Extended Abstracts for the 49th International Symposium of the Princess Takamatsu Cancer Research Fund, 2018

DECIPHERING, SIMULATING AND EDITING OF THE CANCER GENOME

November 6-8, 2018
Tokyo, Japan

Edited by :
Tatsuhiro Shibata, Matthew Meyerson, Peter Campbell, and Hiroyuki Aburatani

Princess Takamatsu Cancer Research Fund

CONTENTS

CURRENT STATUS OF CANCER GENOMICS INCLUDING DATA ANALYSIS

TUMOR HETELOGENEITY AND IMMUNOGENOMICS

NAKAHARA MEMORIAL LECTURE

CLINICAL GENOMICS AND CANCER GENOME EDITING

CONCLUDING REMARKS

INTERROGATING THE ARCHITECURE OF CANCER GENOMES

Peter Campbell
Cancer Genome Project
Wellcome Sanger Institute
Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, UK
(pc8@sanger.ac.uk)

Cancer is driven by mutation. Using massively parallel sequencing technology, we can now sequence the entire genome of cancer samples, allowing the generation of comprehensive catalogues of somatic mutations of all classes. Bespoke algorithms have been developed to identify somatically acquired point mutations, copy number changes and genomic rearrangements, which require extensive validation by confirmatory testing. The findings from our first handful of genomes illustrate the potential for next-generation sequencing to provide unprecedented insights into mutational processes, cellular repair pathways and gene networks associated with cancer development. I will also review possible applications of these technologies in a diagnostic and clinical setting, and the potential routes for translation.

Peter Campbell, PhD

2003	PhD, University of Cambridge
2007	Post-doctoral Fellowship, Sanger Institute
2010-present	Group Leader, Wellcome Sanger Institute

MOLECULAR GENOMIC LANDSCAPE OF LIVER CANCER

Tatsuhiro Shibata
Laboratory of Molecular Medicine
The Institute of Medical Science, The University of Tokyo (IMSUT)
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
(tshibata@ims.u-tokyo.ac.jp)

Division of Cancer Genomics
National Cancer Center Research Institute
5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(tashibat@ncc.go.jp)

1. Whole genome landscape of human cancer towards deconvolution of the cancer genome

Recent advances in sequencing technologies can afford whole genome sequencing approach to cancer at quite reasonable cost. This includes whole genome sequencing (WGS), whole transcriptome sequencing, whole methylation sequencing or other epigenetic approaches such as Chromatin-immunoprecipitation sequencing etc. These analyses identify somatic genetic (nucleotide substitution, insertion/deletion, copy number change, structural rearrangement, mobilization of transposable elements, etc.) and epigenetic (aberrant methylation of CpG, change of active and inactive histone modifications) alterations in cancer cells. Recent studies have further identified the interactions between genetic and epigenetic alterations that include promoter/enhancer hijacking, driver genes regulating epigenetic landscape (IDH1/2, TET2 etc.) and oncogenic non-coding RNA. These whole-genome scale characterizations of the cancer genome can explore the following two questions; trace of positive and negative selections such as identification of cancer driver and anti-driver genes, predisposing germline variations, and trace of carcinogenesis processes in human cancer, including mutational and other genetic/epigenetic signatures, precancerous clonal events, cancer evolution and heterogeneity (Figure 1).

2. Landscape of mutational signatures in hepatocellular carcinoma (HCC)

Mutagenic factors or processes leave unique substitutions with characteristic sequence contexts on DNA, that are called *mutational signatures*. We attempted to examine the relationship between mutational signature with epigenetic contexts. Eight mutational signatures (W1-8) were extracted from WGS data from 266 HCC samples (Figure 2a), and the contribution of each signature was compared between the active and inactive chromatin status (Figure 2b, c). A significant enrichment of W2 (corresponding to COSMIC signature 16) in active chromatin areas (odds ratio: 6.32) was observed irrespective of hepatitis virus status. Signatures W6 (odds ratio: 3.09) and W5 (which corresponds to COSMIC signature 19) (odds ratio: 1.24) were enriched in the active histone-marked genome, whereas signatures W7 (odds ratio: 2.30), W4, W1 (which corresponds to COSMIC signature 1), W3 (which corresponds to COSMIC signature 12) and W8 occurred more frequently in inactive chromatin segments.

3. HBV integration and epigenetic state

The HBV genome integration process is associated with host DNA nicking and recombination. However, the DNA methylation of integrated HBV sequences and the HCC genome around HBV integration sites remain to be thoroughly explored. Analyzing a total of 108 non-cancerous liver tissues and 102 HCC samples from the same patients, 1,010 non-clonal integration events in non-cancerous tissues and 476 clonal integration events in HCCs were identified. Non-clonal HBV integrations in non-cancerous liver tissues were significantly ($P < 2.2e\text{-}16$) enriched in the transcriptionally active chromatin regions, whereas the clonal integration sites in tumors were located significantly more frequently in the inactive chromatin regions ($P = 0.0085$). This suggests that although substantial HBV integration occurs in open and transcribing epigenetic regions, these integration events are negatively selected for in cancerous tissues, probably because of their deleterious effects on hepatocyte growth.

4. Structural rearrangements of integrated HBV genomes

We explored whether structural alterations of HBV genomes in the tumor genome may play a role in the host genome structural rearrangements. Paired reads that both mapped to the HBV genome were used to detect HBV genome rearrangements in five HCC genomes, and the frequency of these rearrangements (per Mb) was extremely higher than that of the tumor genome (42,815-fold increase on average, $P < 2.2e\text{-}16$) (Figure 3a). To verify the rearrangements in the HBV genomes, we obtained long-read (the average length of a sequenced read was 1,783 kb) WGS data for one pair of HCC and non-cancerous liver genomes (HX25). Based on the 16x coverage of PacBio long-read WGS, we de novo

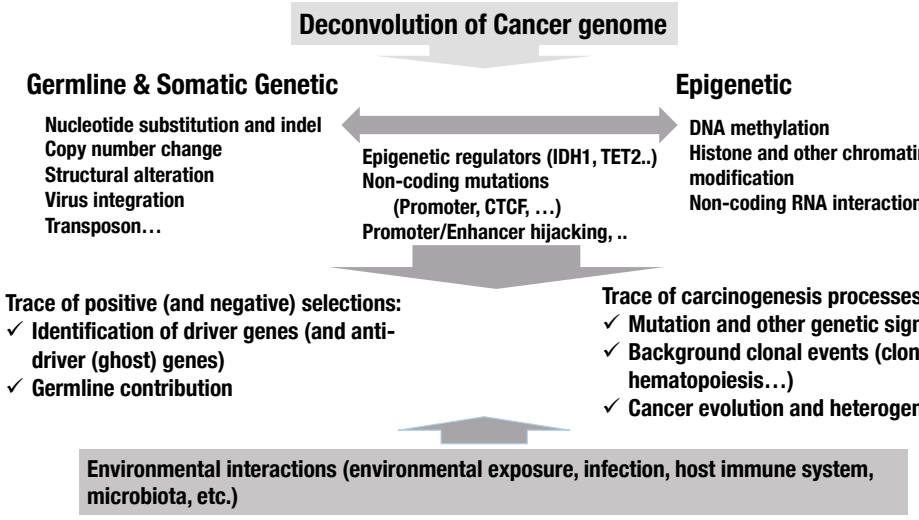
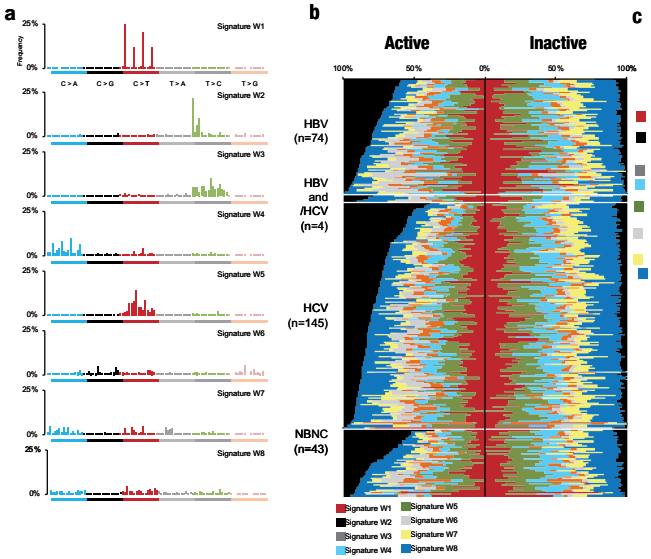


Figure 1 Whole genome landscape of cancer: genetic and epigenetic features and their interactions



Odds ratio of signature

	Active	Inactive
Cosmic Sig1	0.8158689	1.225687
Cosmic Sig16	6.326262	0.1580712
Cosmic Sig12	0.6853502	1.459108
Cosmic Sig4	0.5642329	1.772318
Cosmic Sig19	1.244363	0.8036237
Signature W6	3.093719	0.3232355
Cosmic Sig5	0.4340472	2.303897
Signature W8	0.5635623	1.774427

P< 2.2e-16, fisher's exact test

Figure 2 Distribution of mutational signatures across different chromatin
a. Eight mutational signatures identified from active or inactive chromatin states of 266 HCCs. b. Contribution of the eight mutational signatures to each tumor in active and inactive chromatin states. HBV: hepatitis B virus-related HCCs, HCV: hepatitis C virus-related HCCs, NBNC: HCCs without HBV or HCV infection. c. Odds ratio of the contribution of the signature in active and inactive chromatin.

constructed six independent long consensus reads (5,425-10,251 kb) of the integrated HBV genome. These reads perfectly confirmed the presence of the rearrangements (deletions, inversions and duplications) detected by the Illumina short reads (Figure 3b).

