it primarily learns information related to their frequency, sequence content, and length. Some tokens are almost exclusively localized in repeats, while the vast majority are distributed over different genome elements. The model also learns context and lexical ambiguity. We can extract functional annotations from the genome, purely based on sequence representation to the trained model. We also show possible applications on a selection of different tasks, including prediction of DNA-protein binding. GROVER can be fine-tuned for multiple tasks with additional genomics data to study how DNA is encoding the regulation of genes or stability for example. This can be used to investigate individual predisposition to disease and treatment responses, which to a large extent is encoded in the patients’ general and somatic genomes. Therefore, DNA language models like GROVER have the potential to significantly advance genomics in general, cancer genomics, and somatic genome evolution in particular. Through its transparent design, learned representations and attention can be recovered, thus allowing insight into sequence dependent molecular biological mechanisms.

166

Clonal dynamics of TP53 mutations in male germline cells and the implications for cancer predisposition in offspring

Rashesh Sanghvi[2], Maria Torra I Benach[2], Matthew Neville[2], Andrew Lawson[2], Pantelis Nicola[2], Federico Abascal[2], Tetyana Bayzetinova[2], Joseph Christopher[2], Mette Jorgensen[1], Raheleh Rahbari[2]

1 Great Ormond Street Hospital NHS Foundation Trust, United Kingdom
2 Wellcome Sanger Institute, United Kingdom

Presenter: Rashesh Sanghvi

Heritable genetic variation and disorders stem from germline mutations. Understanding these mechanisms and selection pressures is key for understanding evolution and disease progression. Inherited cancers account for 5-10% of all the cases, impacting individuals across generations. Li-Fraumeni Syndrome (LFS), an autosomal-dominant cancer predisposition syndrome, results from pathogenic variants in TP53 gene, leading to high cancer risk across different life stages. The mutational landscape and tissue-specific cancer transformation causes in LFS individuals remain poorly understood.
Approximately 7-20% of TP53 germline mutations in individuals with LFS occur de novo in the parental germ cells, with the remainder following typical autosomal dominant inheritance. In this study, we analysed whole genome sequences from blood samples of nuclear families (parents and offspring) and adrenocortical tumour biopsies from the children diagnosed with LFS. We found 25% of cases were caused by de novo mutations in TP53, while the rest were inherited. We identified a median of 83 germline de novo mutations per affected child (ranging from 51 to 116 mutations) with parental age accounting for most of the inter-individual variation. Signatures SBS1 and SBS5 explain all of the germline de novo mutations in the children (with median contributions of 15% and 85% respectively). While, in the tumour samples we identified a median of 941 mutations which are explained by SBS1, SBS5 as well as APOBEC signatures. Further, we investigated the clonal dynamics of mutant TP53 in sperm of the fathers (age ranging from 34 to 62 years-old), using a deep targeted sequencing method. In the majority of sperm samples we did not observe any expansion of pathogenic TP53 except in one father, in whom we observed 2.5% of sperm had the TP53, p.V272M pathogenic variant.
This study aims to further investigate the variation in somatic mutation rate and genomic instability caused by germline TP53 mutations. Understanding the clonal dynamics of de novo TP53 mutations in the male germline will better inform recurrence risk in future offspring. Furthermore, by analysing tumour biopsies, it will be possible to ascertain the key evolutionary steps in the development of LFS-associated malignancy and develop targeted screening and therapeutic strategies.

167

Application of Nanopore sequencing for liquid biopsy analysis in children with cancer

Carolin Sauer[1], Nicholas Tovey[6], Debbie Hughes[3], Courtney Himsworth[5], Joanne Stockton[6], Paola Angelini[4], Marwane Bourdim[1], Reda Stankunaite[3], Claire Lynn[3], Marilena Nicolaidou[5], Michael Hubank[3], John Anderson[2], Andrew Beggs[6], Louis Chesler[3], Isidro Cortes Ciriano[1]

1 EMBL-EBI, United Kingdom
2 Great Ormond Street Hospital, United Kingdom
3 The Institute of Cancer Research, United Kingdom
4 The Royal Marsden Hospital, United Kingdom
5 University College London, United Kingdom
6 University of Birmingham, United Kingdom

Presenter: Carolin Sauer

Paediatric cancers are the leading cause of death in children post infancy in the Western world. Comprehensive high-throughput molecular profiling is essential to elucidate the molecular basis of evolving, treatment resistant disease and to more effectively guide clinical decision making. Access to high-quality tumour material for genomic profiling is a challenge in children where tissue biopsies are small. The analysis of cell-free DNA (cfDNA) from liquid biopsies for the detection of circulating tumour-derived DNA (ctDNA) offers a powerful, minimally invasive alternative to tumour profiling. However, at present, ctDNA analysis is limited in sensitivity, specificity, scalability, turnaround time and cost, hindering its implementation into standard clinical care.
Emerging Nanopore sequencing can report on native DNA in either whole-genome or focused format, and are rapid and scalable at low cost, making this platform highly attractive in the clinical setting. Here, we exploit Nanopore sequencing for the multi-modal analysis of cfDNA in paediatric cancer patients, including the analysis of cancer-specific mutations and epigenetic alterations.
Using matched liquid biopsies and tumour tissues obtained from patients with relapsing cancer enrolled in the UK Stratified Medicine Paediatrics study, we demonstrate the utility of Nanopore sequencing to detect clinically relevant somatic aberrations, such as ALK and MYCN amplifications from low volume blood draws. Overall, copy number aberrations detected using Nanopore sequencing were highly concordant with those detected using Illumina whole-genome sequencing. Using the ability of Nanopore sequencing to read out epigenetic modifications, we show novel methylation deconvolution algorithms to accurately specify tissue-of-origin and oncotype. Finally, integration of copy number and methylation data enables monitoring of disease burden, progression, and detection of disease relapse in longitudinal plasma samples. Together, our results suggest that Nanopore-based multi-modal liquid biopsy analysis of ctDNA may present a powerful tool to significantly improve treatment in children with cancer by facilitating early detection, accurate diagnosis, and efficient serial monitoring of disease progression.

168

Challenges of tumour-only variant calling from Amplicon-based sequencing

Irmi Sax[2], Sebastian Dintner[1], Rainer Claus[1], Matthias Schlesner[2]

1 University Hospital Augsburg, Germany
2 University of Augsburg, Germany

Presenter: Irmi Sax

Differentiation between germline and somatic variants is important for selecting proper treatment and medical therapies of patients. However, in some clinical setups there is no matched normal sample available or no sequencing of matched normal sample is performed to save cost and resources. This approach gives rise to the need of tumour-only pipelines that can distinguish germline and somatic variants from tumour tissue only.
Amplicon-based sequencing targets specific regions within the genome. By using target-specific primers only DNA segments of interest are amplified and sequenced.
Using the AmpliSeq for Illumina Focus Panel we captured sequencing information for 52 genes in 268 regions with a mean region size of 106bp covering in total 28kbp of the genome. We followed the GATK best practices for finding single nucleotide somatic variants in 16 patients with Mutect2 using a tumour and a matched normal. The resulting variants were filtered, and the remaining 23 somatic variants served as a validation set of high quality and high confidence.
In the next step the two tools UNMASC and PureCN were used to identify somatic variants from tumour only tissue. UNMASC has identified 10 of the somatic variants (sensitivity 0.43) in the validation set. Running UNMASC successfully was only possible after changing some hardcoded filter thresholds (e.g., minimal number of variants per sample/ segment) and setting those thresholds to values reflecting characteristics of Amplicon sequencing.
PureCN identified 9 variants from the validation set (sensitivity 0.39) after also including low AF variants. PureCN requires at least 20 variants per sample to compute purity and ploidy and to guess the somatic status of a variant. By filtering out low AF candidate variants PureCN was not able to calculate the somatic status in any of the samples as not enough variants per sample were available.
The limitations of the tools are mainly based on the characteristics of Amplicon Sequencing which is a small targeted region resulting in a small number of candidate somatic variants. In the case of small Amplicon Panels other strategies for identifying somatic variants from tumour-only tissue (e.g., variant characterisation with a scoring system) might achieve better results.

169

Rapid and data-driven generation of synthetic NGS Cancer Datasets with SYNGGEN

Riccardo Scandino[1], Federico Calabrese[1], Alessandro Romanel[1]

[1] University of Trento, Italy

Presenter: Riccardo Scandino

Advances in next-generation sequencing (NGS) technologies such as whole-exome sequencing (WES) and targeted sequencing (TS) have revolutionized cancer genomics and precision medicine. However, accurate interpretation of somatic genomics profiling results from NGS requires reliable computational tools. That's where synggen comes in - a powerful tool written in C programming language that enables researchers to rapidly generate realistic synthetic WES and TS datasets for benchmarking.
Synggen closely mimics real-life cancer sequencing scenarios utilizing non-cancer NGS sequencing files in BAM format fo generate reference models and by incorporating user-specified phased germline polymorphisms, complex allele-specific somatic copy number aberrations and point mutations, as well as the clonality of somatic events and overall tumor content of the sample.
To demonstrate the effectiveness of synggen we simulated two liquid biopsy cfDNA scenarios: cancer data at decreasing tumor content, and cancer data simulating temporal sampling from a patient with dynamic tumor sub-clones' populations.
Generating WES reference models using one control sample takes approximately 5 minutes with 4 cores, and 2.5 minutes with 16 cores. Generating a FASTQ file with 100 million reads using the same number of cores requires about 10 minutes and 4 minutes, respectively.

170

DNA methylation changes cause pervasive reorganization of CTCF binding and 3D genome structure in breast cancer cells

Julie Segueni[1], Daan Noordermeer[2], Joanne Edouard[1]

[1] I2BC - CNRS, France
[2] University Paris-Saclay, France

Presenter: Julie Segueni

Mammalian genomes adopt a functional 3D organization where enhancer-promoter interactions are constrained within Topologically Associating Domains (TADs). The CTCF insulator protein has a dual role in this process, with binding at promoters promoting the formation of enhancer-promoter loops (intra-TAD structure) and binding at TAD boundaries preventing the formation of inappropriate loops between neighboring domains. Importantly, perturbations of CTCF binding at specific sites in cancer cells can be caused by both changes to the DNA sequence (mutations) or DNA methylation changes (epi-mutations) (Segueni, Noordermeer 2022).
Following a systematic multi-omics approach in multiple breast cancer cell lines and controls, we have intersected genome-wide CTCF binding, DNA sequence, DNA methylation and its impact on 3D genome organization and gene activity. We first performed precisely-calibrated CTCF ChIP-seq experiments and found that a large number of sites are differentially bound, with a substantial fraction of differential CTCF binding peaks shared among cancer cell lines. Differential CTCF peaks can both be gained and lost and are often localized close to genes associated with breast cancer transformation. Our ongoing analysis indicates that DNA methylation changes, both gains and losses, induce this CTCF binding deregulation. Indeed, chemically induced hypomethylation in both breast cancer and control cell lines confirms the instructive nature of these epi-mutations.
Using high-resolution Hi-C, we assessed the impact of differential CTCF binding on chromatin structure, characterizing considerable 3D genome reorganization at gene loci with perturbed CTCF peaks. Unexpectedly, we find the most drastic examples of reorganization within TADs, at the level of enhancer-promoter loops. Our work thus identifies a pervasive DNA-methylation-guided reorganization of CTCF binding and intra-TAD structure, with a considerable recurrence between independently obtained breast cancer cell lines. Such recurrent patterns of epi-mutations can provide a mechanistic explanation for shared gene deregulation in cancers.