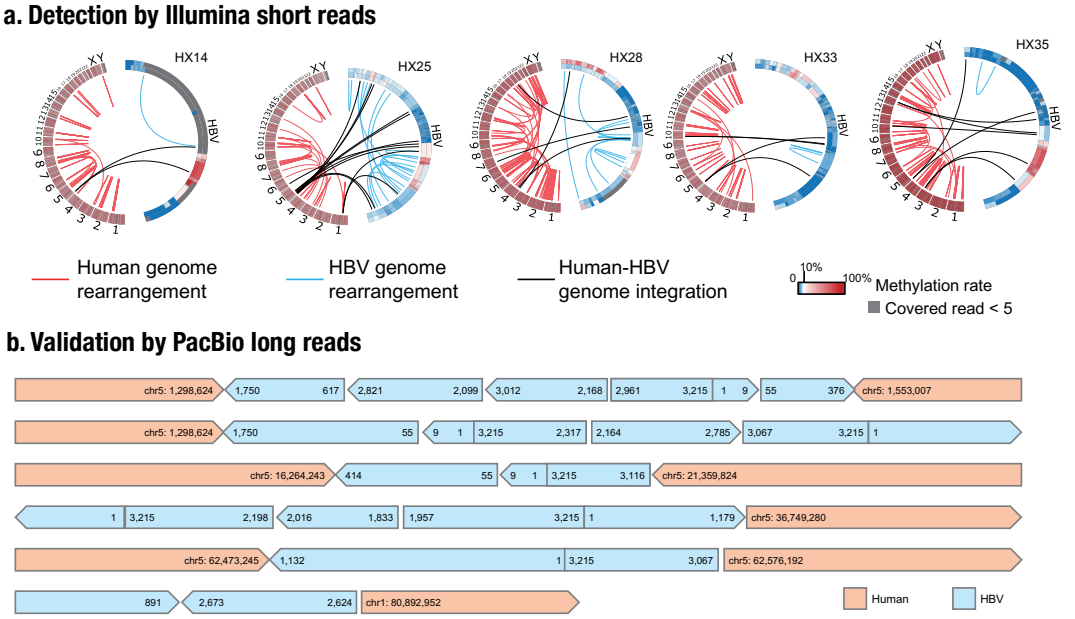


Figure 3 Massive rearrangements in the integrated HBV genome
a. Structural rearrangements of the HBV and human genomes in five samples. Red line: HBV internal rearrangement; black: HBV and human genome breakpoint; blue: human genome rearrangement. b. Six long reads from PacBio WGS of the HX25 tumor including the complex rearrangement of viral and human genome sequences. The integrated or rearranged site was validated by at least two short reads obtained by the Illumina platform.

To further confirm this finding, we performed HBV capture sequencing of a larger HCC cohort and identified more than five rearrangements in 74% (45/61) of cases. Characteristically, tumors with frequent HBV genome rearrangements harbored significantly fewer total somatic mutations (P = 0.04), suggesting that HBV genome instability-associated genetic alterations may complement part of the driver events in HCC.

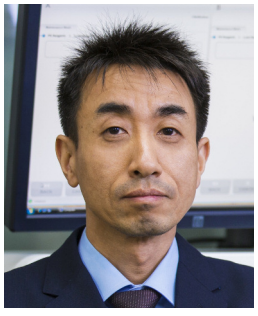
5. Conclusion

This integrative analysis identified interdependency between genetic, viral and epigenetic alterations in liver cancer. Understanding the underlying molecular mechanisms would facilitate the identification of epigenetic driver events as well as carcinogenic processes and ultimately contribute to genome-based treatment and prevention.

References

1. Hama N, Totoki Y, Miura F, Tatsuno K, Saito-Adachi M, Nakamura H, Arai Y, Hosoda F, Urushidate T, Ohashi S, Mukai W, Hiraoka N, Aburatani H, Ito T, Shibata T. Epigenetic landscape influences the liver cancer genome architecture. Nat Commun. 2018 Apr 24;9(1):1643. doi: 10.1038/s41467-018-03999-y.
2. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, Nellore V, Kongpetch S, Ng AWT, Ng LM, Choo SP, Myint SS, Thanan R, Nagarajan S, Lim WK, Ng CCY, Boot A, Liu M, Ong CK, Rajasegaran V, Lie S, Lim AST, Lim TH, Tan J, Loh JL, McPherson JR, Khuntikeo N, Bhudhisawasdi V, Yongvanit P, Wongkham S, Totoki Y, Nakamura H, Arai Y, Yamasaki S, Chow PK, Chung AYF, Ooi LLPJ, Lim KH, Dima S, Duda DG, Popescu I, Broet P, Hsieh SY, Yu MC, Scarpa A, Lai J, Luo DX, Carvalho AL, Vettore AL, Rhee H, Park YN, Alexandrov LB, Gordân R, Rozen SG, Shibata T, Pairojkul C, Teh BT, Tan P. Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. Cancer Discov. 2017 Oct;7(10): 1116-1135.
3. Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell. 2017 Jun 15;169(7):1327-1341.
4. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, Taniguchi H, Kawakami Y, Ueno M, Gotoh K, Ariizumi SI, Wardell CP, Hayami S, Nakamura T, Aikata H, Arihiro K, Boroevich KA, Abe T, Nakano K, Maejima K, Sasaki-Oku A, Ohsawa A, Shibuya T, Nakamura H, Hama N, Hosoda F, Arai Y, Ohashi S, Urushidate T, Nagae G, Yamamoto S, Ueda H, Tatsuno K, Ojima H, Hiraoka N, Okusaka T, Kubo M, Marubashi S, Yamada T, Hirano S, Yamamoto M, Ohdan H, Shimada K, Ishikawa O, Yamaue H, Chayama K, Miyano S, Aburatani H, Shibata T, Nakagawa H. Whole genome mutational landscape and characterization of non-coding and structural mutations in liver cancer. Nat Genet, 2016, 48:500-9.
5. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, Hama N, Hosoda F, Urushidate T, Ohashi S, Hiraoka N, Ojima H, Shimada K, Okusaka T, Kosuge T, Miyagawa S, Shibata T. Genomic spectra of biliary tract cancer. Nat Genet, 2015, 47:1003-10.

6. Totoki Y, Tatsuno K, Covington KR, Ueda H, Creighton CJ, Kato M, Tsuji S, Donehower LA, Slagle BL, Nakamura H, Yamamoto S, Shinbrot E, Hama N, Lehmkuhl M, Hosoda F, Arai Y, Walker K, Dahdouli M, Gotoh K, Nagae G, Gingras MC, Muzny DM, Ojima H, Shimada K, Midorikawa Y, Goss JA, Cotton R, Hayashi A, Shibahara J, Ishikawa S, Guiteau J, Tanaka M, Urushidate T, Ohashi S, Okada N, Doddapaneni H, Wang M, Zhu Y, Dinh H, Okusaka T, Kokudo N, Kosuge T, Takayama T, Fukayama M, Gibbs RA, Wheeler DA, Aburatani H, Shibata T. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat Genet, 2014, 46:1267-73.



Tatsuhiro Shibata, MD, PhD

1990-1994	Graduated from Graduate School of Medicine and Faculty of Medicine, The University of Tokyo (Pathology)
1995-1997	Postdoctoral Researcher, Developmental Biology Center, University of California, Irvine
2010-present	Chief, Division of Cancer Genomics, National Cancer Center Research Institute
2014-present	Professor, Laboratory of Molecular Medicine, The Institute of Medical Science, The University of Tokyo

GERMLINE DETERMINANTS OF THE SOMATIC MUTATION LANDSCAPE IN 2,642 CANCER GENOMES

Jan O. Korbel
Genome Biology
European Molecular Biology Laboratory (EMBL)
Meyerhofstraße 1
69117 Heidelberg, Germany
(korbel@embl.de)

My presentation will cover research from our group on polymorphic genome structural variation, and on genomic somatic DNA rearrangements in cancer. I will provide an update on our efforts to reconstruct patterns of polymorphic genome structural variation through analysis of DNA sequencing data within the Human Genome Structural Variation Consortium, which uses a range of different sequencing platforms including long read and strand-specific sequencing to assess the full spectrum of genetic variation in human genomes. Furthermore, we recently developed computational including cloud-based approaches [1] for performing an analysis of >2,800 deeply sequenced tumor/normal paired genomes in the context of the Pan Cancer Analysis of Whole Genomes (PCAWG) project, to search for commonalities and differences in molecular processes leading to cancer in different tumor entities. I will specifically highlight scientific results of the PCAWG Germline Cancer Genome working group [2,3], which has reconstructed the germline genomes of >2,800 cancer patients to examine how somatic mutation patterns associate with germline genotypes. Based on pan-cancer analysis we recently inferred and subsequently verified experimentally a new mechanism of cancer gene activation involving somatic *neo* topological association domain (neo-TAD) formation mediated by recurrent tandem duplications, which activate the *IGF2* oncogene locus. [4]

References

1. Yakneen, S., Waszak, S. M., Gertz, M. & Korbel, J. O. Enabling rapid cloud-based analysis of thousands of human genomes via Butler. *Nat Biotechnol* in press (2018).
2. Campbell, P., Getz, G., Stuart, J. M., Korbel, J. O. & Stein, L. D. The PanCancer Analysis of Whole Genomes. *bioRxiv*, doi:http://sci-hub.tw/10.1101/185736 (2017).
3. Waszak, S. M., ; *et al.* Germline determinants of the somatic mutation landscape in 2,642 cancer genomes. *BioRxiv* (2017).
4. Weischenfeldt, J. *et al.* Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat Genet* 49, 65-74, doi:10.1038/ng.3722 (2017).

Jan O. Korbel, PhD

1994-2001	Dipl.-Ing. (German engineering degree) degree in Biotechnology, Specialization Medical Biotechnology, from Technical University (TU) Berlin
2001-2005	Doctoral Research Structural & Computational Biology Unit, EMBL Heidelberg
2005-2007	Marie Curie and EMBO Postdoc at Yale University, New Haven, CT
2008-2016	Group Leader at EMBL Heidelberg (since 2014: Additionally Leader of Research Group in the Context of the MMPU of EMBL & Heidelberg University)
2016-present	Group Leader and Senior Scientist at EMBL Heidelberg (Secondary EMBL affiliation with the EMBL-European Bioinformatics Institute, UK)
2018	Winner of the 2018 Pezcoller Foundation - EACR Cancer Researcher Award

MOLECULAR BASIS OF PEDIATRIC T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA

Junko Takita
Department of Pediatrics, Graduate School of Medicine
Kyoto University
54 Kawahara-cho, Shogoin Sakyo-ku, Kyoto 606-8507, Japan
(jtakita@kuhp.kyoto-u.ac.jp)

T-cell acute lymphoblasic leukemia (T-ALL) is an aggressive malignancy which accounts for 10-15% of pediatric leukemia. With appropriately intensive therapy, children with T-ALL have an outcome approaching that of children with B-lineage ALL. However, the outcome of T-ALL patients with primary resistant or relapsed leukemia remains extremely poor [1]. Therefore, current research efforts are focused on the development of therapeutic targets and relevant biomarkers for intractable T-ALL.

Constitutive activation of *NOTCH1* signaling is the most prominent oncogenic pathway in T cell transformation [2]. Furthermore, up to 40% of pediatric T-ALL cease carry *TAL1* gene aberrations mainly due to small deletions in chromosome 1p32 and translocations involving *TAL1* locus [3]. In addition, recently, a *TAL1* super enhancer abnormality was reported, which shows aberrant expression of *TAL1* without gene fusion [4]. However, since the prognostic significance of these genetic alterations in T-ALL is not clear, genetic basis which contributes aggressive phenotype or progression of pediatric T-ALL is still to be elucidated.