171

In silico testing of hypotheses for the effect of smoking on somatic evolution in the healthy human lung

Hugh Selway-Clarke[2], Kate Gowers[2], Vitor de Sousa Teixeira[2], Calum Gabbutt[1], Carlos Martinez Ruiz[2], Ahmed Alhendi[2], Benjamin Simons[3], Sam Janes[2], Nicholas McGranahan[2], Adam Pennycuick[2]

1 Institute of Cancer Research, United Kingdom
2 University College London, United Kingdom
3 University of Cambridge, United Kingdom

Presenter: Hugh Selway-Clarke

Recent single-cell genomic analysis of healthy lung tissue (Yoshida et al, Nature 2020) has shown remarkable intra-tissue heterogeneity in the degree of effect smoking has on mutational burden, as well as an expansion of less-mutated basal cell sub-populations after smoking cessation. These two findings suggest potential mechanisms for somatic evolution in the healthy human lung, which forms the backdrop for lung cancer formation. Here, we use computational modelling, based on a model of lung homeostasis previously verified by lineage tracing (Teixeira et al, eLife 2013), to assess the ability of these hypotheses to reproduce observations. Applying a Bayesian inference framework to simulations of basal lung cell populations over the course of patients' lifetimes, we find preliminary evidence for a protected sub-population of basal cells in the lung which are less affected by smoking. The simulations suggest that this protected sub-population, in combination with immune targeting of highly mutated cells being dampened during smoking, can best reproduce the unexpected dynamics seen in the data. With further testing and validation in epidemiological datasets, this mechanistic understanding will allow for better public engagement on smoking and streamline future research into the early detection and prevention of lung cancer.

172

RUVBL1 and RUVBL2 as novel druggable DNA damage response regulators in the N-Myc regulatory network in neuroblastoma

Joachim Tetteh Siaw[1], Arne Claeys[2], Wei-Yun Lai[2], Marcus Borenäs[2], Elien Hilgert[1], Sarah-Lee Bekaert[1], Frank Speleman[2], Kaat Durinck[1], Bengt Hallberg[2], Ruth Palmer[2], Jimmy Van den Eynden[2]

[1] Ghent University, Belgium
[2] University of Gothenburg, Sweden

Presenter: Joachim Tetteh Siaw

High-risk neuroblastoma (NB) accounts for about 50% of all cases. These tumours are characterized by MYCN amplification and high MYC gene expression and patients have a high relapse rate despite intensive therapies, hence the need for safer and more effective drugs. Strategies to develop inhibitors that directly target the MYC proteins have been elusive.
Based on in silico molecular signature score and network analyses, we identified RUVBL2 as a key interactor of MYC. Kaplan-Meier survival and multivariate Cox regression analyses using public NB datasets demonstrated that expression of RUVBL2 and its interaction partner RUVBL1 are strong and independent predictors for both overall and event-free survival in NB patients. Using different types of NB cell lines, we experimentally demonstrated that transient knockdown of RUVBL1/2 or pharmacological inhibition using CB-6644 resulted in cell cycle arrest, cell growth arrest and a DNA Damage Response (DDR) through regulation of ATR and ATM. Additionally we confirmed that RUVBL1/2 transcriptionally regulate MYCN and MYC.
Our work demonstrates that RUVBL1 and RUVBL2 are novel regulators of the DDR, with therapeutic and independent prognostic potential in high-risk NB.

173

Unraveling the spatial architecture of Cancer Hallmarks

Mustafa Sibai[4], Sergi Cervilla[4], Eva Musulen[4], Daniela Grases[4], Rossana Lazcano[8], Chia-Kuei Mo[6], Verónica Dávalos[4], Arola Fortian[3], Margarita Romeo[3], Adria Bernat[3], Collin Tokheim[2], Enrique Grande[5], Francisco X Real[7], Jordi Barretina[3], Alexander J Lazar[8], Li Ding[6], Manel Esteller[4], Matthew H Bailey[1], Eduard Porta-Pardo[4]

[1] Brigham Young University, United States of America
[2] Dana-Farber Cancer Institute, United States of America
[3] Institut de Recerca Germans Trias i Pujol (IGTP), Spain
[4] Josep Carreras Leukaemia Research Institute (IJC), Spain
[5] MD Anderson Cancer Center Madrid, Spain
[6] Siteman Cancer Center, Washington University in St. Louis, United States of America
[7] Spanish National Cancer Research Centre (CNIO), Spain
[8] The University of Texas MD Anderson Cancer Center, United States of America

Presenter: Mustafa Sibai

Tumors are complex ecosystems in which different cell types coexist and interact to promote malignancy. The concept of Cancer Hallmarks outlined by Hanahan and Weinberg distills this complexity into a set of underlying principles (phenotypes) that govern tumor growth. Here, we exploit this abstraction to explore the physical distribution of 13 Cancer Hallmarks across 63 primary untreated tumors from 10 cancer types using spatial transcriptomics.
We show that Hallmark activity is spatially organized–with 7 out of 13 Hallmarks consistently more active in cancer cells than within the non-cancerous tumor microenvironment (TME). The opposite is true for the remaining six Hallmarks. We further show that Hallmark activity is spatially organized even within its associated compartment (cancer cells or TME).
Additionally, we discovered that while genomic distance between tumor subclones correlates with differences in Cancer Hallmark activity, clone-Hallmark specialization occurs only in 3% of pairs of all tumor subclones. Conversely, over 7% of pairs of all tumor subclones exhibited genomic divergence and phenotypic convergence (i.e. genomic divergence without any difference in Hallmark activity).
Finally, we demonstrate interdependent relationships between Cancer Hallmarks at the junctions of TME and cancer compartments. In particular, we show that Hallmarks of the TME collaborate for configuring the spatial phenotypic landscape of cancer cells. By contrast, cancer cells tend to phenotypically diverge based on their relative distance to the TME. Our model suggests that the core of cancer cells (the area that is further from the TME) engages in “Resisting Cell Death”, “Enabling Replicative Immortality”, “Genome Instability and Mutation” and “Deregulating Cellular Energetics". On the other hand, cancer cells at the interface with the TME mostly engage in “Evading Growth Suppressors” and “Senescence”, followed by “Non-Mutational Epigenetic Reprogramming” being closer to the core.
In conclusion, including the spatial dimension, particularly through the lens of Cancer Hallmarks, can improve our high-level understanding of tumor ecology.

174

Identification of genomic and transcriptomic aberrations of clinical and biological relevance in pediatric T-ALL: data from a tertiary care centre of India

Minu Singh[1], Pankaj Sharma[1], Prateek Bhatia[1], Amita Trehan[1], Rozy Thakur[1], Sreejesh Sreedharanunni[1]

[1] Postgraduate Institute of Medical Education & Research, India

Presenter: Minu Singh

T-cell acute lymphoblastic leukemia (T-ALL) is a genetically heterogeneous disease with poor prognosis and inferior outcome. To better understand the disease biology, we comprehensively analysed the transcriptome (n=25) including gene fusion and aberrant gene expression along with whole exome sequencing (n=17) in pediatric T-ALL patients from an Indian cohort and correlated the findings with clinical characteristics and disease outcome.

Based on RNA sequencing data we identified 16 cases (64%) with fusion/rearrangement of genes. The most common fusion gene noted was STIL::TAL1 in 4 patients (16%) followed by NUP21::ABL1, TCF7::SPI1, LMO1::RIC3, DIAPH1::JAK2, SETD2::CCDC12, and RCBTB2::LPAR6, one (4%) in each case. Further, we noted 3 novel fusion in our cohort viz. ETV6::HDAC8, MED12::IRF2BPL and LYZ::SLC2A4; in one case (4%) each. We also found KMT2A and MLL10 rearrangement in 2 (8%) and 1 case (4%) each, respectively. Aberrant expression was noted in RAG2(80%), RAG1(64%), MYCN(52%), NKX3-1(52%), NKX3-2 (32%), TLX3(28%), LMO1(20%) and MYB(16%) genes. Based on exome sequencing data, the most commonly mutated gene noted was NOTCH1(35%) followed by WT1(23%), FBXW7(12%), KRAS(12%), PHF6(12%) and JAK3(12%). Further, we found that 88.2% of cases showed deletion of CDKN2A/B/MTAP genes. NOTCH1 mutation was noted to be significantly associated with TLX3 over-expression (p= 0.04). When examined for correlation with clinical outcome, cases with STIL::TAL1 fusion and NOTCH1, FBXW7 mutations showed better event frees survival (EFS) and lower relapse rate (RR), however the data was not significant. Patients with RAG2 overexpression showed better EFS (p=0.01) and overall survival (p=0.01) for a median follow up of 22 months. Further, cases with overexpression of MYB was noted to have poor EFS (p=0.041) and high RR (p=0.045).

In our pilot study of comprehensive genomics in T-ALL patients, we found that each of the 25 cases harbored at least one major genetic abnormality that may be involved in driving leukemogenesis in these patients. Despite the major limitaion of small cohort size of the study, we demonstrated the frequencies of genetic alterations in Indian cohort of pediatric T-ALL patients and is a salient addtion to current genomics data sets available in T-ALL.

175

Genome-wide somatic mutation analysis of formalin-fixed paraffin-embedded sinonasal adenocarcinomas

Lauri Sipilä[2], Riku Katainen[2], Mervi Aavikko[2], Janne Ravantti[2], Iikki Donner[2], Rainer Lehtonen[2], Ilmo Leivo[3], Henrik Wolff[1], Reetta Holmila[1], Kirsti Husgafvel-Pursiainen[1], Lauri Aaltonen[2]

1 Institute of Occupational Health, Finland
2 University of Helsinki, Finland
3 University of Turku, Finland

Presenter: Lauri Sipilä

Ba**ckground:** Sinonasal adenocarcinoma (SNAC) is a rare cancer, highly significantly enriched in individuals with occupational exposure to wood dust. Studies have discovered histological subtype segregation trends by the type of wood dust, with hardwood dust driving tumor development towards intestinal-type adenocarcinoma (ITAC), and softwood dust toward non-intestinal-type adenocarcinoma (non-ITAC). While chronic inflammation appears to have a role especially in the development of ITACs, the mechanisms behind the tumorigenic effects of wood dust remain largely unknown.
Methods: We whole-genome sequenced formalin-fixed paraffin-embedded (FFPE) SNAC samples from ten wood dust-exposed and six non-exposed individuals, with partial tobacco exposure data. Sequences were analyzed for the presence of single-nucleotide mutation, doublet mutation, and copy number (CN) variation signatures matching COSMIC database signatures. Driver mutations and CN variant regions were characterized.
Results: Mutation burden as measured by total single-nucleotide variant (SNV) count was higher in wood dust-exposed samples (p=0.016). Reactive oxygen species (ROS) damage-related signature mutations were almost exclusively identified in ITAC subtype samples (p=0.00055). Tobacco smoke mutation signatures were observed in samples with tobacco exposure or missing information, but not in non-exposed samples. CN signature analysis identified tetraploidy signature CN2 and homologous recombination deficiency signature CN17, with borderline significant differences in event count by subtype. CN variation included recurrent gains in cancer gene census genes TERT, SDHA, RAC1, ETV1, PCM1, and MYC.
Conclusions: Our analysis identified distinct mutational characteristics by SNAC subtype. This suggests that mutation signature data may eventually become useful for documentation of occupation-related cancer, while exact mechanisms behind wood dust-driven carcinogenesis remain elusive. The presence of homologous recombination deficiency signatures implies an opportunity for treatment, but due to small sample size further studies are needed.