Here, we report comprehensive profiling of 121 cases of pediatric T-ALL using RNA sequencing and/or targeted capture sequencing through which we identified new recurrent gene fusions involving *SPI1* (*STMN1-SPI1* and *TCF7-SPI1*) [5] (Fig. 1A). Cases positive for fusions involving *SPI1* (encoding PU.1), which accounted for 3.9% (7/181) of the total examined pediatric T-ALL cases. The fusion-positive samples invariably showed markedly elevated *PU.1* expression (Fig. 1B), most likely reflecting high-level fusion transcripts from the rearranged allele under the control of a heterologous promoter from *TCF7* or *STMN1*, which was demonstrated to be highly expressed in the fusion-positive

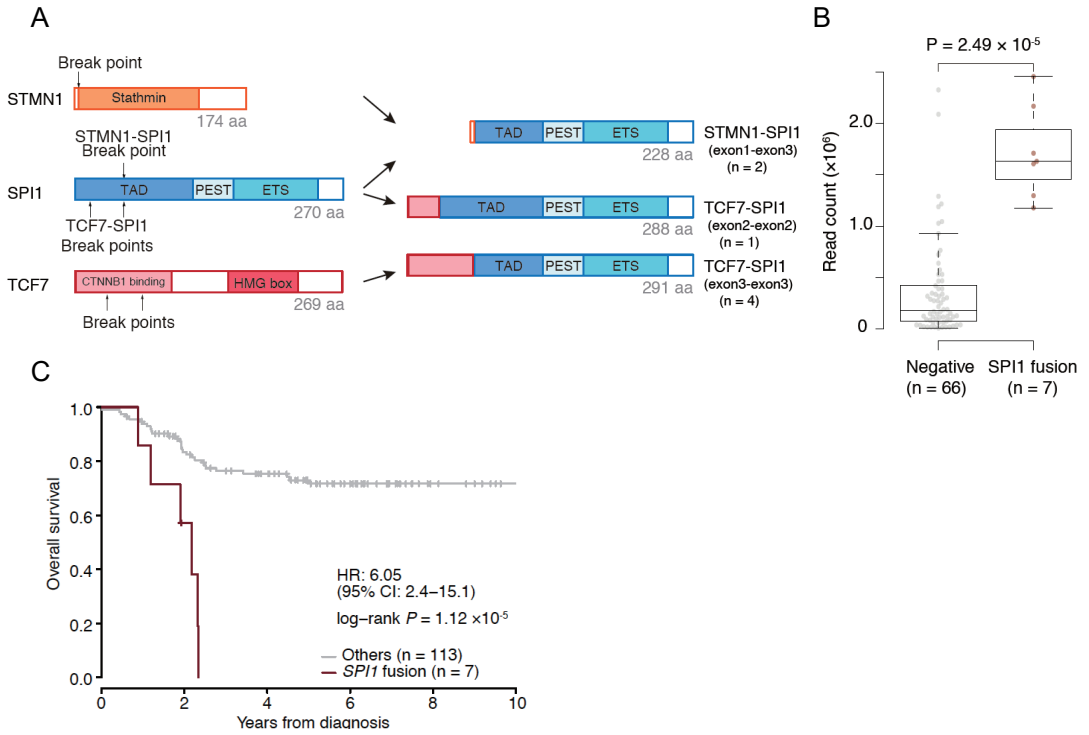


Figure 1 *SPI1* fusions in pediatric T-ALL
(A) Schematic representation of *SPI1* fusions. (B) Comparison of expression levels of *SPI1* between *SPI1* fusion-positive cases (n=7) and negative cases (n=116). (C) Kaplan-Meier survival curves of overall survival for *SPI1* fusion positive cases (n=7) and negative cases (n=113).

samples. To address the functional activities of *SPI1* fusions, we performed luciferase assays in HeLa cells carrying a reporter containing the *SPI1*-responsive promoter sequence. In comparison to empty vector, transfection with the *SPI1* fusion constructs resulted in a marked increase in luciferase activity comparable to that of wild type *SPI1*. This increase was almost completely absent when the sequence encoding the DNA-binding ETS domain was deleted from *SPI1*. This result suggested that *SPI1* fusions retain the transcriptional activity comparable to wild type *SPI1*. Then next, we evaluated the effect of *SPI1* fusions on T cell proliferation. DN T cells from wild-type mice thymus were transduced with constructs encoding *SPI1* fusions and seeded to in vitro cell culture to examine the effect of *SPI1* fusions on cell proliferation. *SPI1* fusion transduced cells showed significantly higher proliferation rates than cells transduced with empty vector. We next investigated the effects of *SPI1* fusion on T cell development. Mouse DN1 and DN2 thymocytes were transduced to express wild-type *SPI1* or each *SPI1* fusion and cultured on TSt4/DLL1 mouse stromal cells. Cells expressing SPI1 fusions and those expressing wild-type SPI1 showed a

differentiation block during DN T cell development. There were significantly reduced numbers of DP, SP, and DN4 T cells in comparison to control cells, with an increase in the number of immature DN1 and DN2 T cells. We also confirmed similar results by transplantation of SPI1 fusion-transduced primary bone marrow cells.

To characterize *SPI1*-fusion-positive T-ALL, we investigated gene expression profile of our T-ALL cases. Using two-step unsupervised consensus clustering, we obtained 5 stable clusters. Among these clusters, four clusters, TAL1-RA, -RB, TLX, and ETP, largely recapitulated distinct T-ALL subtypes characterized in previous expression array studies. However, the remaining cluster was newly identified and consisted of only SPI1-fusion-positive cases, suggesting that these cases represent a unique subtype of pediatric T-ALL distinguishable from the known T-ALL subtypes. Although their small sample size precluded accurate evaluation, *SPI1* fusion-positive T-ALL shared several genetic abnormalities with other T-ALL subtypes, such as frequent *NOTCH1* mutations and *CDKN2A* deletions, while other mutations commonly found in other T-ALL subtypes did not seem to be frequent, except for altered *RAS* pathway genes, which were found in 57% cases with *SPI1* fusions.

Finally, we also evaluated the effects of *SPI1* fusions on the clinical outcomes. When compared to *SPI1* fusion negative cases or cases in non *SPI* fusion clusters, *SPI1* fusion positive cases showed significantly shorter overall survival, with a median survival time of 2.2 years (Fig. 1B). Six out of the seven cases died within 3 years from diagnosis with early relapse.

In conclusion, we have described novel recurrent fusions involving *SPI1* associated with pediatric T-ALL. Exhibiting unique cytological and gene expression profiles, *SPI1* fusion-positive T-ALL had a uniformly dismal clinical outcome. To the best of our knowledge, these are the first genetic lesions associated with a very poor prognosis in pediatric T-ALL, although their impacts on survival need to be confirmed in additional cases. Patients with *SPI1* fusions seem to be incurable with current standard chemotherapy, which underscores the importance of detecting this subset of patients for more intensive therapy, including allogeneic stem-cell transplantation, to improve their clinical outcomes.

References

1. Herold, R., von Stackelberg, A., Hartmann, R., Eisenreich, B. & Henze, G. Acute lymphoblastic leukemia-relapse study of the Berlin-Frankfurt-Munster Group (ALL-REZ BFM) experience: early treatment intensity makes the difference. J Clin Oncol 2004; 22, 569-570.

2. Weng AP, Ferrando AA, Lee W, Morris JP 4th, Silverman LB, Sanchez-Irizarry C, Blacklow SC, Look AT, Aster JC. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 2004; 306:269-271.
3. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, Behm FG, Pui CH, Downing JR, Gilliland DG, Lander ES, Golub TR, Look AT. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell 2002; 1, 75-87.
4. Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, Etchin J, Lawton L, Sallan SE, Silverman LB, Loh ML, Hunger SP, Sanda T, Young RA, Look AT. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element.Science. 2014; 346:1373-1377.
5. Seki M, Kimura S, Isobe T, Yoshida K, Ueno H, Nakajima-Takagi Y, Changshan Wang, Lin L, Kon A, Suzuki H, Shiozawa Y, Kataoka K, Fujii Y, Shiraishi Y, Chiba K, Tanaka H, Shimamura T, Masuda K, Kawamoto H, Ohki K, Kato M, Arakawa Y, Koh K, Hanada R, Moritake H, Akiyama M, Kobayashi R, Deguchi T, Hashii Y, Imamura T, Sato A, Kiyokawa A, Oka A, Hayashi Y, Takagi M, Manabe A, Ohara A, Horibe K, Sanada M, Iwama A, Mano H, Miyano S, Ogawa S, and Takita J. Recurrent PU.1 *(SPI1)* fusions in high-risk pediatric T cell acute lymphoblastic leukemia. Nature Genet. 2017; 49:1274-1281.

Junko Takita, MD, PhD

1991-1992	Resident, Tokyo University Hospital
1992-1993	Resident, Yaizu City Hospital
1993-1996	Research Resident, National Cancer Research Institute
2003-2005	Assistant Professor, Tokyo University Hospital
2005-2013	Lecutuer, Tokyo University Hospital
2013-2018	Associate Professor, Tokyo University Hospital
2018-present	Professor, Kyoto University

DECODING CHALLENGING CANCERS

Sean Grimmond
Center for Cancer Research
University of Melbourne
Victorian Comprehensive Cancer Centre
Level 10, 305 Grattan Street, Melbourne, VIC 3000, Australia
(sean.grimmond@unimelb.edu.au)

Over the past 8 years, Prof Grimmond has led Australia's International Cancer Genome Consortium (ICGC) efforts, carrying out integrated whole-genome & transcriptome-based surveys of large Pancreatic, Ovarian, Neuroendocrine, Esophageal cancer and Melanoma cohorts and using them to determine the major root causes of somatic mutation, the common somatic mutations driving oncogenesis, mapping out clinically-relevant molecular taxonomies and providing a global survey of readily actionable mutations in this disease [1-4].

The greatest area of focus of the Australian ICGC effort has been on Pancreatic Cancer. Pancreatic Adenocarcinoma (PAAD) is currently the 4th leading cause of cancer death in Western societies and is projected to become the 2nd leading cause within a decade [5]. Having a median survival measured in months, 5-year survival prospects of 7%, and therapeutic advances over the last 30+ years on providing only incremental improvement, there is an urgent need to better understand the molecular pathology and oncogenesis of Pancreatic Cancer to improve patient selection for current treatments, and to develop novel therapeutic strategies.