176

Identification of a mutational signature of dietary acrylamide in renal cancer genomes

Kim Smits[2], François Virard[1], Bérénice Chavanel[1], Maria Zhivagui[3], Frederick Beland[4], Michael Korenjak[1], Leo Schouten[2], Jiri Zavadil[1]

1 International Agency for Research on Cancer, France
2 Maastricht University, The Netherlands
3 University of California San Diego, United States of America
4 US Food and Drug Administration, National Center for Toxicological Research, United States of America

Presenter: Jiri Zavadil

The chemical acrylamide (ACR) is a probable human and established rodent carcinogen found in heated starchy foods and in the tobacco smoke. Epidemiological studies suggest elevated non-significant risk of clear-cell renal cell carcinoma (ccRCCs) due to the dietary ACR intake. We previously observed in silico that 70% of 111 ccRCC PCAWG genomes carry the putative mutational signature of ACR (PMID 30846532). Here we aimed to establish a molecular link between the ACR intake and ccRCC by analyzing mutational signatures in a unique tumor set with well-documented dietary ACR exposure history.
Within the prospective Netherlands Cohort Study on Diet and Cancer (NLCS) involving 120,852 subjects of whom 480 developed RCC, we selected a set of never-smokers with ccRCC and a history of high versus low (10 cases/group) dietary ACR intake, as assessed by a food frequency questionnaire reflecting chemical analysis of relevant Dutch foods (PMID 18469268). DNA isolated from tumor and non-tumor FFPE tissue pairs was whole-genome sequenced, and somatic mutations were processed by the SigProfilerExtractor tool to identify the de novo as well as known COSMIC mutational signatures. The Mutational Signature Analysis (MSA) tool was used for optimized computational per-sample signature assignment.
The NLCS ccRCC genomes harbored the COSMIC signatures SBS1, SBS5, SBS40 in all samples, in proportions irrespective of the dietary ACR exposure groups. In contrast, the MSA signature assignment showed relative enrichment of the previously reported ACR mutational signature in the high-exposure (6 of 10 [60%] cases) compared to the low-exposure patient group (3 of 10 [30%] cases).
Our results indicate for the first time ACR-related genome-scale mutagenesis in ccRCC with documented dietary ACR exposure history, and its potential contribution to tumor formation. These findings, to be corroborated in an expanded patient set, may have important implications as evidence supporting the reduction of human exposure to acrylamide and the related cancer prevention measures.
Funding: WCRF International, Grant No. SG_2020_089

177

Inference of pathway functional interactions in pediatric cancer

Anastasia Spinou[1], Richard Gremmen[1], Puck Veen[1], Joanna von Berg[1], Jarno Drost[2], Patrick Kemmeren[1]

1 Princess Máxima Center for Pediatric Oncology, The Netherlands
2 Princess Maxima Center for Pediatric Oncology; Oncode Institute, The Netherlands

Presenter: Anastasia Spinou

Despite the constant improvement of treatments for pediatric cancer and an overall 5-year survival spanning up to 80% or 90% (in high-income countries), cancer remains the leading disease-related cause of death in children. Side effects of existing therapies and rare cases, that are challenging to treat, highlight the necessity for new targeted therapies. Developing targeted therapies requires a deep understanding of each pediatric cancer type and unraveling the complexities of their genomic landscape. Genetic interactions, combinations of two mutated genes that lead to unexpected cell phenotypes such as synthetic sickness/lethality, can aid in identifying potential therapeutic targets as well as understanding tumor biology. These functional interactions can be beneficial to investigate not only between genes but also between pathways to uncover synthetic sickness/lethality and advantageous pathway relationships. Investigating pathway functional interactions is relevant, because, in cancer, the mutated gene(s) of a pathway can modify the pathway activity which can lead to a cancerous phenotype. But, also, more than one modified pathway can lead to a cancerous phenotype due to pathways performing related functions. Here, we infer pathway functional interactions by applying a statistical test (WeSME/rediscover) on patient pathway mutation profiles that aggregate information from individual gene mutations. We use a selection of pathways from the Reactome database to associate each mutated gene with its corresponding pathway. Pathways group the mutated genes of each cancer type into a functional unit and increase power for uncovering relationships containing genes of low mutation frequency which otherwise would not be observed by testing gene relationships only. These relationships between pathways may be synthetic sick/lethal or advantageous and will subsequently be functionally validated using pre-clinical models of the respective tumor entity. Overall, this approach provides a better understanding of genetic interactions, finds novel relationships between genes, and uncovers pathway relationships that could potentially assist in finding new therapeutic targets.

178

Quantifying fitness effects of structural variants with SimChA

Adam Streck[2], Cody Duncan[1], Felix Schifferdecker[1], Roland Schwarz[3]

[1] Institute for Computational Cancer Biology, Centre for Integrated Oncology (Uniklinik Köln), Germany
[2] Institute for Computational Cancer Biology, Centre for Integrated Oncology (Uniklinik Köln), Berlin Institute for Medical Systems Biology (Max Delbrück), Germany
[3] Institute for Computational Cancer Biology, Centre for Integrated Oncology (Uniklinik Köln), BIFOLD (Berlin Institute for the Foundations of Learning, Germany

Presenter: Cody Duncan

Over the course of cancer evolution, cells accumulate new mutations, leading to intra-tumour heterogeneity. These subclones compete with each other for survival, selecting for advantageous traits and punishing negative ones. But cancer is not a homogeneous disease. Different genotypes thrive in different micro-environments within a tumour, leading to the spectrum of different cancer subtypes observed in patients. Understanding how cancer's evolutionary path impacts subclonal fitness is one of the main challenges in cancer biology: can we build a model of cancer cell fitness from first principles and use it to simulate realistic cancer genomes?
In our talk, I will present our recent efforts to quantify fitness from three core physical properties of the genome: cellular stress, TSG/OG score, and essentiality. We test the robustness of this fitness score by demonstrating that the majority of samples from PCAWG have positive fitness relative to the healthy genome. With the scoring method, we then developed a new simulation tool - SimChA, which simulates the effects of a large number of structural variants on genomes. Using an MCMC approach, we can fit genotypes to a target fitness sampled from the fitness distribution of a dataset, and consequently compare the average copy-number profile of the observed data to the simulated dataset. We show that fitness is a necessary component to reproducing the patterns of mutation we observed in data, and that the MCMC algorithm selects for the location of mutation, rather than the event itself. Finally, we evaluate the fitness dynamics of genotypes over the course of cancer evolution and investigate cancer-type specific fitness differences.

179

Nanopore sequencing reveals structural features of somatic and germline retrotransposon insertions

Päivi Sulo[5], Tatiana Cajuso[1], Aurora Taira[5], Heli Kuisma[5], Niko Välimäki[5], Anna Lepistö[2], Laura Renkonen-Sinisalo[2], Selja Koskensalo[2], Toni Seppälä[2], Ari Ristimäki[2], Teijo Kuopio[4], Jan Böhm[3], Jukka-Pekka Mecklin[3], Annukka Pasanen[6], Oskari Heikinheimo[6], Ralf Bützow[2], Kathleen Burns[1], Lauri Aaltonen[5], Kimmo Palin[5]

1 Dana-Farber Cancer Institute, United States of America
2 Helsinki University Central Hospital, University of Helsinki, Finland
3 Jyväskylä Central Hospital, University of Jyväskylä, Finland
4 Univeristy of Jyväskylä, Finland
5 University of Helsinki, Finland
6 University of Helsinki and Helsinki University Hospital, Finland

Presenter: Päivi Sulo

Retrotransposons are DNA sequences with the ability to create copies of themselves and subsequently integrate the copies into novel loci within the genome. Retrotransposons in the human genome predominantly lack the capability for transposition. The elements that remain transposition competent are mainly epigenetically repressed in adult tissues. However, somatic retrotransposition has been detected in multiple cancer types. Furthermore, it has been reported that somatic retrotransposition can contribute to tumorigenesis and is associated with poor survival in colorectal cancer.
Novel retrotransposon insertions have been challenging to study with short-read sequencing technologies. However, the rise of long-read technologies has enabled the research of structural features of the insertions. To detect and annotate both somatic and germline retrotransposon sequences, we created a novel pipeline, Transposon Detection in Oxford Nanopore Sequencing data (TraDetIONS). By applying TraDetIONS to a dataset of 62 colorectal cancers and 104 uterine leiomyomas, we were able to detect 1539 somatic insertions in the colorectal samples. Notably, uterine leiomyomas, benign neoplasms originating from mesenchymal tissue, exhibited no somatic insertions.
Through a comparative analysis of somatic and germline insertions, differences in terms of transposon type, insertion length, and preference for target sites were revealed. This suggests that the diverging conditions of tumor and germline cells impact the process of transposition. The research into the differences between somatic and germline transposition processes can illuminate the mechanisms in the activation of somatic retrotransposons in tumorigenesis.

180

Enhanced in vitro culture of leukemic cells: insights from collagen scaffolds and carboxymethyl cellulose-polyethylene glycol gel

Hana Svozilova[3], Lucy Vojtová[1], Jana Dorazilová[1], Jana Bruknerová[4], Veronika Poláková[1], Lenka Radová[4], Michael Doubek[4], Karla Plevová[4], Sarka Pospisilova[2]

[1] CEITEC, Brno University of Technology, Czech Republic
[2] CEITEC, Masaryk University & University Hospital Brno, Czech Republic
[3] Masaryk University, Czech Republic
[4] University Hospital Brno, FCEITEC, Masaryk University Brno, Czech Republic

Presenter: Hana Svozilova

Background: Studying chronic lymphocytic leukemia (CLL) in vitro is challenging due to its complexity and dependency of CLL cells on their microenvironment. An integral part of the natural CLL microenvironment is the three-dimensional (3D) spatial organization, which facilitates frequent cell-to-cell and cell-to-matrix contacts. Therefore, we aimed to mimic the natural tissue architecture by implementing a 3D in vitro culture. We hypothesized that compared to conventional culture, the additional dimension and increased cell-to-matrix contacts would enhance the prosurvival stimuli in CLL cells.
Methods: We cultured CLL cells in two materials: (i) collagen scaffolds, or (ii) gel composed of carboxymethyl cellulose and polyethylene glycol (CMC-PEG). We assessed cell distribution, morphology, and viability via microscopy, and measured the metabolic activity by AlamarBlue assay. Gene expression (MYC, VCAM1, MCL1, CXCR4, CCL4) was analyzed using qPCR to understand the effects of novel culture approaches on adhesion, apoptosis, and intercellular interactions of CLL cells co-cultured with bone marrow stromal cells (BMSCs) in 3D.
Results: The materials facilitated cell-to-cell and cell-to-matrix interactions due to scaffold structure and aggregate generation. CLL cells in CMC-PEG displayed similar or higher metabolic activity than in conventional culture. Compared to conventional culture, we observed the following effects: (i) a lower expression of VCAM1 in both materials, (ii) a higher expression of CCL4 in collagen scaffolds, and (iii) a lower expression of CXCR4 and MCL1 (proapoptotic transcript variant 2) in collagen scaffolds, while it was higher in the CMC-PEG gel.
Conclusion: Our findings suggest that the introduction of material into the in vitro 3D culture impacts CLL cells' apoptosis and interactions. While the expression of proapoptotic MCL1 is either suppressed (collagen) or elevated (CMC-PEG) in 3D materials, culture in both CMC-PEG and collagen can enhance the expression of certain genes associated with CLL-BMSCs interaction.
Acknowledgments: Supported by MEYS CZ MUNI/A/1224/2022, MH CZ RVO FNBr 65269705, and NPO Programme EXCELES LX22NPO5102. Core facility support by CELLIM and CFG funded by MEYS CZ LM2018129, and NCMG LM2023067, respectively.

181

Measuring the interplay between chromosomal instability and whole-genome doubling in human cancer

Dalil Taher[1], Mathieu Parmentier[1], Maxime Tarabichi[1]

[1] Université Libre de Bruxelles, Belgium

Presenter: Dalil Taher

Whole genome doubling (WGD) is a prevalent macroevolutionary event across cancer types. It provides a mechanism to escape Mullers' ratchet. Also in cell lines, WGD was linked to an increased tolerance to aneuploidies leading to chromosomal instability. This would mean that WGD provides a diverse substrate for selection with large leaps in the evolutionary landscape. To date, WGD has often been inferred from sequencing data, which is mathematically underdetermined and was shown to be incorrect across cancer types. Moreover sequencing-based estimates make the assumption that only smaller-scale copy-number events can be subclonal while the ploidy is always clonal, which we know to be untrue. Here, to get a cross-cancer picture of heterogeneity at the ploidy level, the frequency of subclonal WGD and its relationship to the tissue phenotype, we performed ploidy assessment of isolated nuclei through FISH analysis, microscopy imaging on isolated nuclei, as well as FACS followed by a single-cell whole-genome sequencing. We have benchmarked a range of nuclei extraction protocols to achieve precise nucleus extraction from delicate tissues while also preserving potential nucleus heterogeneity. Our aim is to ensure the accurate representation of nuclei populations, allowing us to delve into the intricacies of cellular composition within tissues. We are currently engaged in benchmarking various ploidy validation approaches, exploring a spectrum of techniques including gold standard technics like FISH (Fluorescence In Situ Hybridization), FACS (Fluorescence Activated Cell Sorting), but also image cytometry, histochemistry, and single-cell DNA sequencing (scDNA-seq) on FACS sorted nuclei. We identified a small fraction of ovarian and colorectal cancer and a large fraction of sarcomas with subclonal WGD, for which we separately sorted single-nuclei from the different ploidy populations for single-cell sequencing. Thanks to our setup, we will reconstruct phylogenetic relationships between the different ploidy populations using copy-number events as genetic barcodes. This will allow us to directly measure the impact of ongoing WGD on chromosomal instability in vivo across human cancers. We will also train neural networks on corresponding FFPE H&E images to infer the presence of WGD across cancer types.

182

Epigenetics meets metabolomics: studying the link between methylation and metabolism in BRAF mutated colorectal cancer

Kristiina Rajamäki[3], Aurora Taira[3], Kimmo Palin[3], Riku Katainen[3], Lauri Aaltonen[3], Toni Seppälä[1], Ari Ristimäki[1], Selja Koskensalo[1], Laura Renkonen-Sinisalo[1], Anna Lepistö[1], Jukka-Pekka Mecklin[2]

1 Helsinki University Central Hospital, University of Helsinki, Finland
2 Jyväskylä Central Hospital, University of Jyväskylä, Finland
3 University of Helsinki, Finland

Presenter: Aurora Taira

Somatic missense mutations of the BRAF gene occur in a variety of human cancers including colon cancer (CRC). BRAF is part of the RAS-RAF-MEK-ERK-MAP signaling pathway which regulates cellular responses to growth signals. Mutations at codon 600 (V600E) lead to overactivation of the signaling pathway altering the regulation of essential processes such as cell proliferation and cell death. BRAF mutated tumors have high DNA methylation levels. In general, aberrant methylation is a hallmark of cancer, disturbs the regulation of a variety of genes, and predisposes cells to malignant development. Cancer cells are known to acquire different metabolic adaptations in order to maintain malignant properties and sustained growth. Multiple studies have analyzed metabolite levels from blood or stool of CRC patients. However, studies on metabolite levels in tumor tissue samples have remained low, especially when it comes to studies linking metabolic changes to driver gene mutations in CRC. Tumors harboring gene mutations affecting the TCA cycle, such as FH mutated uterine leimyomas and SDH deficient small gastrointestinal tumors, suggest that epigenetic patterns and metabolic changes are linked: metabolites can disturb the activity of epigenetic regulators and may lead to accumulation of methylation. The aim of this project is to characterize epigenetic and metabolic patterns in BRAF mutated CRCs and to study the potential link between these factors. We limit our analysis to microsatellite stable tumors. Microsatellite instability (MSI) has been linked with methylation and thus methylation changes induced by MSI and BRAF mutations are difficult to separate from heterogenous sample sets. We have performed a global metabolite analysis for 24 BRAF mutated and 24 wild type tumors and obtained global methylation levels from 7 BRAF mutated and 26 wild type tumors together with 7 normal colon tissue samples using whole genome bisulfite sequencing data. In addition, RNA-sequencing data and mutational signatures from another set of tumors will be analyzed. Preliminary results indicate differential methylation at GLIS1 binding sites between BRAF mutated and wild type tumors. Moreover, BRAF mutated tumors showed low vitamin C levels. Vitamin C is a known cofactor of TET enzymes, which play a key role in regulating DNA demethylation.

183

An update: COSMIC - Catalogue Of Somatic Mutations in Cancer

Jonathan Teague[1], Zbyslaw Sondka[1], Nidhi Bindal Dhir[1], Sari Ward[1], Ilaria Fasanella[1], Susan Haller[1], Karen McLaren[1], Madiha Ahmed[1], Joanna Argasinska[1], David Beare[1], Denise Carvalho-Silva[1], Manpreet Singh Chawla[1], Stephen Duke[1], Balazs Hetenyi[1], Leonie Hodges[1], Alex Holmes[1], Stephen Jupe[1], Rachel Lyne[1], Madhumita Madhumita[1], Thomas Maurel[1], Helen Schuilenburg[1], Sumodh Nair[1], Avirup Guha Neogi[1], Helder Pedro[1], Amaia Sangrador-Vegas[1], Zoe Sheard[1], Michael Starkey[1], Jennifer Wilding[1], Siew-Yit Yong[1]

[1] Wellcome Sanger Institute, United Kingdom

Presenter: Jonathan Teague

As COSMIC approaches its 20th anniversary it has become an established and sustainable bioinformatics repository of somatic mutations in human cancer. As of release v98, COSMIC has manually curated nearly 30,000 scientific publications with just over 1.5 million cancer samples and over 20 million mutations (5 million protein-coding). Base-pair substitutions, small insertions/deletions, gene fusions, structural variation and copy number changes are captured along with expression and methylation data. Extensive sample metadata is also curated, providing in-depth phenotype information allowing the linking of tumour types to somatic mutations and drug response. Curated mutations are accessible through downloadable files and websites.

To further aid researchers, COSMIC has developed a series of accompanying resources with extra levels of information. The first resource to be established was the **Cancer Gene Census (CGC)**, which is a listing of cancer genes with supporting evidence from the scientific literature. In 2017, the CGC was amended with the **Hallmarks of Cancer**, which provides additional manual curation describing how CGC genes functionally contribute to cancer development, based around 10 cancer hallmarks. At the mutation level, the **Cancer Mutation Census (CMC)** classifies and tiers each mutation into whether it is likely to be a driver or passenger event. Recently, **Actionability** has been added to the suite of resources that catalogues drugs in clinical development or approved for treating patients with a defined cancer type and variant status. **COSMIC-3D** allows mutations to be scrutinised at the protein level, while the **Mutational Signatures** reference set aids the identification of mutational processes that drive cancer development.

Currently, COSMIC is focusing on modernising its IT infrastructure. Primarily, this will enable the streamlining of our curation systems to keep pace with the increased generation of genomic data. Furthermore, this will allow us to refine our existing resources and develop new features in collaboration with the scientific community. In addition, we are also engaging with the community to develop data reporting and processing standards in somatic oncology and tools to allow such standards to be adopted.

184

HDAC4 targeting in FBXW7 mutated CRC re-sensitizes cells to Oxaliplatin treatment

Vanessa Tolotto[1], Monica Colitti[1], Eros Di Giorgio[1], Claudio Brancolini[1]

[1] University of Udine, Italy

Presenter: Vanessa Tolotto

Colorectal cancer (CRC) is one of the most common cancers and adjuvant therapy including Oxaliplatin (OxPt) is considered the standard for advanced CRC. 17% of CRC patients bear loss-of-function (LOF) mutations in FBXW7, a tumour suppressor gene that encodes for the substrate recognition component of an SCF ubiquitin ligase complex. FBXW7 LOF induces resistance to chemotherapeutics as OxPt, while the treatment with Vorinostat, a pan HDAC inhibitor, sensitizes FBXW7 mutated cancer cells to chemotherapy.
We have recently described the role played by HDAC4, a class IIa HDAC, in supervising the assembly of end-resection complex involved in homologous recombination repair (HR). In cells defective for HR repair, HDAC4 is proteasomal degraded; however, the detailed mechanism of HDAC4 degradation is unknown.
HDAC4 is degraded at the protein level in HR impaired cells. This process requires the phosphorylation by GSK3β at specific sites (S298 and S302). We identified FBXW7 as the E3 ligase involved in HDAC4 degradation by means of an in silico screening of SILAC experiments and with an in vitro secondary screening. FBXW7 promotes K48 polyubiquitination of HDAC4, as demonstrated by CoIP and in vitro ubiquitylation assay. Importantly, HDAC4 and FBXW7 anti-correlate in tissue-microarray of grade 2-3 CRC biopsies and HDAC4 expression is higher in cancer samples in respect to matched healthy tissues. CRC cells and patients' derived spheroids defective in FBXW7 display resistance to OxPt. We confirmed that FBXW7 -/- HCT-116 are resistant to OxPt and OxPt does not trigger HDAC4 degradation. HDAC4 depletion by siRNAs or by the treatment with a specific PROTAC compound restores OxPt sensitivity in these cells, leading to the induction of apoptosis.
With our work we clarified the mechanism of HDAC4 degradation and our preliminary data identify HDAC4 as a druggable target to overcome OxPt resistance in FBXW7 mutated colorectal cancer, towards a precision cancer medicine approach.

185

Epigenome-instructed pan-cancer discovery of non-coding cancer drivers

Marketa Tomkova[2], Jakub Tomek[2], Julie Chow[1], John McPherson[1], David Segal[1], Fereydoun Hormozdiari[1]

1 University of California Davis, United States of America
2 University of Oxford, United Kingdom

Presenter: Marketa Tomkova

Identifying DNA mutations that drive cancer is a major goal in cancer research, critical for precision oncology and understanding tumour development. While pathogenic mutations in the protein-coding genome have been extensively characterized, the discovery of cancer drivers in the non-coding genome and distant enhancers has proved extremely challenging, with only a handful of well-established non-coding drivers found so far. The challenges of reliable identification of cancer driver mutations in the non-coding genome consist of (i) a 50 times larger size of the non-coding vs. coding genome, (ii) modelling the background mutation rate and predicting the functional effect of the non-coding mutations, (iii) often unknown and potentially tissue-specific function of the non-coding regions. Here, we address these challenges by multi-omics integration of tissue-matched genomics, epigenomics, and transcriptomics data.
We developed Dr.Nod, a computational framework for **D**iscovery of **r**egulatory **NO**n-coding **D**river mutations that are associated with dysregulated gene expression, using tissue-matched enhancer-gene annotations. Applying Dr.Nod to the PCAWG dataset revealed a 4.4-fold enrichment of known cancer driver genes in the target regions of the candidate non-coding regulatory driver mutations. In particular, oncogenes are enriched in the upregulated targets, and, to a lesser extent, tumour suppressor genes are enriched in the downregulated targets. Enrichment of both gain and loss of transcription factor binding sites is observed in the regulatory driver mutations. Interestingly, more than half of the detected non-coding regulatory driver mutations are over 20 kbp distant from their target gene, demonstrating the importance of long-range chromosomal interactions. Moreover, our results show the importance of tissue-matched enhancer-gene maps, functional impact of mutations, and a complex background mutagenesis model for the prediction of non-coding regulatory drivers.
In summary, our study suggests a widespread pan-cancer mechanism of oncogene upregulation by non-coding mutations in their cis-regulatory elements.

Marketa Tomkova is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

186

Clonal evolution trajectories of mature and immature teratomas

Joseph Christopher[4], Maria Torra I Benach[4], Rashesh Sanghvi[4], Thomas R. W. Oliver[4], Jose Espejo Valle-Inclan[1], Lia Chappell[4], hyunchul jung[4], Mimy Pham[4], Matthew Murray[3], Nicholas Coleman[3], Mette Jorgensen[2], Isidro Cortes Ciriano[1], Peter Campbell[4], Raheleh Rahbari[4]

1 EMBL-EBI, United Kingdom
2 Great Ormond Street Hospital NHS Foundation Trust, United Kingdom
3 University of Cambridge, United Kingdom
4 Wellcome Sanger Institute, United Kingdom

Presenter: Maria Torra I Benach

Germ cell tumours (GCTs) are highly heterogeneous tumours derived from primordial germ cells. Amongst these, teratomas emerge as a distinctive GCT subtype, encompassing diverse cell types from all three primordial germ layers. Teratomas can be categorized into mature and immature types. Whilst mature teratomas are composed of well-differentiated tissues and are mostly benign, immature teratomas present embryonal-like features, entailing an indeterminate malignant potential and a lack of precise prognostic tools for clinical risk assessment. Thus, highlighting the need to fully characterise the complex cellular composition of teratomas at the histological level and explore the molecular underpinnings of malignancy. Here we report a comprehensive multi-omics study, coupling bulk sequencing techniques with the precision of laser-captured microscopy to analyse mature and immature teratomas at unprecedented histological and spatial resolution. Through the generation of 314 genomes, 394 transcriptomes and 68 methylomes, we discern distinctive attributes of gonadal and non-gonadal teratomas. Moreover, we elucidate the phylogenetic relationships between diverse histological units within these tumours and unravel their characteristic mutational landscapes and copy number profiles. Given the developmental origin of teratomas, we undertake a comparative exploration of their transcriptomic and methylation profiles, providing a deeper insight into their underlying molecular signatures. In aggregate, our findings collectively enhance the comprehension of tumour evolution, elucidating disparities that demarcate immature from mature teratomas.