These findings have been the impetus behind molecularly-targeted PAAD clinical trials [6] and provided foundations for in-silico modelling of cancer evolution and tumour cellular make up [1]. While the genome-exploratory efforts of the ICGC and The Cancer Genome Atlas (TCGA) have revolutionized our understanding of many cancers, these studies are far from complete. The original experimental design adopted had significant limitations: - (i) small cohort sizes and higher than expected somatic mutation rates left many studies underpowered, compromising the comprehensive detection of infrequent but

potent mutations in key genes, (ii) a reliance on exome sequencing left us blind to the role of non-coding / regulatory driver mutations, (iii) genomic analyses on a single early -stage tumour sample per patient hampered on our ability to study intramural heterogeneity and the temporal order of somatic damage, (iv) a reliance on standard bulk tumour tissue sequencing confounded our ability to accurately assign genomic events and transcriptomic features to specific cellular compartments – tumour vs. stroma – within malignancies.

Given these challenges, our recent activities have focused on :- i)Dramatically expanding our PAAD mutational atlas to include > 1000 whole genomes, exomes and transcriptomes of pancreatic pre-malignant lesions, primary resected tumours and metastases, ii) Studying uncommon molecular subtypes in pancreatic cancer (such as KRAS wildtype and long term survivors) iii) Use single cell sequencing approaches to resolve intra-tumour genomic and transcriptomic heterogeneity to define the cellular composition and cell-specific activation states during PAAD progression and iv) coupling genomic analysis of prospective pancreatic cancers drug sensitivity screening using patient matched organoid models.

Early ICGC-based studies into Pancreatic Cancer revealed a complex mutational landscape. While the classical instigators of premalignant progression (KRAS, CDKN2A, TP53 and SMAD4 mutations) are common place, they frequently occur in combination with inactivating mutations in ARID1A, TGFBR2, RNF43, KDM6A (5-15% of patients). Furthermore, studies of 450 tumours via exome and whole genome analysis discovered a "long tail" of 40 genes under selection, attracting highly-impactful mutations at low prevalence (1-5% of patients). Focal copy number changes and chromosomal rearrangements, often caused by genomic catastrophes, add additional key tumour suppressor (TP53, CDKN2A) and oncogenes (e.g. GATA6, CCNE1, ERRB2, MET, MYC, MIB1) [1] to this landscape. Intriguingly, the function of these recurrently mutated genes fall into 12 core processes we now consider central to Pancreatic Cancer tumorigenesis.

In an effort to see whether we had saturated driver gene discovery based on position selection / recurrent mutation, we expanded this study to include the CAN-ICGC &TCGA alongside the expanded AUS-ICGC cohort identifies 60 significantly mutated genes. When these genes are viewed in the context of "pathways", we find the mutated genes under selection fall into MAP kinase signalling, TGF-beta signalling, WNT signalling, NOTCH signalling, ROBO-SLIT signalling, Cell Cycle, DNA damage repair, the SWI-SNF complex, Chromatin modification, and RNA processing.

With so much of early ICGC & the TCGA's focus on exome rather than whole-genome sequencing, the exploration of noncoding driver mutations has been limited. CIA-Grimmond and colleagues recently carried out a genome-wide study of recurrent mutations in known regulatory regions resulting in coordinate perturbation of proximal

gene expression in 305 PAAD whole cancer genomes from the Australian and Canadian ICGC-PAAD efforts [7]. This study showed that mutations in regulatory elements is a genuine alternate mechanism for damaging core pancreatic cancer pathways like WNT signalling and Axon guidance, as well as impacting on new processes like Cell adhesion, Regulation of transcription and Homeobox regulation.

More recently, others have shown that structural rearrangements are capable of activating cancer-promoting genes through promoting ectopic enhancer activity [8,9]. Exploratory studies into these so called "enhancer hijacking" events within the AUS-PAAD cohort have found recurrent duplication and amplification of the super enhancer located upstream of the MYC oncogene which have been shown to promote its overexpression in breast cancer [9]. Taken together these exploratory findings warrant analysis of larger cohorts to validate these early explorations and appropriately power deeper investigations.

Studying driver gene composition in clinically significant segments of the PAAD cohort: This cohort will be suitably powered to investigate the underlying genetics in two uncommon but clinically significant segments: - (i) KRAS+/+ PAADs. Making up only 7 % of all PAAD, these tumours must rely on an alternate oncogenic mutation to drive them. In the recent TCGA study [10], 3/11 KRAS +/+ PAADs possessed oncogenic BRAF fusions & mutations. This cohort contains WGS/WTS for >100 KRAS wildtype tumours allowing us to robustly screen for oncogenic events (gain of function mutations, gene fusions and gene amplifications) and contains a wealth of MAPK signalling gain-of-function mutations (eg BRAF fusions & in frame indels, NRG fusions, RAF1 fusions, N and HRAS hotspot mutations) opening up the opportunity to re-purpose BRAFi strategies or EGFRi targeting of MAPK independent tumours as testable therapeutic avenues KRAS+/+ tumours.

Over recent years, the TCGA and ICGC efforts have also demonstrated that global RNA expression profiling is a powerful way to gain insight into the consequences of somatic mutation, the transcriptional programs promoting malignancy and molecular taxonomies present in otherwise histologically-homogenous cancers. We recently carried out a global analysis of bulk tissue from >260s resectable ICGC PAADs via Whole Transcriptome Sequencing (WTS) and/or microarray expression profiling and identified 4 robust PAAD expression subtypes [5], now known as: Pancreatic Progenitor, ADEX (Abnormally Differentiated-Endocrine & eXocrine), Squamous, and Immunogenic PAADs. While traditional whole-tissue RNAseq studies focus on deconvolving cell-specific gene expression patterns from the admixtures of tumour, stromal, immune, endothelial cell etc. expression in a tumour. Single-cell sequencing technology now makes this possible. We are using single-nucleus RNASeq from cryopreserved tissues to study surplus material from the extensively characterised PAADs studied as part of the AUS_ICGC program.

Given that PAAD is such an aggressive malignancy (median survival is 7 months), testing multiple therapeutic strategies in patients is rarely possible. As a consequence, systematic screening of drug sensitivity in patient-matched, ex-vivo tumour models has been proposed as an alternative. Prof Grimmond and colleagues have previously generated xenograft models for patients consented into the Australian ICGC program. Unfortunately, the lag time and costs render them clinically impractical.

Organoid tumour models are an increasing popular alternative for solid tumours Over the last 2 years, Prof Grimmond has led an Avner Foundation supported program with researchers from University of Melbourne &WEHI to establish generation of pancreatic cancer organoid cultures, carry genomic analysis and plate-based drug sensitivity profiling to the standard of care agents within 6-8 weeks of biopsy (See Figure 1). Where extreme drug sensitivities are observed, the study's ethical framework permits return data to patients. Similar to published studies, the models maintain genomic fidelity to their primary lesion their drug sensitivity profile recapitulates patient therapeutic responses. The next step for this program is to escalate the scale of drug screening performed on each patient-matched organoid generated, creating a much-needed knowledgebase of the spectrum of responses to dozens-100s of agents, which includes all *standard of care* drugs, *re-purposed cancer drugs* for which actionable mutations have been seen in PAAD cancer atlases novel agents targeting potential *druggable leads* that are identified in the previous aims. This will determine how the allelic series of mutations seen in driver genes, and their common combinations influence drug sensitivity.