187

Stepwise transcriptional progression of myeloid leukaemia associated with Down syndrome

Mi Trinh[4], Matthew Young[4], Conor Parks[4], Elena Prigmore[4], Agnes Oszlanczi[4], Laura Jardine[2], Konstantin Schuschel[3], Jan-Henning Klusmann[3], Jack Bartram[1], Sam Behjati[4]

1 Great Ormond Street Hospital for Children, United Kingdom
2 Newcastle University, United Kingdom
3 University Hospital Frankfurt, Germany
4 Wellcome Sanger Institute, United Kingdom

Presenter: Mi Trinh

Down syndrome (DS) is a genetic condition caused by the presence of an extra copy of chromosome 21 (trisomy 21 or T21). Children born with DS have a substantially higher risk of developing childhood leukaemia. Approximately 30% of DS infants develop a preleukemic condition called transient abnormal myelopoiesis (TAM). TAM is strictly associated with GATA1 truncating mutations in a T21 background. Peculiarly, most TAM cases spontaneously resolve within the first few months of life. However, by the age of five, approximately 20% of cases progress to megakaryoblastic/erythroid leukaemia, known as myeloid leukaemia associated with DS (MLDS). Current literature suggests that T21 causes perturbed myelopoiesis within the developing foetal liver, and the GATA1 mutation then further predisposes the individual to leukaemogenesis (TAM).
We aim to further characterise the mechanism underlying this multi-step development of leukaemogenesis in DS: from diploid to T21, to the acquisition of GATA1 mutations in TAM, and finally, acquisition of additional transforming mutations in MLDS. Here, we directly interrogate these transcriptional changes at single-cell resolution, using primary human samples. We generated whole genome and single-cell RNA sequencing data from T21 human foetal livers, and bone marrow aspirates or peripheral blood from children with MLDS or TAM. We showed that T21 haematopoietic stem cells are enriched for the transcriptional programme of megakaryocyte-erythrocyte progenitors (MEPs), which may underpin the expansion of MEPs observed in T21 foetal liver. TAM/MLDS blast cells generally resemble MEPs transcriptome, but also show heterogeneous signatures of differentiation towards erythrocytes, megakaryocytes, and mast cells. The degree of differentiation was then assessed for potential correlation with clinical features. Ongoing work is focused on defining the transcriptional consequences of GATA1 mutation in the context of T21, and molecular signatures of refractory MLDS.
Overall, our results provide in-depth mechanistic insights into the transcriptional consequences of T21 in early haematopoiesis and the stepwise progression of TAM and MLDS, offering the possibility of more meaningful clinical interpretations and new therapeutic targets for these conditions.

188
Cancelled

189

Complex structural variation is prevalent and highly pathogenic in pediatric solid tumors

Ianthe van Belzen[1], Marc van Tuil[1], Shashi Badloe[1], Alex Janse[1], Eugene Verwiel[1], Marcel Santoso[1], Sam de Vos[1], John Baker-Hernandez[1], Hindrik Kerstens[1], Michael Meister[3], Jarno Drost[3], Marry M. van den Heuvel-Eibrink[4], Johannes Merks[1], Jan Molenaar[2], Weng Chuan Peng[1], Bas Top[1], Frank Holstege[1], Patrick Kemmeren[1], Jayne Hehir-Kwa[1]

1 Princess Máxima Center for Pediatric Oncology, The Netherlands
2 Princess Maxima Center for Pediatric Oncology; Utrecht University, The Netherlands
3 Princess Maxima Center for Pediatric Oncology; Oncode Institute, The Netherlands
4 Princess Maxima Center for Pediatric Oncology; UMCU-Wilhelmina Children's Hospital, The Netherlands

Presenter: Ianthe van Belzen

Backgrond: In pediatric cancer, structural variants (SVs) and copy number (CN) alterations can contribute to cancer initiation and progression, but also aid diagnosis and treatment stratification. The few studies considering complex rearrangements have found associations with tumor aggressiveness or poor outcome, but their prevalence and biological relevance across pediatric solid tumors remains unknown.
Results: In an unselected cohort of 120 primary tumors, we systematically characterized patterns of extrachromosomal DNA, chromoplexy and chromothripsis across five solid pediatric cancer types: neuroblastoma, Ewing sarcoma, Wilms tumor, hepatoblastoma and rhabdomyosarcoma. Complex SVs were identified in 56 tumors (47%) and instances of complex classes occurred in multiple cancer types, indicating shared mutational mechanisms. Hotspot regions recurrently mutated by complex SVs often overlap with cancer genes and tend to be cancer-type specific, hinting at selection pressures. In total, we identified (potentially) pathogenic complex SVs in 42 tumors that affect cancer driver genes or regions previously associated with poor outcome. Half of which were drivers known to arise from complex rearrangements, e.g. MYCN amplifications due to ecDNA and EWSR1::FLI1 fusions due to chromoplexy. In addition, we identified novel candidate complex events for follow-up analyses in 21 tumors. Recurrent examples include chromoplexy in WT1 in Wilms tumors, focal chromothripsis with 1p loss in hepatoblastomas and ecDNAs amplifying MDM2 in rhabdomyosarcomas.
Conclusions: Complex SVs are prevalent in pediatric solid tumors. They affect known cancer driver genes and represent a type of genomic variation which currently remains unexplored. Moreover, carrying complex SVs seems to be associated with adverse clinical events like progression, relapse or death. Further research is needed to establish whether complex SVs are useful for risk stratification or can be exploited by targeted treatments.

190

Effect of bladder cancer subtypes on response to immunotherapy

Sergio Vázquez[1], Lilian Marie Boll[1], Júlia Perera-Bel[1], Marta Espinosa Camarena[1], Joaquim Bellmunt[1], M.Mar Albà[2]

1 Hospital del Mar Research Institute, Spain
2 Hospital del Mar Research Institute, Catalan Institute for Research and Advanced Studies

Presenter: Sergio Vázquez

Bladder cancer tumors can be divided into different subtypes: luminal (6%), luminal-infiltrated (19%), basal-squamous (35%), neuronal (5%), and luminal-papillary (35%)(Robertson et al., 2017). The use of immune checkpoint inhibitors (ICI) is becoming ever more common in the treatment of bladder cancer, but how the subtype impacts the response to ICI is not well understood. Here we have collected tumor gene expression and mutation data from a total of 583 patients (398 with ICI response information), and investigated the factors determining the response in the different subtypes.
We have confirmed that the neuronal subtype shows a higher than average response to immunotherapy (Kim et al., 2019) but that, contrary to the commonly-held assumption, luminal-infiltrated and basal-squamous do not. The latter subtypes are characterized by higher amounts of immune cells, and we find that a high tumor mutational burden (TMB) is in general also required to respond to the treatment. We identify several unexpected similarities between the luminal and neuronal subtypes, including high TMB values and low TGF-β gene expression. Intriguingly, in these subtypes, the response does not appear to depend on the TMB and TGF-β levels but is positively associated with the amount of T cells and macrophages. The findings of the study can be used to improve the prediction of the response to immunotherapy in bladder cancer.

191

Accurate comparison of insertion and deletion mutation rates using sequence composition correction with novel sequence ambiguity scoring

Jan Verburg[1], Martin Taylor[1]

[1] MRC Human Genetics Unit, MRC Institute of Genetics and Cancer, The University of Edinburgh, United Kingdom

Presenter: Jan Verburg

Dysregulated cellular processes and exposure to exogenous agents are commonly attributed to cancer, and the patterns of mutations associated can be summarised as ‘signatures’. Constructing mutational signatures is becoming increasingly important in the clinical treatment of patients. Though insertions and deletions contribute to a significant proportion of the mutations in cancer, much of the research into signatures has been restricted to single nucleotide substitutions due to their consistent trinucleotide context. Indels are folded into 83 distinct categories when creating the indel mutational signature due to varying event size and sequence context length. Indel mutational signatures only allow fair comparison of regions with the same sequence composition. This impedes investigating underlying mechanisms of indel-causing processes. Without a null expectation or other genomic region to serve as comparison, it is impossible to make a meaningful deduction from observed mutations. This underlines a need for an approach to account for variability in sequence composition in order to make accurate comparison of indel mutation rates across the genome and better resolve aetiology of the events and processes leading to cancer. We propose a novel framework and metric that quantifies sequence ambiguity inherent in indel alignment to systematically score sequence context. In combination with sequence size, we retain sequence identity across each of the ID83 categories. This sequence ambiguity scoring framework provides means for sequence composition correction and to generate null expectations from any sequence. Moreover, it allows for compositionally adjusted indel rates to be compared between genomic regions including between species. This has previously been impossible when using indel mutational signatures. We show that compositionally adjusted indel rates in ID83 format allow for direct comparison between coding and non-coding regions and show compositionally adjusted indel rates in early/late replicating regions in mismatch repair deficient tumours. In short, systematically scoring indels for sequence ambiguity opens up new avenues to explore context specific mechanisms of indel causing mutational processes.

192

Optimizing POLE and POLD1 variant interpretation: gene-specific classification guidelines and in vitro system for functional assessment

Julen Viana-Errasti[1], Sandra Garcia[1], Mariona Terradas[1], Gabriel Capella[1], Pilar Mur[1], Laura Valle[1]

[1] Bellvitge Biomedical Research Institute (IDIBELL), Spain

Presenter: Julen Viana-Errasti

Germline pathogenic variants within exonuclease domain (ED) of polymerases epsilon and delta (encoded by POLE and POLD1 genes, respectively) cause an autosomal dominant cancer syndrome characterized by increased risk of gastrointestinal polyps, and colorectal, endometrial, ovarian and breast cancers, among other tumors. Proofreading-defective tumors of either constitutional or somatic origin, tend to accumulate thousands of mutations (>10-100 mutations per Mb), which correspond to very specific mutational spectra: mutational signature SBS10a for POLE and SBS10d for POLD1 proofreading deficient tumors. Accurate variant classification of POLE and POLD1 is crucial for a precise clinical management, both in the hereditary cancer field as for precision medicine in cancer. However, interpretation of variants in these genes is challenging, and there is an impending need of reliable functional evidence to improve variant classification.

Our group has developed gene-specific recommendations for POLE and POLD1 variant classification, which have been applied to 128 ED variants identified in patients and/or reported in public databases, allowing the classification of 34 variants as (likely) pathogenic or benign. In parallel, we are currently developing an in vitro system to assess, in a high throughput manner, the functional impact of any POLE and POLD1 ED variant. This method consists of performing high throughput editing of the human haploid HAP1 cell line with CRISPR/Cas9 to produce cellular models that harbor every reported variant in the ED of the polymerase genes, or every possible variant at each nucleotide of the ED. As endpoint, we evaluate the mutation burden and spectrum (mutational signatures) of the variants accumulated by the edited cell lines to assess the functional effect of the variant under study. A pilot study has been performed modeling 4 pathogenic ED variants, and we are optimizing the system for the high throughput application.

193

A pan-cancer copy number profile database from published array-based studies

Ioana-Antonia Vlaicu[1], Maxime Tarabichi[1]

[1] Université Libre de Bruxelles, Belgium

Presenter: Ioana-Antonia Vlaicu

Copy Number Aberrations (CNAs) accumulate throughout tumor development and are linked to prognosis. CNAs can amplify oncogenes or delete tumor suppressor genes, but even simple gains or losses have large-scale effects on the transcriptomes through gene-dosage effects. Thus, they are important to consider when studying cancer genomes and the reprogramming of the metabolic pathways of cancer cells.
To date, the data of hundreds of studies using microarrays on cancer patients have been published, and the genomic data resulting from these studies is publicly available on databases such as Gene Expression Omnibus and ArrayExpress. These valuable resources can thus be used to infer the copy number profiles of thousands of cancer patients and provide an easy way to access them in the form of a database.
Using state-of-the-art tools and novel pipelines, we constructed a database of curated copy number profiles from this published cancer array data, and compiled them in an accessible web application.
To cover a large majority of high-resolution genome-wide studies, we ran ASCAT and ASCAT.sc to infer the copy number profiles from Affymetrix SNP arrays and Illumina methylation arrays, respectively. Two automated pipelines were implemented to derive copy-number profiles in batches without the need for manual intervention. In particular, patient sex was inferred and ASCAT was run in tumor-only mode, avoiding the need to manually match tumor and normal samples. Our web application allows the retrieval and inspection of an extensive resource of cancer related copy number profiles. Users have access to the generated CNPs, along with additional illustrative and explorative figures.
This aims to allow any member of the scientific community to quickly explore the landscape of CNAs across cancer types, and also to model their prevalence for understanding their role in cancer progression, to potentially identify signatures of mutational processes and markers for clinical diagnosis and prognosis, as well as aid the identification of driver genomic regions.

194

“Hide and seek” retroelement activity in hematological malignancies

Anastasiya Volakhava[2], Sarka Pavlova[6], Kristyna Zavacka[3], Karol Pal[1], Marcela Krzyzankova[3], Hana Synackova[5], Sarka Pospisilova[6], Ilgar Mamedov[4], Karla Plevová[7]

1 CEITEC, Czech Republic
2 CEITEC Masaryk University, Czech Republic
3 CEITEC, Masaryk University, University Hospital Brno, Czech Republic
4 Masaryk University, Czech Republic
5 University Hospital Brno, Czech Republic
6 University Hospital Brno, Masaryk University, Czech Republic
7 University Hospital Brno, Masaryk University & Central European Institute of Technology, Masaryk University, Czech Republic

Presenter: Anastasiya Volakhava

Retroelements (REs), which function via a “copy-and-paste” mechanism, comprise nearly half of the human genome. The Long Interspersed Nucleic Elements, type 1 (LINE-1 or L1) are the only active autonomous REs. They are able to retrotranspose other RNAs including Alu and SVA REs, and occasionally protein-coding RNAs. Retroelements are silenced via multiple mechanisms, but genomic instability of cancer cells often leads to aberrant disruption of RE repression and enhances their transposition activity. The main goal of our research is to explore RE activity in chronic lymphocytic leukemia (CLL) and myelodysplastic syndrome (MDS) and to study the impact of therapy and TP53 inactivation on RE activity.
To identify tumor-specific RE insertions, we adopted a highly sensitive amplicon NGS protocol for localizing insertions of REs from Alu-Ya5, Alu-Yb8, or L1-HS families into their target genomic regions. In total, (i) 99 samples from 17 MDS and 21 CLL patients, and (ii) 60 samples from 4 leukemic cell lines were analyzed.
To observe the transcriptional activity of REs in CLL primary samples (N=42) and cell lines (N=12), RNA-seq data analysis using the TElocal tool was performed.
The method sensitivity of the targeted amplicon-based NGS method was evaluated revealing the 1% detection threshold for the proportion of cells with specific RE insertion. Following this result, we identified 23 candidates for new tumor-specific RE insertions. The candidate insertion evaluation showed no de novo somatic retrotransposition events in the tested cohort at the set detection limit.
The results of RNA-seq analysis showed lower TE expression in CLL patient samples taken after expansion of TP53 mutation, and in TP53-mutated cell lines. The most differentially expressed families were represented by older RE families, likely due to the co-expression effect of nearby protein-coding regions.
To summarize, the frequency of RE somatic insertions in hematological malignancies is low. A possible hypothesis is that the studied leukemia types have a low tolerance to RE insertional activity, as it can lead to leukemic cell elimination.
Supported by GACR 19-11299S, NUVR LX22NPO5102, and MH-CZ RVO 65269705.

Anastasiya Volakhava is the recipient of an EMBL Advanced Training Centre Corporate Partnership Programme fellowship.

195

The Dutch childhood cancer genome project: characterizing tumor drivers

Ianthe van Belzen[1], Joanna von Berg[1], Fleur Fleur[1], Anastasia Spinou[1], Lennart Kester[1], John Baker-Hernandez[1], Richard Gremmen[1], Alex Janse[1], Shashi Badloe[1], Sam de Vos[1], Eugene Verwiel[1], Marc van Tuil[1], Hindrik Kerstens[1], Jayne Hehir-Kwa[1], Bas Top[1], Frank Holstege[1], Patrick Kemmeren[1]

[1] Princess Máxima Center for Pediatric Oncology, The Netherlands

Presenter: Joanna von Berg

Childhood cancer is the main cause of disease-related death in children in high-income countries. Identifying genetic mutations leading to tumor formation is crucial for diagnosis, risk stratification and targeted therapies. Previous work has uncovered the first tumor driver candidates in childhood cancer. These studies were mostly performed with whole-exome sequencing (WES), facilitating discovery of tumor driver genes but of little value for non-coding or structural variant drivers. To date, most pediatric cancer genomics studies have focused on individual variant types: (coding) SNVs, copy number aberrations or gene fusions. More efforts are needed to come to an integrative approach that also accounts for functional consequences of mutations. At the Princess Máxima Center – the Dutch national institute for pediatric oncology care and research – we are collecting whole-genome sequencing (WGS) from tumor and normal tissue and RNA-sequencing (RNA-seq) from tumor tissue. The centralized position of the institute enables us to uniformly generate data of patients across the country, leading to a representative national dataset with little to no batch effects.

As of August 2023, the dataset consists of 1625 patients; 842 with all data available. Standard secondary analyses include: germline variant calling (SNV and small indels), somatic variant calling (SNV and small indels, and CNA), gene expression quantification, and gene-fusion identification. SV calling will be performed for the whole cohort later in time. Although our aim is to integrate different variant types, we need to compare the future results to established methods that identify coding tumor drivers. We have used Intogen - which combines several previously described complementary methods for gene prioritization - to identify candidate tumor driver genes. We stratified on different levels: tumor type, primary group (according to ICCC-3), and domain (solid, hematological or CNS tumors). Preliminary pan-cancer analysis identified 25 genes which are known tumor driver candidates, and five genes which are potentially novel tumor drivers. We are also preparing to run a method developed by Dietlein et al in collaboration with the authors, to identify non-coding genomic regions that could be a cancer driver if mutated.

196

Extrachromosomal DNA promotes drug resistance in pancreatic ductal adenocarcinoma cells

Tim Vorberg[3], Manuel Reitberger[3], Bernardo Rodriguez[2], Maja Starostecka[2], Ornella Kossi[1], Vanessa Vogel[3], Jan O. Korbel[2], Andreas Trumpp[3], Martin Sprick[3]

1 Biolabs, Germany
2 EMBL Heidelberg, Germany
3 German Cancer Research Center (DKFZ), Germany

Presenter: Tim Vorberg

Pancreatic ductal adenocarcinoma (PDAC) is the fourth leading cause of cancer-related deaths in Western Countries and has been predicted to become the second leading cause of cancer-related deaths by 2030. Difficulties in early detection as well as strong chemoresistance of the disease prevent the successful treatment of PDAC patients. Although we have previously identified CYP3A5 as a mediator of paclitaxel resistance, further targets remain elusive. Thus, we set out to identify these CYP3A5-independent mechanisms which could lead to the identification of novel biomarkers and drug targets to improve PDAC treatment.
Multi-drug resistance protein 1 (ABCB1) was highly upregulated in our differential gene expression analysis and confirmed as a CYP3A5-independent mediator of acquired paclitaxel resistance using cell viability assays. On the one hand, we identified genomic rearrangements leading to the expression of a functional ABCB1-SEMA3C fusion transcript. On the other hand, ABCB1 expression and copy number were significantly enhanced by the generation of extrachromosomal DNA (ecDNA) upon drug treatment. As a next step, we combined targeted enrichment of the ecDNA with short-read sequencing as well as long-read optical mapping approaches for a deeper characterization of these molecules. This strategy revealed a precise structure of the most abundant ecDNAs and further supported their strong intra-tumoral heterogeneity.
Our results suggest the clinical relevance of ABCB1 as a mediator of paclitaxel resistance in PDAC and we plan to validate our findings in patient samples. However, no clinically applicable drug is available due to severe side effects. Therefore, we want to further analyze the regulatory pathways associated with enhanced ABCB1 expression which might offer unknown therapeutic potential.

197

Profiling the complex rearrangement architecture of sarcoma

Sara Waise[1], Tom Lesluyes[1], Jonas Demeulemeester[6], Maxime Tarabichi[4], Nischalan Pillay[3], Adrienne Flanagan[5], Peter Van Loo[2]

1 The Francis Crick Institute, United Kingdom
2 The University of Texas MD Anderson Cancer Center, United States of America
3 UCL Cancer Institute, Sarcoma Genetics, United Kingdom
4 Université Libre de Bruxelles, Belgium
5 University College London Cancer Institute, Royal National Orthopaedic Hospital, United Kingdom
6 VIB - KU Leuven Center for Cancer Biology, Laboratory of Integrative Cancer Genomics, Belgium

Presenter: Sara Waise

Genomic rearrangements are key mutational processes in bone and soft tissue tumours, used for both disease classification and as prognostic and predictive biomarkers. However, the mutational processes and rearrangement architecture underlying many of these events remain poorly characterised. Recent data have indicated that sarcomas show particularly high frequencies of complex rearrangement events, including patterns which do not fit those of known mutational mechanisms. As the largest whole genome sequencing (WGS) cohort of sarcomas to date, the Genomics England (GE) 100,000 Genomes project represents a unique dataset in which to profile these events.
Structural variants (SVs) were identified in WGS data from 978 GE samples using an optimised approach based on 5 SV callers. SVs were filtered to retain only those passing the quality filter, and to remove variants present in either matched germline samples or >1% of a panel of normal variants. Caller performance was evaluated by comparison with consensus SV calls identified by the Pan-Cancer Analysis of Whole Genome Consortium. SVs identified by at least two calls were taken forward for further analysis. Chromothriptic events, extrachromosomal DNA (ecDNA) and genes enriched for SVs were identified using established algorithms.
Consistent with previous reports, the prevalence of SVs varies by tumour type, with particularly high numbers observed in liposarcoma and myxofibrosarcoma. As described previously, liposarcomas showed the highest rates of chromothripsis. We have examined some tumour types at higher granularity than previously available, demonstrating directly that the rates of complex rearrangement events vary across subtypes. For example, chromothriptic events are identified in 100% of well-differentiated liposarcomas, but less than 10% of myxoid liposarcomas. Similarly, the presence of ecDNA varies by tumour type, with the novel observation of particularly high rates in angiosarcoma.
Despite more recent advances in histological classification, survival for patients with sarcoma has remained largely unchanged for 40 years. Characterisation of the mutational processes underlying these rearrangements will shed light on the pathogenesis of these tumours.

198

Genomic analysis of WNT medulloblastoma reveals drivers of monosomy 6

Evan Wang[3], Jessica Taylor[1], Hiromichi Suzuki[2], Richard Gilbertson[1], Michael Taylor[3]

1 Cancer Research UK Cambridge Institute, University of Cambridge, United Kingdom
2 National Cancer Center Research Institute, Japan
3 The Hospital for Sick Children, University of Toronto, Canada

Presenter: Evan Wang

Medulloblastoma is the most common malignant pediatric tumor. The WNT subgroup (WNT-MB) is characterized by hyperactivation of CTNNB1, the downstream effector of the WNT signaling pathway. Additionally, loss of chromosome 6 has been identified in 85% of WNT-MB, however, we have a limited understanding of the genes on chromosome 6 driving this alteration. To characterize the genomic and transcriptomic landscape of this subgroup, we performed RNA-seq on 74 cases of WNT-MB. Clustering of transcriptomes revealed separation of monosomy 6 and disomy 6 tumors, revealing differences in neuronal and RNA processing gene signatures. Recurrent mutations were identified in SHH pathway genes, which are typically associated with SHH-MB, in a transcriptomic cluster enriched for disomy 6, suggesting dual activation of SHH and WNT signaling in a subset of WNT-MB. We also identified recurrent mutations in genes located on chromosome 6, including the transcription factor FOXO3. FOXO3 was expressed in lower rhombic lip progenitor cells, the cell of origin for WNT-MB, while expression was decreased in WNT-MB compared to other medulloblastoma subgroups. Mice carrying homozygous deletion of Foxo3 in hindbrain neural progenitors developed medulloblastoma, whereas mice with mono- or bi-allelic Foxo3 deletion in conjunction with constitutively active Ctnnb1 displayed dramatically increased cerebellar hyperplasia compared to mice with Ctnnb1 activation alone. Our analyses uncover remarkable genetic heterogeneity within WNT-MB and suggest loss of FOXO3 is a significant driver of monosomy 6.