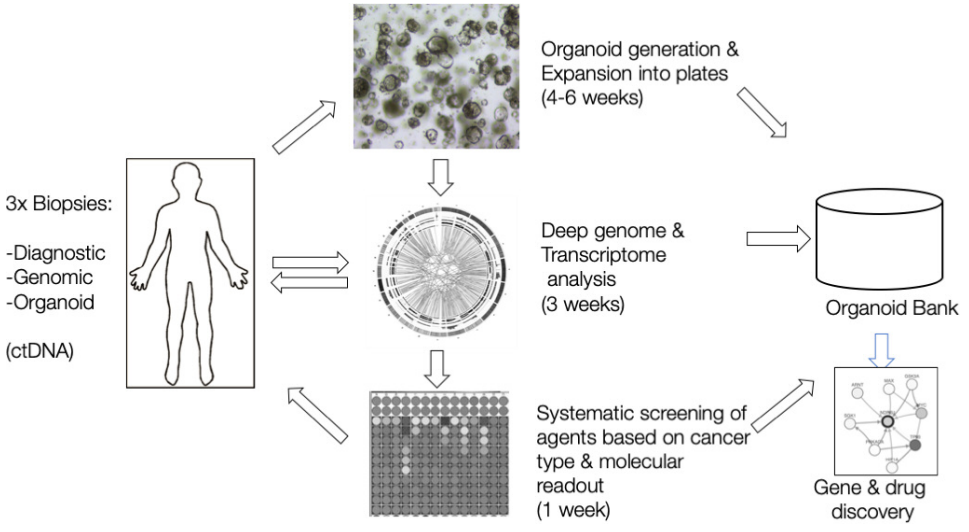


Figure 1

References

1. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, Nourse C, Murtaugh L, Harliwong I, Idrisoglu S, Manning S, Nourbakhsh E, Wani S, Fink L, Holmes O, Chin V, Anderson MJ, Kazakoff S, Leonard C, Newell F, Waddell N, Wood S, Xu Q, Wilson PJ, Cloonan N, Kassahn KS, Taylor D, Quek K, Robertson A, Pantano L, Mincarelli L, Sanchez LN, Evers L, Wu J, Pinese M, Cowley MJ, Jones MD, Colvin EK, Nagrial AM, Humphrey ES, Chantrill LA, Mawson A, Humphris J, Chou A, Pajic M, Scarlett CJ, Pinho AV, Giry-Laterriere M, Rooman I, Samra JS, Kench JG, Lovell JA, Merrett ND, Toon CW, Epari K, Nguyen NQ, Barbour A, Zeps N, Moran-Jones K, Jamieson NB, Graham JS, Duthie F, Oien K, Hair J, Grützmann R, Maitra A, Iacobuzio-Donahue CA, Wolfgang CL, Morgan RA, Lawlor RT, Corbo V, Bassi C, Rusev B, Capelli P, Salvia R, Tortora G, Mukhopadhyay D, Petersen GM; Australian Pancreatic Cancer Genome Initiative, Munzy DM, Fisher WE, Karim SA, Eshleman JR, Hruban RH, Pilarsky C, Morton JP, Sansom OJ, Scarpa A, Musgrove EA, Bailey UM, Hofmann O, Sutherland RL, Wheeler DA, Gill AJ, Gibbs RA, Pearson JV, Waddell N, Biankin AV, Grimmond SM. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature. 2016;531:47-52.
2. Patch AM, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, Nones K, Cowin P, Alsop K, Bailey PJ, Kassahn KS, Newell F, Quinn MC, Kazakoff S, Quek K, Wilhelm-Benartzi C, Curry E, Leong HS; Australian Ovarian Cancer Study Group, Hamilton A, Mileshkin L, Au-Yeung G, Kennedy C, Hung J, Chiew YE, Harnett P, Friedlander M, Quinn M, Pyman J, Cordner S, O'Brien P, Leditschke J, Young G, Strachan K, Waring P, Azar W, Mitchell C, Traficante N, Hendley J, Thorne H, Shackleton M, Miller DK, Arnau GM, Tothill RW, Holloway TP, Semple T, Harliwong I, Nourse C, Nourbakhsh E, Manning S, Idrisoglu S, Bruxner TJ, Christ AN, Poudel B, Holmes O, Anderson M, Leonard C, Lonie A, Hall N, Wood S, Taylor DF, Xu Q, Fink JL, Waddell N, Drapkin R, Stronach E, Gabra H, Brown R, Jewell A, Nagaraj SH, Markham E, Wilson PJ, Ellul J, McNally O, Doyle MA, Vedururu R, Stewart C, Lengyel E, Pearson JV, Waddell N, deFazio A, Grimmond SM, Bowtell DD. Whole Genome characterization of chemoresistant ovarian cancer. Nature. 2015;521:489-494.
3. Scarpa A, Chang DK, Nones K, Corbo V, Patch AM, Bailey P, Lawlor RT, Johns AL, Miller DK, Mafficini A, Rusev B, Scardoni M, Antonello D, Barbi S, Sikora KO, Cingarlini S, Vicentini C, McKay S, Quinn MC, Bruxner TJ, Christ AN, Harliwong I, Idrisoglu S, McLean S, Nourse C, Nourbakhsh E, Wilson PJ, Anderson MJ, Fink JL, Newell F, Waddell N, Holmes O, Kazakoff SH, Leonard C, Wood S, Xu Q, Nagaraj SH, Amato E, Dalai I, Bersani S, Cataldo I, Dei Tos AP, Capelli P, Davì MV, Landoni L, Malpaga A, Miotto M, Whitehall VL, Leggett BA, Harris JL, Harris J, Jones MD,

Humphris J, Chantrill LA, Chin V, Nagrial AM, Pajic M, Scarlett CJ, Pinho A, Rooman I, Toon C, Wu J, Pinese M, Cowley M, Barbour A, Mawson A, Humphrey ES, Colvin EK, Chou A, Lovell JA, Jamieson NB, Duthie F, Gingras MC, Fisher WE, Dagg RA, Lau LM, Lee M, Pickett HA, Reddel RR, Samra JS, Kench JG, Merrett ND, Epari K, Nguyen NQ, Zeps N, Falconi M, Simbolo M, Butturini G, Van Buren G, Partelli S, Fassan M; Australian Pancreatic Cancer Genome Initiative, Khanna KK, Gill AJ, Wheeler DA, Gibbs RA, Musgrove EA, Bassi C, Tortora G, Pederzoli P, Pearson JV, Waddell N, Biankin AV, Grimmond SM. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature. 2017;543:65–71.

4. Hayward NK, Wilmott JS, Waddell N, Johansson PA, Field MA, Nones K, Patch AM, Kakavand H, Alexandrov LB, Burke H, Jakrot V, Kazakoff S, Holmes O, Leonard C, Sabarinathan R, Mularoni L, Wood S, Xu Q, Waddell N, Tembe V, Pupo GM, De Paoli-Iseppi R, Vilain RE, Shang P, Lau LMS, Dagg RA, Schramm SJ, Pritchard A, Dutton-Regester K, Newell F, Fitzgerald A, Shang CA, Grimmond SM, Pickett HA, Yang JY, Stretch JR, Behren A, Kefford RF, Hersey P, Long GV, Cebon J, Shackleton M, Spillane AJ, Saw RPM, López-Bigas N, Pearson JV, Thompson JF, Scolyer RA, Mann GJ. Whole-genome landscapes of major melanoma subtypes. Nature. 2017; 545:175-180.
5. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer research. 2014;74:2913-2921.
6. Chantrill L, Nagrial A, Watson C, Johns A, Martyn-Smith M, Simpson S, Mead S, Jones M, Samra JS, Gill A, Watson N, Chin V, Humphris J, Chou A, Brown B, Morey A, Pajic M, Grimmond SM, Chang DK, Thomas D, Sebastian L, Sjoquist K, Yip S, Pavlakis N, Asghari R, Harvey S, Grimison P, Simes J, Biankin AV. Precision Medicine for Advanced Pancreas Cancer: The Individualized Molecular Pancreatic Cancer Therapy (IMPaCT) Trial. Clin Cancer Res. 2017;21:2029-2037.
7. Feigin ME, Garvin T, Bailey P, Waddell N, Chang DK, Kelley DR, Shuai S, Gallinger S, McPherson JD, Grimmond SM, Khurana E, Stein LD, Biankin AV, Schatz MC, Tuveson DA. Recurrent noncoding regulatory mutations in pancreatic ductal adenocarcinoma. Nat Genet 2017;49:825-833.
8. Northcott PA, Lee C, Zichner T, Stütz AM, Erkek S, Kawauchi D, Shih DJ, Hovestadt V, Zapatka M, Sturm D, Jones DT, Kool M, Remke M, Cavalli FM, Zuyderduyn S, Bader GD, VandenBerg S, Esparza LA, Ryzhova M, Wang W, Wittmann A, Stark S, Sieber L, Seker-Cin H, Linke L, Kratochwil F, Jäger N, Buchhalter I, Imbusch CD, Zipprich G, Raeder B, Schmidt S, Diessl N, Wolf S, Wiemann S, Brors B, Lawerenz C, Eils J, Warnatz HJ, Risch T, Yaspo ML, Weber UD, Bartholomae CC, von Kalle C, Turányi E, Hauser P, Sanden E, Darabi A, Siesjö P, Sterba J, Zitterbart K, Sumerauer D, van Sluis P, Versteeg R, Volckmann R, Koster J, Schuhmann MU, Ebinger M, Grimes HL, Robinson GW, Gajjar A,

Mynarek M, von Hoff K, Rutkowski S, Pietsch T, Scheurlen W, Felsberg J, Reifenberger G, Kulozik AE, von Deimling A, Witt O, Eils R, Gilbertson RJ, Korshunov A, Taylor MD, Lichter P, Korbel JO, Wechsler-Reya RJ, Pfister SM. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature. 2014;511:428-434.

9. Glodzik D, Morganella S, Davies H, Simpson PT, Li Y, Zou X, Diez-Perez J, Staaf J, Alexandrov LB, Smid M, Brinkman AB, Rye IH, Russnes H, Raine K, Purdie CA, Lakhani SR, Thompson AM, Birney E, Stunnenberg HG, van de Vijver MJ, Martens JW, Børresen-Dale AL, Richardson AL, Kong G, Viari A, Easton D, Evan G, Campbell PJ, Stratton MR, Nik-Zainal S. A somatic-mutational process recurrently duplicates germline susceptibility loci and tissue-specific super-enhancers in breast cancers. Nat Genet. 2017;49:341-348.
10. Cancer Genome Atlas Consortium: Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell. 2017; 32:185-203.

Sean Grimmond, BSc, PhD, FFS, RCPA

2009-2014	Director, Queensland Centre for Medical Genomics
2013-2015	Chair of Medical Genomics, University of Glasgow
2014-2015	Co-director of the Scottish Genome Partnership
2016-present	Director & The Bertalli Chair in Cancer Medicine University of Melbourne Centre for Cancer Research Head, Department of Clinical Pathology University of Melbourne Centre for Cancer Research

COMPREHENSIVE GENOMIC ANALYSIS OF RARE GASTROENTEROLOGICAL CANCERS

Shinichi Yachida
Department of Cancer Genome Informatics
Graduate School of Medicine, Osaka University
2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
(syachida@cgi.med.osaka-u.ac.jp)

Comprehensive genetic and epigenetic analyses for major cancers have been performed as TCGA (The Cancer Genome Atlas) and ICGC (International Cancer Genome Consortium) projects. We have learned a lot of organ-specific and non-organ-specific alterations of cancer genome from these projects. However, genomic features of rare cancers are still unknown.

Among rare cancers, we conducted an in-depth genomic analysis on ampullary carcinomas from Japanese and American patient cohorts [1]. Carcinoma of the ampulla of Vater is a highly malignant neoplasm. Three distinct epithelial linings (duodenal, biliary, and pancreatic) converge at the ampulla of Vater, with pancreatic and biliary epithelium merging within the ampulla of Vater to form the epithelium of the ampulla. Ampullary carcinomas can be classified into two histological phenotypes, intestinal-type and pancreatobiliary-type. Carcinomas of the pancreatobiliary subtype are found to be more aggressive than those of the intestinal subtype in most studies [2]. These phenotypes have different pathogenic and clinical characteristics.

To gain insight into the genetic basis of this tumor type, we performed the exome sequencing in a discovery set of 60 ampullary carcinomas and ten duodenal carcinomas. We next selected 92 genes which were recurrently altered in the discovery screen, or which were well-documented components of a pathway, or which were potentially targetable, since alterations in these genes are most likely to be clinically relevant. In total, 172 ampullary carcinomas and 18 non-ampullary duodenal carcinomas were investigated.

Twenty-four genes were significantly mutated driver genes in 172 ampullary carcinomas. We identified a characteristic significantly mutated driver gene (*ELF3*) as well

as previously known driver genes (*TP53*, *KRAS*, *APC*, and others) (Figure). We compared significantly mutated genes between Japanese and American patients. We found that *ELF3* mutations occur in ampullary carcinomas across racial lines. There were differences between the genomic landscapes of the intestinal phenotype and those of the pancreatobiliary phenotype. Among the significantly mutated genes, high-ranking genes based on the prevalence of mutations were similar between intestinal-type ampullary carcinomas and colorectal carcinomas, i.e., *APC*, *TP53*, *KRAS* and between pancreaticobiliary-type carcinomas and pancreatic carcinomas, i.e., *KRAS*, *TP53*, *SMAD4*.