199

Mutational processes in tumour-adjacent normal kidneys across countries with varying RCC incidence rates

Michael Stratton[3], Yichen Wang[3], Sarah Moody[3], Behnoush Abedi-Ardekani[1], Tatsuhiro Shibata[2], Paul Brennan[1], Yvette Hooks[4], Laura Humphreys[3], Jingwei Wang[3]

1 International Agency for Research on Cancer, France
2 National Cancer Center Research Institute, Japan
3 Wellcome Sanger Institute, United Kingdom

Presenter: Yichen Wang

In the past years, a series of whole genome sequencing studies in cancer across multiple continents have been conducted to uncover unknown causes of cancer through their mutational signatures, which were left in cancer genomes by past exposures to mutagenic agents. Distinct mutational signatures were detected in Renal Cell Carcinoma (RCC) samples from high RCC incidence regions. However, whether the detected mutational signatures are present in normal tissue, and thus likely to be due to exogenous exposures, or are actually a consequence of early neoplastic transformation remains unknown.
Until recently, detection of somatic mutations in highly polyclonal normal tissues such as the kidney has been challenging. However, a recent developed sequencing technology, Nanorate sequencing (Nanoseq), can cross-reference information from copies of both strands of a single DNA molecule to reduce sequencing errors to 10^{-9}. With Nanoseq, we are able to detect somatic mutations in polyclonal tissues like the normal kidney.
In this study, we used Nanoseq to sequence normal kidney tissues across countries with varying RCC incidence. Subsequently, we ran agnostic mutational signature extraction to see whether the region-specific signatures found in cancers can also be extracted from the matched normal. We demonstrated that a distinctive T>C mutational signature enriched in Japanese RCC was also present in their paired normal kidney samples. A strong transcriptional strand bias of this signature provides circumstantial evidence that is likely to be caused by an environmental agent causing bulky adducts on adenine. A subset of normal kidney samples from Serbia and Romania had mutations likely caused by the mutagenic plant extract aristolochic acid, although the mutation burden of related signatures in normal kidneys is lower than in cancer. By utilizing laser-capture microdissection to isolate distinct microscopic structures within the normal kidney, including glomeruli, proximal tubules, and medulla, we observed a substantial enrichment of aristolochic acid signatures specifically in proximal tubules as opposed to glomeruli and medulla. This finding suggests that proximal tubules may act as a sensitive sensor for exogenous mutagens, possibly due to their unique metabolic and reabsorption functions.

200

Establishing a rapid autopsy program to explore cancer evolution: preliminary experience within the uk regulatory framework

Jamie Weaver[2], David Wedge[3], Pedro Oliveira[2], Yeng Ang[1], Matthew Hall[2]

1 Northern Care Alliance NHS Foundation Trust, United Kingdom
2 The Christie NHS Foundation Trust, United Kingdom
3 The University of Manchester, United Kingdom

Presenter: Jamie Weaver

Introduction:
Large scale cancer genomics projects have explored the genomic features of early resectable malignancy and single sites of metastatic disease. However little information is available regarding the heterogeneity in the terminal stages of malignancy when resistance to all standard of care therapies has developed. Donation of tumour material after death via rapid autopsy allows extensive sampling of tumours facilitating characterisation of cancer evolution. True rapid autopsy - 6 hours - has signifant regulatory, patient and family experience implications. In the UK regulations around certifcation and registration of death create signifcant issues for implementation of rapid research autopsy protocols.
Methods:
As part of our prelimnary work in developing a local rapid autopsy program in the North west region of the UK we have conducted interviews with patients and regulatory representatives for care after death services; Medical examiners, Coroners representatives and Mortuary leads.
Results:
Patient representatives and families were broadly postive. Of 10 patients surveyed 8 of 10 would consent to be part of a rapid autopsy program. Patient representatives particularly expressed interest in support in the management of the logistics after death. In particular patient family members cited contact with primary cancer management site in last days of life and support in logistics during the early post-death period as positive factors. Regional Coroners representatives and Medical examiner representatives across the North West area (United Kingdom) were supportive of a rapid autopsy process allowing the establishment of a bespoke regional protocol for delivery of this research program.
Conclusion:
Rapid autopsy for donation of cancer material is feasible but requires signifcant input from a dedicated investigatory team and region specific protocols in the UK. Patients with terminal cancer generally expressed positive views about consenting to such a research program. Patients and family members expressed greater positivity about the rapid process (hours) over a more delayed cancer autopsy (days). These findings have lead to a larger project exploring support for cancer patients and families during last days of life and in the post-death period.

201

Tracing the origin of hepatoblastoma

Anna Wenger[4], Tim Coorens[1], Thomas Dowe[2], Maesha Deheragoda[2], Nigel Heaton[2], Foad Rouhani[3], Sam Behjati[4]

1 Broad Institute of MIT and Harvard, United States of America
2 King's College Hospital, United Kingdom
3 The Francis Crick Institute, United Kingdom
4 Wellcome Sanger Institute, United Kingdom

Presenter: Anna Wenger

One in every four children with cancer does not survive, and the origin and development of paediatric cancers is still largely unknown. Cancers are believed to develop from malignant transformation of a single cell, but studies of Wilms tumour (a childhood kidney tumour) unexpectedly identified premalignant clones of normal kidney cells (driven by H19 loss of imprinting) with the potential to generate tumours. This project expands on these findings to determine if a similar clonal expansion in the liver underpins childhood liver tumour hepatoblastoma, which has casual similarities to Wilms tumour. Such an expansion would be marked by cancer mutations present already in histologically normal liver tissue. In addition, the mutation burden and signatures present in normal liver and hepatoblastoma can shed further light on the processes driving this cancer.
To characterize the origin and evolution of hepatoblastoma, we have performed extensive sampling of hepatoblastoma tumours (2-9 samples per child) and paired histologically normal liver tissue (16-22 samples per child) from three children. These samples were sequenced with whole-genome sequencing (WGS) and a novel, sensitive duplex sequencing technique (NanoSeq), which can detect rare mutations. WGS data from all three children showed shared mutations in the tumour and normal liver, while absent from the blood. The variant allele fraction was as high as 23%, indicating that around half of the cells within that normal liver biopsy were derived from a single cell that also gave rise to the tumour. Further, NanoSeq revealed an extraordinarily high mutation burden in both the hepatoblastoma and paired normal liver tissue (1000-4000 mutations per diploid genome) compared to liver from a child without hepatoblastoma (650 mutations). The mutations in the hepatoblastoma patients, including those in the normal liver samples, were mainly caused by platinum-based treatment (mutational signatures SBS31 and SBS35), and de novo extraction of mutational signatures pointed to a novel signature present in the tumour and surrounding liver samples, warranting further studies.
In conclusion, we show that cancer mutations are present already in healthy liver suggesting that hepatoblastoma arise through clonal expansion of these cancer precursor cells in the normal liver.

202

Differentiation states of paediatric B-cell acute lymphoblastic leukaemia

Holly Whitfield[3], Laura Jardine[2], Angus Hodder[1], Sam Behjati[3], Jack Bartram[1]

1 Great Ormond Street Hospital for Children, United Kingdom
2 Newcastle University, United Kingdom
3 Wellcome Sanger Institute, United Kingdom

Presenter: Holly Whitfield

Acute lymphoblastic leukaemia (ALL) remains a life-threatening disease that overwhelms the bone marrow with malignant cells of a particular haematopoietic lineage. Almost 80% of ALL cases implicate the B lineage (B-ALL) where, particularly for children, treatment precipitates a range of chronic health problems and a reduced life expectancy. Current risk stratification is based on the presence of specific genetic aberrations and other clinical factors but does not consider transcriptome-based cell phenotype. We wanted to assess the relationship between immature differentiation states and high-risk genetic subgroups.
We sequenced the single-cell transcriptomes and whole genomes of bone marrow biopsies from 32 children with B-ALL, spanning disease treatment. Our dataset includes bone marrow biopsies at diagnosis (D0), post-treatment biopsies (D29), and further samples for patients who relapse. This allows us to assess the bone marrow regeneration profiles of patients and better understand how therapeutic response is mediated by cancer phenotype.
Analysis of > 300,000 cells shows transcriptional diversity across genetic subtypes and distinct transcriptional signals in refractory patients. Using the fetal cell atlas as a reference, we projected B-ALL cells onto a trajectory of healthy haematopoiesis to identify differentiation states. We find that immature differentiation programs might be more expressed in some subgroups or individuals. Our results suggest that the accuracy of risk prediction at diagnosis may be improved by transcriptomic assessment. Improved early-disease risk-stratification will aid the identification of patients whose treatment can be safely de-escalated or modified to spare them the late effects of intensive treatment.

203

Profiling copy number mutational signatures in KRAS mutant non-small cell lung cancer

Laura Woodhouse[2], David Wedge[1], Colin Lindsay[1], Aliah Hawari[2]

1 The University of Manchester, United Kingdom
2 The University of Manchester, Manchester Cancer Research Centre, United Kingdom

Presenter: Laura Woodhouse

Background: Lung cancer is the 3rd most common cancer in the UK and leading cause of cancer mortality worldwide. KRAS mutant cancers are the largest molecular subset of non-small cell lung cancer (NSCLC). Immunotherapy is standard 1st line therapy for advanced stage NSCLC, but reliable biomarkers for patient selection are lacking. Mutational signatures are a genetic imprint of mutational processes that have occurred during carcinogenesis, and preliminary evidence suggests a role in predicting immunotherapy response.
Methods: Tumour whole genome sequencing (WGS) data is available through the 100 000 Genomes Project (Genomics England). Copy number (CN) signature extraction was performed using SigProfiler (COSMIC version 3.3), and KRAS mutations profiled using Ensembl Variant Effect Predictor (VEP).
Results: 1058 NSCLC patients were included; 734 KRAS wild type (69%) and 324 KRAS mutant (31%). Common KRAS variants were present, including G12C (38%), G12V (16.4%) and G12D (14.1%). Nine CN signatures were extracted and present in both KRAS mutant and wild type cohorts. KRAS mutant cancers had lower CN signature activity when compared to wild type, except for the normal diploid signature (CN1, $p = 0.017$, OR 1.38). Within the KRAS mutant group 10 variants were analysed alongside wild type, with significant variation in CN signature expression seen across subtypes. KRAS G12D was positively associated with diploidy (CN1, $p<0.05$) and negatively associated with loss of heterozygosity (CN9, 13, $p<0.05$).
Discussion: KRAS mutant NSCLC is more chromosomally stable when compared to KRAS wild type. G12D is commonly associated with never smokers, and the more chromosomally stable phenotype observed could be explained by the association between tobacco smoke and chromosomal instability. Further work is ongoing to comprehensively profile the full set of mutational signatures, including known smoking signatures. Using reverse translation and clinical data integration, mutational signatures could provide prognostic information and inform predictive models.