Functional studies demonstrated that *ELF3* silencing in normal human epithelial cells enhances their motility and invasion. Since an immortalized normal epithelial cell line has not been established from ampullary cells, we used an immortalized normal epithelial cell line of common bile duct origin, designated HBDEC2-3H10, and an immortalized normal epithelial cell line of duodenal mucosa origin, designated HDuodEC3. These lines were selected for functional analyses because *ELF3* mutations have also been observed in 7/74 (9.5%) common bile duct carcinomas in our study [3] and 1/18 (5.6%) duodenal carcinomas in the present study. To investigate the consequences of the loss-of-function mutation in *ELF3*, three human *ELF3*-specific small interfering RNA (siRNA) oligonucleotides were utilized to knock down ELF3 expression in the HBDEC2-3H10 cells. Invasion/migration assay using Matrigel invasion chambers and control inserts demonstrated that invasive activities and motilities in *ELF3* knockdown cells were significantly increased compared with control cells. HDuodEC3 cells treated with ELF3 siRNAs showed similar phenotypic changes in terms of cell invasion and motility. Consistent with the present data, aggressive invasion phenotype (extended cell bodies into the Matrigel matrix) of *ELF3* knockdown cells was observed in time-lapse images of 3D cell invasion assay. Quantitative RT-PCR analysis for the expression of matrix metalloproteinase-1 and -9 (MMP1 and MMP9) further supported this observation, showing higher expression levels of MMP1 and MMP9 in *ELF3* knockdown cells compared with control cells. Knockdown of *ELF3* is associated with epithelial-to-mesenchymal transition (EMT): immunofluorescence and quantitative RT-PCR analysis showed that the expression of vimentin, which is a mesenchymal marker of EMT and a regulator of cell migration, was increased in *ELF3* knockdown cells compared with the control cells. By contrast, the expression of the epithelial marker, cytokeratin 19 (CK19), was decreased in cells with ELF3 deficiency. In addition, key regulators of EMT, such as ZEB1, ZEB2, and TWIST1, were upregulated in *ELF3* knockdown cells.

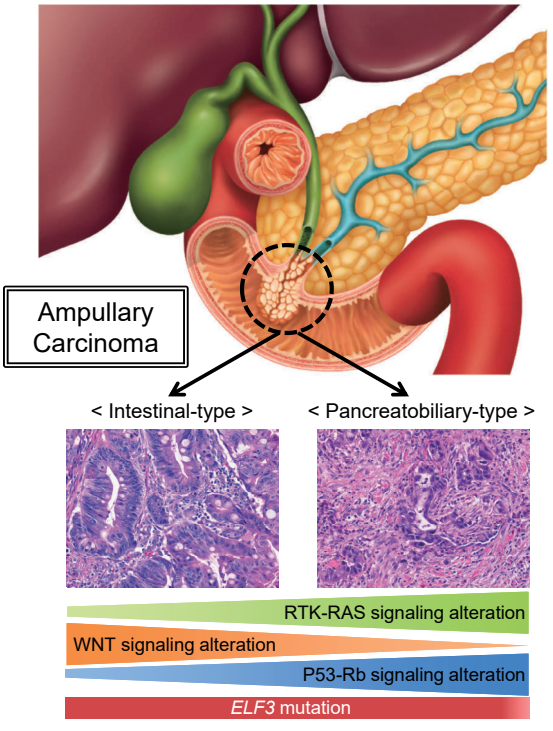
Neuroendocrine neoplasms (NENs) are rare and heterogenous malignancies, which arise across organs (e.g., esophagus, stomach, duodenum, pancreas, bile ducts and colorectum), with two major subtypes: neuroendocrine tumor (NET) and neuroendocrine carcinoma (NEC). Previous studies have shown that NETs and NECs have distinct

prognoses and responses to treatment. The genomic features of pancreatic NET have been identified, then *DAXX*/*ATRX*, *MEN1*, and mTOR pathway are frequently altered [4,5]. The histology of pancreatic NECs is similar to that of small/large cell lung carcinomas. We previously investigated alterations of *KRAS*, *CDKN2A*/*p16*, *TP53*, *SMAD4*/*DPC4*, *DAXX*, *ATRX*, *PTEN*, *Bcl2*, and *RB1* by immunohistochemistry and/or targeted exomic sequencing in 19 pancreatic NECs, indicating small cell NECs of pancreas are genetically similar to large cell NECs, and these genetic changes are distinct from those reported pancreatic NETs [6].

Currently, we are focusing on gastroenterological NECs. While in this process, we have performed an in-depth analysis of the genomic abnormalities of these carcinomas through a multi-center collaboration to establish a potential basis for treatments of this disease.

References

1. Yachida S, Wood LD, Suzuki M, Takai E, Totoki Y, Kato M, Luchini C, Arai Y, Nakamura H, Hama N, Elzawahry A, Hosoda F, Shirota T, Morimoto N, Hori K, Funazaki J, Tanaka H, Morizane C, Okusaka T, Nara S, Shimada K, Hiraoka N, Taniguchi H, Higuchi R, Oshima M, Okano K, Hirono S, Mizuma M, Arihiro K, Yamamoto M, Unno M, Yamaue H, Weiss MJ, Wolfgang CL, Furukawa T, Nakagama H, Vogelstein B, Kiyono T, Hruban RH, Shibata T. Genomic sequencing identifies ELF3 as a



driver of ampullary carcinoma. Cancer Cell. 2016;29: 229-240.

2. Okano K, Oshima M, Yachida S, Kushida Y, Kato K, Kamada H, Wato M, Nishihira T, Fukuda Y, Maeba T, Inoue H, Masaki T, Suzuki Y. Factors predicting survival and pathological subtype in patients with ampullary adenocarcinoma. J Surg Oncol. 2014;110: 156-162.
3. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, Hama N, Hosoda F, Urushidate T, Ohashi S, Hiraoka N, Ojima H, Shimada K, Okusaka T, Kosuge T, Miyagawa S, Shibata T. Genomic spectra of biliary tract cancer. Nat Genet. 2015;47: 1003-10.
4. Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti MA, Velculescu VE, Diaz LA Jr, Vogelstein B, Kinzler KW, Hruban RH, Papadopoulos N. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science 2011; 331: 1199-1203.
5. Scarpa A, Chang DK, Nones K, Corbo V, Patch AM, Bailey P, Lawlor RT, Johns AL, Miller DK, Mafficini A, Rusev B, Scardoni M, Antonello D, Barbi S, Sikora KO, Cingarlini S, Vicentini C, McKay S, Quinn MC, Bruxner TJ, Christ AN, Harliwong I, Idrisoglu S, McLean S, Nourse C, Nourbakhsh E, Wilson PJ, Anderson MJ, Fink JL, Newell F, Waddell N, Holmes O, Kazakoff SH, Leonard C, Wood S, Xu Q, Nagaraj SH, Amato E, Dalai I, Bersani S, Cataldo I, Dei Tos AP, Capelli P, Davì MV, Landoni L, Malpaga A, Miotto M, Whitehall VL, Leggett BA, Harris JL, Harris J, Jones MD, Humphris J, Chantrill LA, Chin V, Nagrial AM, Pajic M, Scarlett CJ, Pinho A, Rooman I, Toon C, Wu J, Pinese M, Cowley M, Barbour A, Mawson A, Humphrey ES, Colvin EK, Chou A, Lovell JA, Jamieson NB, Duthie F, Gingras MC, Fisher WE, Dagg RA, Lau LM, Lee M, Pickett HA, Reddel RR, Samra JS, Kench JG, Merrett ND, Epari K, Nguyen NQ, Zeps N, Falconi M, Simbolo M, Butturini G, Van Buren G, Partelli S, Fassan M; Australian Pancreatic Cancer Genome Initiative, Khanna KK, Gill AJ, Wheeler DA, Gibbs RA, Musgrove EA, Bassi C, Tortora G, Pederzoli P, Pearson JV, Waddell N, Biankin AV, Grimmond SM. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature 2017;543: 65-71.
6. Yachida S, Vakiani E, White CM, Zhong Y, Saunders T, Morgan R, de Wilde RF, Maitra A, Hicks J, Demarzo AM, Shi C, Sharma R, Laheru D, Edil BH, Wolfgang CL, Schulick RD, Hruban RH, Tang LH, Klimstra DS, Iacobuzio-Donahue CA. Small cell and large cell neuroendocrine carcinomas of the pancreas are genetically similar and distinct from well-differentiated pancreatic neuroendocrine tumors. Am J Surg Pathol. 2012;36: 173-184.

Shinichi Yachida, MD, PhD

1988-1994	M.D., Tottori University
1994-1998	Ph.D. in Surgery, Kagawa University
1998-1999	Clinical Fellow, Kagawa University Hospital
1999-2002	Resident, National Cancer Center Hospital
2002-2007	Assistant Professor, Surgery, Kagawa University
2007-2010	Postdoctoral Fellow, Johns Hopkins Medical Institutions
2011-2012	Assistant Professor, Surgery, Kagawa University
2012-2017	Laboratory Head, National Cancer Center Research Institute
2017-present	Professor, Graduate School of Medicine, Osaka University

GENOMIC AND EPIGENOMIC PROFILES OF ASIAN ENDEMIC CANCERS

Patrick Tan
Programme in Cancer & Stem Cell Biology
Duke-NUS Medical School
8 College Road, Singapore 169857
(gmstanp@duke-nus.edu.sg)

In 2012, 14 million new cancer cases were diagnosed and there were 8.2 million cancer-related deaths. Of these, almost half (49%) of newly diagnosed cancer cases were from Asia, contributing to 55% of all global cancer mortality. This number is expected to dramatically increase in coming years, and will be disproportionately felt in Asia. Unfortunately, because different cancers can display tremendous geographic and regional variation, many of the major cancers relevant to Asia are different from those in Western countries, and for these Asian cancers comparatively little is known about their underlying molecular genetics. For example, both liver and gastric cancer are endemic to Asia but have few targeted treatment options. There are also certain types of cancer such as peripheral T-cell lymphomas, where Asian-dominant subentities (Natural Killer T-cell Lymphoma (NKTCL) are associated with dismal prognosis.

Notably, many malignancies with high prevalence in Asia are caused by exposures to carcinogens, such as infectious agents and chemical toxins. Such cancers provide important "natural experiments" for understanding how environmental perturbations can disrupt normal cellular processes to ultimately drive tumor development, at both the genetic and epigenetic level. Over the past decade, our lab has taken a "team science" approach towards the molecular dissection of Asian cancers, working with Prof Bin Tean Teh (National Cancer Centre Singapore) and Prof Steven Rozen (Duke-NUS Medical School Singapore). Our results have elucidated new cancer genes associated with biliary tract cancer [1-3], breast fibroepithelial tumors [4,5] and NKTCL [6]. We have also gained insights into how natural compounds such as aristolochic acid, which is found in certain traditional Chinese medicines, can contribute to urinary tract and liver cancers [7].

In this talk, I will focus on gastric cancer, which is highly endemic in Asia and a leading cause of global cancer mortality. Earlier work from our team revealed that epigenomic alterations are prevalent in gastric cancer, as revealed through driver mutations in chromatin modifier genes such as *ARID1A* (Figure 1, ref. 8). More recent work from our group has revealed a surprising link between epigenomic alterations in gastric cancer and the ability of tumors to evade host anti-tumor activity. Specifically, by mapping epigenomic histone modifications associated with gene promoters, we found that gastric cancers pervasively utilize alternative promoters, resulting in shortened proteins which exhibit lower immunogenicity (Figure 2, ref 9). We have also focused on the precision prevention of gastric cancer, by asking if molecular profiles of gastric pre-malignant conditions (intestinal metaplasia) can predict which patients will progress to gastric cancer development. This work was performed by leveraging on the Gastric Cancer Epidemiology

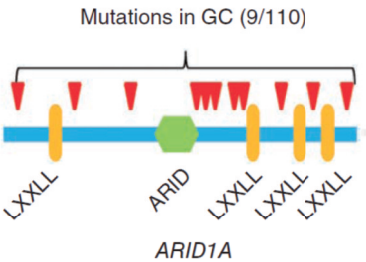


Figure 1 ARID1A mutations in gastric cancer. (a) Eight percent (9/110) of tumors harbor ARID1A somatic mutations. Triangles indicate inactivating mutations. ARID, ATrich interactive domain(green hexagon); LXXLL, C-terminal leucine-rich LXXLL motif (yellow ovals). From Zang et al., (2012).

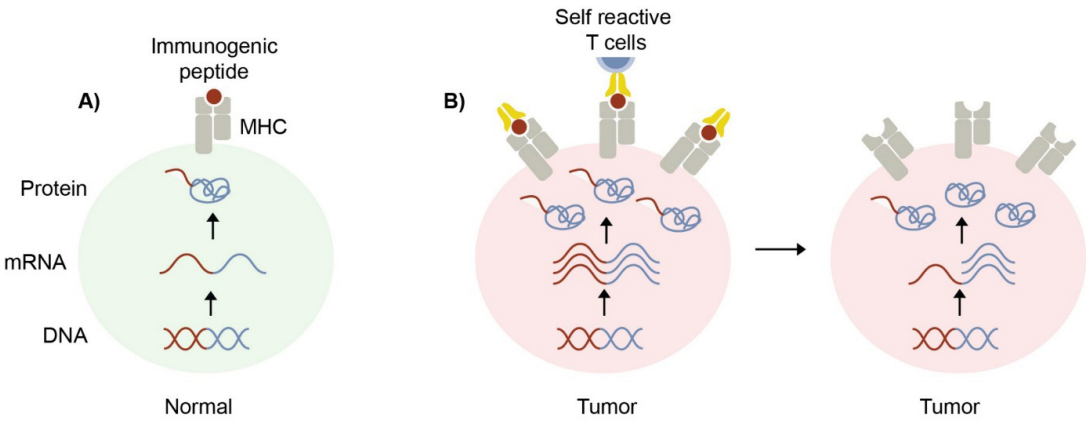


Figure 2 Epigenomic Alterations in Promoter Usage Facilitate Evasion of Host Immunity (from Qamra et al., 2017).

Program (GCEP), a 10-year prospective cohort study (Figure 3). By analyzing GCEP samples, we were able to identify a subgroup of patients with high-risk of gastric cancer progression, and patients with previously-undetected *Helicobacter pylori* infection, the causative agent of gastric cancer [10].

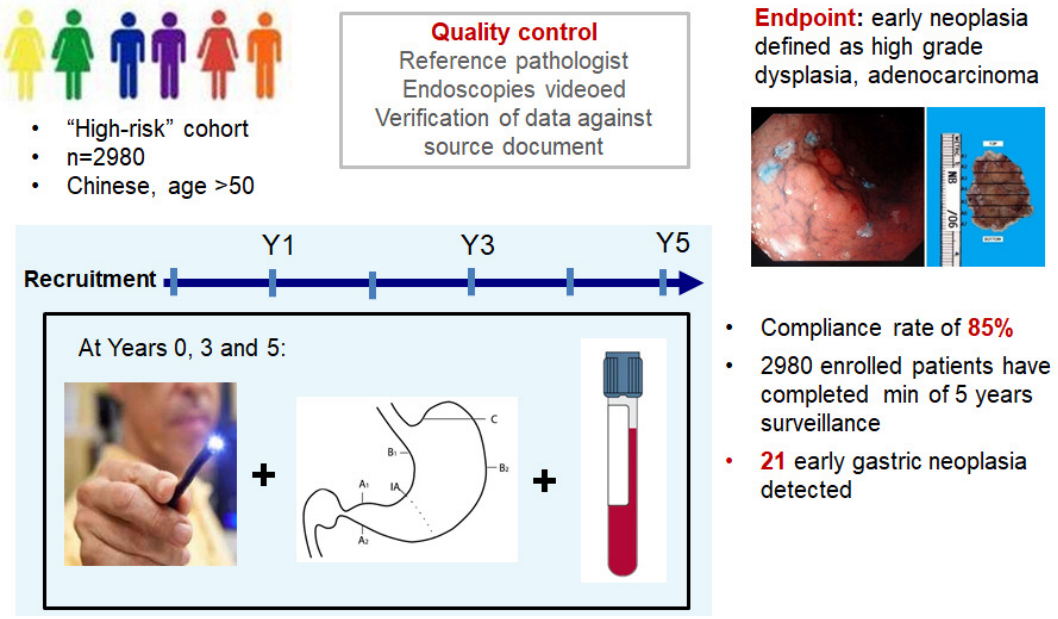
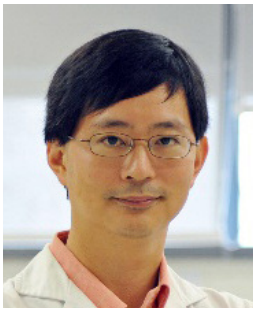


Figure 3 Overview of Gastric Cancer Epidemiology Program

References

1. Ong CK, Subimerb C, Pairojkul C, Wongkham S, Cutcutache I, Yu W, et al. Exome sequencing of liver fluke-associated cholangiocarcinoma. Nature Genetics. 2012 Jun;44(6):690-3
2. Chan-On W, Nairismagi ML, Ong CK, Lim WK, Dima S, Pairojkul C, et al. Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. Nature Genetics. 2013 Dec;45(12):1474-8
3. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, et al. Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. Cancer Discovery 2017 Oct;7(10):1116-1135.
4. Lim WK, Ong CK, Tan, Thike AA, Ng CC, Rajasegaran V, et al. Exome sequencing identifies highly recurrent MED12 somatic mutations in breast fibroadenoma. Nature Genetics. 2014 Aug; 46(8):877-80.

5. Tan J, Ong CK, Lim WK, Ng CC, Thike AA, et al. Genomic landscapes of breast fibroepithelial tumors. Nature Genetics. 2015 Nov;47(11):1341-5.
6. Koo GC, Tan SY, Tang T, Poon SL, Allen GE et al. Janus kinase 3-activating mutations identified in natural killer/T-cell lymphoma. Cancer Discovery. 2012 Jul;2(7):591-7
7. Poon SL, Pang ST, McPherson JR, Yu W, Huang KK, Guan P, et al. Genome-wide mutational signatures of aristolochic acid and its application as a screening tool. Science translational medicine. 2013 Aug 7;5(197):197ra01.
8. Zang ZJ, Cutcutache I, Poon SL, Zhang SL, McPherson JR, Tao J, et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. Nature genetics. 2012 May;44(5):570-4.
9. Qamra A, Xing M, Padmanabhan N, Kwok JJT, Zhang S, et al. Epigenomic Promoter Alterations Amplify Gene Isoform and Immunogenic Diversity in Gastric Adenocarcinoma. Cancer Discovery 2017 Jun;7(6):630-651.
10. Huang KK, Ramnarayanan K, Zhu F, Srivastava S, Xu C, et al. Genomic and Epigenomic Profiling of High-Risk Intestinal Metaplasia Reveals Molecular Determinants of Progression to Gastric Cancer. Cancer Cell. 2018 Jan 8;33(1):137-150.



Patrick Tan, MD, PhD

1988-1992	Harvard University, B.A. (Biochemistry)
1992-2000	Stanford University School of Medicine, M.D., Ph.D (Developmental Biology)
2006-present	Principal Investigator (Adjunct), National Cancer Centre, Singapore
2009-present	Director, Duke-NUS Genome Biology Facility
2012-present	Senior Principal Investigator, Cancer Science Institute of Singapore (NUS)
2012-present	Professor (Tenured), Cancer and Stem Cell Biology, Duke-NUS Medical School
2015-present	Director, SingHealth Duke-NUS Institute of Precision Medicine (PRISM)
2016-present	Deputy Executive Director, Biomedical Research Council, A*STAR

CODING AND NON-CODING DRIVERS IN >2,500 WHOLE CANCER GENOMES

Gad Getz
Cancer Center and Dept. of Pathology, Massachusetts General Hospital
55 Fruit St., Boston, MA 02114, USA
Cancer Program, Broad Institute of Harvard and MIT
415 Main St., Cambridge, MA 02142, USA
Harvard Medical School
250 Longwood Ave, Boston, MA 02115, USA
(gadgetz@broadinstitute.org)

Discovery of cancer genes has traditionally focused on the identification of protein-coding drivers [1]. In the past few years, a few non-coding drivers were discovered and functionally validated, including mutations in the promoter of *TERT* across many cancers [2] and promoter of *FOXA1* in breast cancer [3]. Here, I will present the work of the Drivers and Functional Interpretation Group [4] within the ICGC/TCGA Pan-Cancer Analysis of Whole Genomes (PCAWG) project [5], in which we performed a comprehensive analysis of putative cancer driver mutations in both protein-coding and non-coding genomic regions across >2,500 whole cancer genomes. We developed a statistically rigorous strategy for combining significance levels from multiple driver discovery methods that can overcome limitations of individual methods. Our analyses confirm previously reported elements, raise doubts about others, and identify novel candidate elements across 27 cancer types and 15 pan-cancer sets. Novel recurrent events were found in the promoters or 5'UTRs of *TP53*, *RNF34*, and *MTG2*; in the 3'UTRs of *NFKBIZ* and *TOB1*; and in the RNA gene *RMRP*. We provide evidence that recurrent mutations in the RNA genes *NEAT1* and *MALAT1* are subject to a localized mutational process and are not under positive selection.