204

Mapping the somatic mutations during the evolutionary transition from oral leukoplakia to oral squamous cell carcinoma

Ting Yang[1], Ludmil Alexandrov[1]

[1] University of California San Diego, United States of America

Presenter: Ting Yang

Head and neck cancer is the sixth most prevalent malignancy worldwide with over 10,000 death per year in the United States alone. The main risks for developing head and neck cancer are tobacco exposure, alcohol consumption, and human papilloma virus (HPV) infection. Oral cavity squamous cell carcinoma (OSCC) encompasses the majority of all oral cancers and is the most predominant subtype of HPV-negative head and neck cancer. OSCCs are often preceded by oral potentially malignant disorders with oral leukoplakias (OL) being the most common precursor lesion. In this project, our primary goal is to evaluate the evolutionary role of somatic mutations in the transition from OL into OSCC by compressively profiling over 300 pairs OL and OSCC. Here, we present a pilot of the project where we examine whole-exome sequencing data from 36 OSCCs and 33 OLs from two cohort: (i) an India cohort encompassing 28 paired oral gingivobuccal cancer (GBC) and OL; (ii) a Taiwan cohort including 8 OSCC and 5 OL cases. Molecular events were identified by using our previously developed bioinformatics pipelines to reveal: (i) driver genes, (ii) mutation burden, (iii) mutational signatures, and (iv) copy number alterations. Amongst the 28 Indian cases, the top three most commonly mutated driver genes were CASP8 (9/28 samples), TP53 (2/28), and PIK3CA (2/28). Further, the median tumor mutational burden (TMB) per Mb for GBC was 5.43, whereas the TMB for OL was 2.12 (p-value 0.01). Genes that were significantly mutated in GBC but not OL samples were TP53 and FBXW7 (Fisher Exact Test's p-values < 0.05). While comparing oral cancer samples between the India and Taiwan cohort, HRAS and NOTCH1 were not among the top 30 mutated genes in the Taiwanese cohort. Taken together, the results from our pilot demonstrate the feasibility of our student and its ability to elucidate the somatic mutations during the evolutionary transition from oral leukoplakia to oral squamous cell carcinoma.

205

Identifying significant genomic information in cancer through integrating multi-omics datasets

Zhihao Yao[4], Magnar Bjørås[3], Victor Jin[2], Junbai Wang[1]

[1] Akershus University Hospital and University of Oslo, Norway
[2] Medical College of Wisconsin, United States of America
[3] Norwegian University of Science and Technology, Norway
[4] Olso University Hospital, Norway

Presenter: Zhihao Yao

The advancement of both sequencing technology and computational methods has revolutionized the field of cancer genomics, brings new requirements and challenges to biologist, clinicians, and mathematicians. In previous studies, we proposed a novel methodology to investigate genome-wide regulatory mechanisms in follicular lymphoma. Potential biomarkers for follicular lymphoma are identified through integrating whole genome sequencing, methylation, gene expression, topologically associated domain information in regulatory mutation prediction. Meanwhile, we developed a non-encoding single nucleotide variants (SNVs) analysis tool bpb3, BayesPI-BAR version 3, aiming to identify the functional mutation blocks (FMBs) by integrating genome sequencing and transcriptome data. By testing the datasets from follicular lymphoma and melanoma, bpb3 automatically and robustly identifies FMBs. Recently, we proposed a novel mathematical framework for integrating Hi-C data with many other epigenetic modification data, like enhancer, promoter, repressor etc., for investigating intra-chromosomal community interactions. By comparing genomic and network topological features through differential intra-chromosomal interaction network analysis between tamoxifen-sensitive cell line and tamoxifen-resistant cell line, we identified significant genes with biological functions.

206
Deciphering FGFR3-TACC3 oncogenic fusions

Julia Yemelyanenko Lyalenko[1], Jinhyuk Bhin[3], Sjoerd Klarenbeek[1], Ji-Ying Song[1], Catrin Lutz[1], Marieke van de Ven[1], Shridar Ganesan[2], Lodewyk Wessels[1], Daniel Zingg[1], Jos Jonkers[1]

1 Netherlands Cancer Institute, The Netherlands
2 Rutgers Cancer Institute of New Jersey, United States of America
3 Yonsei University College of Medicine, Republic of Korea

Presenter: Julia Yemelyanenko Lyalenko

Chromosomal rearrangements of the fibroblast growth factor receptor (*FGFR*) genes that give rise to fusions are one of several mechanisms by which the FGF/FGFR signaling axis can become deregulated and result in enhanced signaling in cancer. We have previously identified truncation of exon (E) 18 of *FGFR2* as a potent single-driver alteration in cancer, independently of the rearrangement (RE) partner. In contrast, the same does not appear to hold true for its E18-truncated ortholog *FGFR3*.
We mined human oncogenomic datasets from Hartwig Medical Foundation (>2,500 WGS profiles) and Foundation Medicine (>200,000 hybrid-capture panel-seq profiles) for alterations affecting *FGFR3*. Examination of the structural variants affecting *FGFR3* uncovered that 85% of all E18-truncating *FGFR3* REs involved transforming acidic coiled-coil-containing protein 3 (*TACC3*) as the downstream fusion partner gene. Additionally, there was a clear predominance of self-interacting domains (95%) among all *FGFR3* RE partners, with the coiled-coil domain being the most recurrent, suggesting enhanced receptor dimerization and downstream signaling capacity for the majority of the *FGFR3* fusions.
Our findings led to the generation of a compendium of *FGFR3* structural variants which we then functionally tested. *In vitro* testing of mouse mammary epithelial cells expressing *Fgfr3ΔE18-Tacc3E7-E16* fusion variants showed that both E18-truncation and a fusion partner were required for 3D outgrowth and signaling induction. *In vivo* evaluation of the oncogenic capacity of *Fgfr3* variants using somatic delivery of lentiviruses to the mouse mammary gland and lung found that *Fgfr3ΔE18-Tacc3E7-E16* fusion variants rapidly induced mammary and lung tumor formation in wild-type and *Wap-Cre;Cdh1F/F*, and in wild-type and *Trp53F/F* mice, respectively. In contrast, *Fgfr3full-length(FL)*, *Fgfr3ΔE18* and *Fgfr3FL–Tacc3E7-E16* were non-tumorigenic. Noteworthy, mammary tumors driven by *Fgfr3ΔE18-Tacc3E7-E16* fusions were sensitive to the FGFR inhibitor AZD4547 during a drug intervention study.
In summary, our findings show that E18-truncating *FGFR3* alterations are recurrent across human cancers and, as opposed to *FGFR2*, depend both on E18-truncation and an additional fusion partner with dimerizing capacity such as *TACC3* for oncogenic activity.

207

Improved methods of analysis in functional genomics screens: application to screens of tumour microenvironment stress

Hanting Yong[1], Dean Singleton[1], Saumya Bhatta[1], Stephen Jamieson[1], Tet-Woo Lee[1]

[1] University of Auckland, New Zealand

Presenter: Hanting Yong

Discovering new tumour-selective therapeutic targets is important for developing novel cancer drugs and deepening our insight into tumour biology. One characteristic that sets tumour cells apart from normal cells is the hypoxic, nutrient-scarce and acidic nature of the tumour microenvironment (TME) stemming from chaotic and leaky tumour vasculature. This area of research has long been of interest but the complexities of tumour biology mean there remains much to be unravelled regarding how tumour cells adapt to the harsh TME conditions.

One powerful approach to investigate gene function while minimising bias from prior knowledge is through functional genomic screens. To find TME stress tolerance mechanisms, we have conducted whole-genome pooled CRISPR-Cas9 knockout screens in HCT-116/54C, a colorectal carcinoma subline. The cells were transduced with the whole-genome Brunello guide RNA (gRNA) library, and the resultant knockout cell population subjected to a selective pressure of chronic hypoxia, glucose deprivation or acidosis. Differential survival under selection was then used to infer the functional importance of targeted genes.

Initial analyses of our data using the commonly-used tool MAGeCK were sub-optimal due to the confounding effect of differential growth in our stressed versus control conditions, and also revealed limitations of the alpha-robust rank aggregation (α-RRA) method for combining gRNA-level results to gene-level results. To improve analysis of our TME screen data, we developed growth-corrected voom/limma (GCVL), a method for data-driven estimation of growth covariates that are then incorporated into the voom/limma linear model analysis pipeline to properly model differential growth between conditions. We also investigated several methods of integrating gRNA-level results to gene-level results – in some cases methods conventionally used for pathway-level analysis of RNA sequencing data, such as CAMERA, ROAST and gene set test, possess some advantages over methods specifically developed for functional genomics such as α-RRA.

Our results not only reveal new insights into the biology underlying TME stress tolerance, but also highlight important considerations when analysing pooled screen data, maximising the potential of this tool to generate promising gene candidates.

208

mSigPortal: A comprehensive platform for interactive mutational signature analysis in cancer genomics

Tongwu Zhang[1], Jian Sang[1], Alyssa Klein[1], Aaron Ge[1], Madhu Kanigicherla[1], Phyllip Cho[1], Thuong Nguyen[1], Brian Park[1], Kailing Chen[1], Jonas S. Almeida[1], Stephen Chanock[1], Maria Teresa Landi[1]

[1] National Cancer Institute, United States of America

Presenter: Tongwu Zhang

Mutational signatures represent distinct mutation patterns resulting from specific mutagenic processes. These signatures offer insights into tumor etiology, biomarker identification, diagnosis, and therapeutic strategies. Unfortunately, current repositories of mutational signatures are spread across various websites. Moreover, these resources do not enable the re-analysis of publicly available or user-generated data to assess mutational signatures. Additionally, they require computational and bioinformatic expertise, limiting their use for the broader research community.

To bridge these gaps, we present mSigPortal, a comprehensive, user-friendly platform that facilitates interactive exploration and analysis of mutational signatures. mSigPortal aggregates the most comprehensive mutational signature data from extensive cancer genomics studies, encompassing mutational profiles, deciphered signatures, associated etiological information, and other genomic data. This platform offers various functionalities, empowering users to interactively and comprehensively explore, integrate, and visualize signature-related data. Furthermore, mSigPortal incorporates state-of-the-art methodologies for signature analysis, introducing a unified framework for assessment accuracy. By revisiting mutational signature analysis through mSigPortal, novel perspectives on cancer etiology and the discovery of rare mutational signatures become attainable.

Aligned with the FAIR principle, mSigPortal not only provides Application Programming Interfaces (API) and Software Development Kits (SDK) to streamline distributed data processing workflows and graphical representation of mutational signature data but also eliminates the necessity for server-side data aggregation. These resources can be seamlessly integrated into web applications or portals, enhancing accessibility.

In summary, mSigPortal offers cutting-edge techniques with an intuitive framework, streamlining the exploration and analysis of mutational signatures, thereby enriching our understanding of mutagenesis in cancer development. This platform bridges the gap between intricate analytical methods and wider scientific accessibility, fostering advancements in cancer genomics research.

209

Analysis of the tumor immune microenvironment in advanced salivary gland cancers

Erika Zuljan[3], Damian Tobias Rieke[2], Eric Blanc[1], Benjamin Von Der Emde[2], Iris Piwonski[2], Frederick Klauschen[2], Ingeborg Tinhofer[2], Andreas Mock[4], Peter Horak[5], Ulrich Keller[2], Konrad Klinghammer[2], Stefan Fröhling[6], Sebastian Ochsenreither[2], Ulrich Keilholz[2], Dieter Beule[1]

1 Berlin Institute of Health at Charité (BIH), Germany
2 Charité, Germany
3 Charité-Universitätsmedizin Berlin, Germany
4 Ludwig-Maximilians-Universität München (LMU), Germany
5 National Center for Tumor Diseases (NCT), Germany
6 National Center for Tumor Diseases (NCT) and German Cancer Research Center (DKFZ), Germany

Presenter: Erika Zuljan

Salivary gland cancers (SGC) are rare and heterogeneous tumors. Adenoid cystic carcinoma (ACC) is among the most common histologies and characterized by an immune depleted microenvironment. Advanced SGC lack established treatment options and show poor response to immunotherapy. The aim of this study is to characterize the tumor immune microenvironment (TIME) in advanced SGC, with a focus on ACC, to identify potential therapeutic strategies.
Advanced SGC from the DKTK MASTER program were analyzed using bulk RNA-seq (n=95), exome (n=55), genome (n=50) and single cell RNA sequencing (n=13). Results were validated immunohistochemically (n=14).
RNA-seq immune deconvolution showed an overall lower immune cell infiltration in ACC compared to other SGC entities. A small subset of ACC showed high immune infiltration. TMB in ACC was significantly lower than in other SGC and was not associated with inflammation. No association was found between inflammation and other clinical parameters. The immune checkpoint VTCN1 was found to be significantly overexpressed in ACC, which was confirmed by immunohistochemistry and single cell RNA seq. Bulk-based immune infiltration was validated by single cell data.
These data suggest an immune-high subgroup in advanced ACC. A clinical trial of a VTCN1-directed therapy is ongoing.

The index number refers to the abstract number and not the page number.

A

B

C

H

K

L

L

L

M

M

R

S

S

T

T

V

W

Z

www.embl.org/events