Perhaps the most striking finding is the relative paucity of point mutations (single nucleotide variations or short insertions and deletions) driving cancer in non-coding genes and regulatory elements. While larger studies will provide greater statistical power and likely identify novel non-coding driver events (as well as additional coding drivers), aggregate analyses of the promoters of known cancer genes predict a modest yield, roughly at a 1 to 10 ratio, of driver mutations in non-coding regulatory sequences compared to

coding regions. This is in contrast to the many non-coding germline variants that affect common diseases and may be explained by smaller territory of strong functional effects in non-coding elements compared coding ones.

References

1. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Nature. 2014 Jan 23;505(7484):495-501. doi: 10.1038/nature12912. Epub 2014 Jan 5.
2. Rheinbay E, Parasuraman P, Grimsby J, Tiao G, Engreitz JM, Kim J, Lawrence MS, Taylor-Weiner A, Rodriguez-Cuevas S, Rosenberg M, Hess J, Stewart C, Maruvka YE, Stojanov P, Cortes ML, Seepo S, Cibulskis C, Tracy A, Pugh TJ, Lee J, Zheng Z, Ellisen LW, Iafrate AJ, Boehm JS, Gabriel SB, Meyerson M, Golub TR, Baselga J, Hidalgo-Miranda A, Shioda T, Bernards A, Lander ES, Getz G. Nature. 2017 Jul 6;547(7661):55-60. doi: 10.1038/nature22992. Epub 2017 Jun 28.
3. Rheinbay E, Nielsen MM, Abascal F, Tiao G, Hornshøj H, Hess JM, Pedersen RI, Feuerbach L, Sabarinathan R, Madsen HT, Kim J, Mularoni L, Shuai S, Camaioni AAL, Herrmann C, Maruvka YE, Shen C, Amin SB, Bertl J, Dhingra P, Diamanti K, Gonzalez-Perez A, Guo Q, Haradhvala NJ, Isaev K, Juul M, Komorowski K, Sushant K, Lee D, Lochovsky L, Liu EM, Pich O, Tamborero D, Umer HM, Uusküla-Reimand L, Wadelius C, Wadi L, Zhang J, Boroevich KA, Carlevaro-Fita J, Chakravarty D, Yiu CW, Chan YW, Fonseca NA, Hamilton MP, Hong C, Kahles A, Kim Y, Lehmann K-V, Andrew T, Johnson A, Kahraman A, Park K, Saksena G, Sieverling L, Sinnott-Armstrong NA, Campbell PJ, Hobolth A, Kellis M, Lawrence MS, Raphael B, Rubin MA, Sander C, Stein LD, Stuart J, Tsunoda T, Wheeler DA, Johnson R, Reimand J, Gerstein MB, Khurana E, Lopez-Bigas N, Martincorena I, Pedersen JS, Getz G, PCAWG Drivers and Functional Interpretation Group, ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Network. bioRxiv 237313. doi: 10.1101/237313
4. Campbell PJ, Getz G, Stuart JM, Korbel JO, Stein LD, ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Network. bioRxiv 162784. doi: 10.1101/162784

Gad Getz, PhD

2003	Ph.D. Physics, Weizmann Institute of Science, Rehovot, Israel
2004-2007	Postdoctoral Associate, Cancer Program, Broad Institute of Harvard and MIT
2008-2010	Head, Cancer Genome Computational Analysis
2010-present	Director, Cancer Genome Computational Analysis, Broad Institute
2013-present	Director of Bioinformatics Program, Massachusetts General Hospital Cancer Center & Department of Pathology
2013-2018	Assoc. Professor of Pathology, Mass. General Hospital and Harvard Medical School
2018-present	Professor of Pathology, Mass. General Hospital and Harvard Medical School

SYSTEMATIC IDENTIFICATION OF SPLICING ASSOCIATED VARIANTS TOWARD PRECISION MEDICINE

Yuichi Shiraishi
Center for Cancer Genomics and Advanced Therapeutics
National Cancer Center
5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(yuishira@ncc.go.jp)

Splicing defects caused by somatic variants have long been implicated in cancer development. The advances in high-throughput sequencing technology have provided the opportunity to investigate the relationship between somatic variants and splicing alterations, especially beyond the variants affecting highly conserved canonical splice sites (GT-AG dinucleotides). However, due to the complexity of splicing regulation, systematic characterization of somatic variants inducing splicing alterations has been a challenging issue.

We have developed a novel approach (SAVNet, https://github.com/friend1ws/SAVNet) based on a rigorous Bayesian theory to identify somatic variants causing splicing alterations by disrupting or creating splicing donor/acceptor motifs. Through this approach, we performed a comprehensive analysis of 8,976 primary cancer samples across 31 cancer types deposited in The Cancer Genome Atlas (TCGA), constructing a catalogue of 14,438 splicing-associated variants (SAVs) [1]. Such a large collection of SAVs enabled characterization of their positional distribution, genomic features, underlying mutational processes and list of frequently affected genes. We have generated a high-resolution profile of SAVs disrupting or creating splice sites. Notably, in addition to those disrupting GT-AG canonical splice bases, a substantial number of somatic variants affect non-canonical bases of splice sites (including previously unidentified +3 and +5 bases of donor sites, Figure 1) or newly create splice sites. Mutation signature analysis has revealed the relative contribution of each signature to SAV generation: smoking signature is more frequently associated with SAVs, whereas ultraviolet exposure and aberrant activity of the error-prone polymerase POLE have less impact on SAV generation. We showed that as many as 14.7% of samples harbored at least one SAVs in cancer-related genes (Figure 2), particularly in

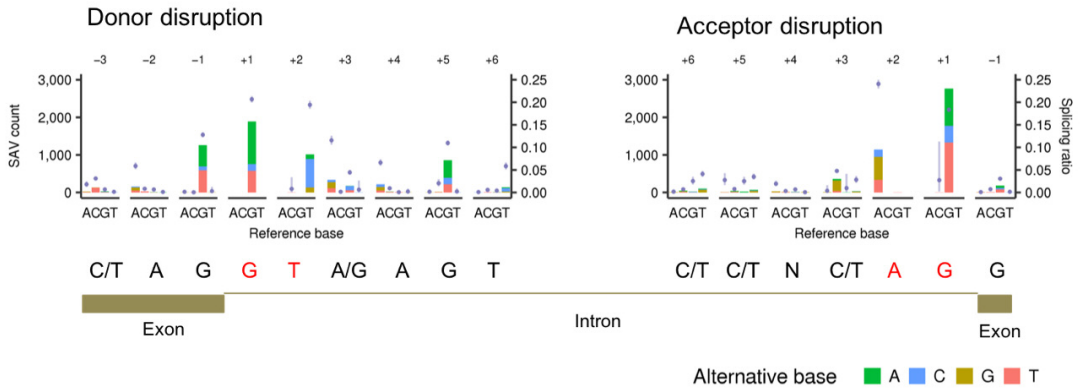


Figure 1 Base substitution patterns of SAVs at each exonic and intronic position of splice donor and acceptor sites. Extracted and modified from Shiraishi et al., Genome research, 2018.

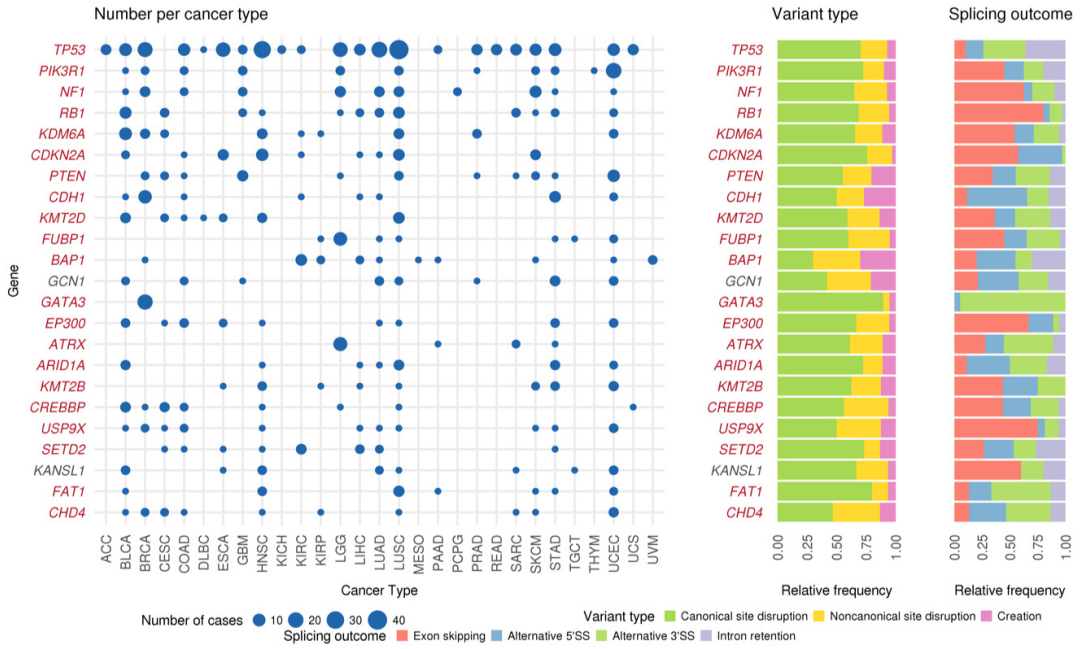


Figure 2 Landscape of SAVs in frequently altered genes across cancer types. The point size indicates the number of affected samples. Genes are sorted by the total number of SAVs in all cancer types and known cancer-related genes are shown in red. Extracted and modified from Shiraishi et al., Genome research, 2018.

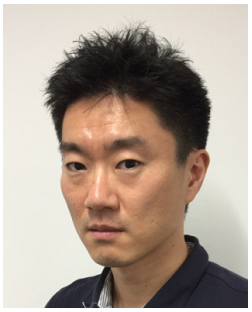
tumor suppressor genes (TSG). Importantly, among these, 6.7% of samples had SAVs disrupting non-canonical bases or creating novel splice motifs, which would not be identified by the conventional analysis focusing on those at canonical splice sites. In addition, we comprehensively described the type and position of SAVs identified in well-known TSGs, such as *TP53*, *PIK3R1*, and *CDKN2A*, which will help understanding the biological properties of these genes.

Additionally, SAVNet have been used in the PCAWG (Pan-Cancer Analysis of Whole Genomes) project, especially focusing on somatic variants in deep intronic regions, and have contributed to identification of several novel mechanisms of new exon generation by deep intronic variants [2].

Collectively, these findings obtained by systematic analysis of SAVs will give highly valuable insights into cancer genetic, biology, and medicine, especially into precision medicine.

References

1. Shiraishi Y, Kataoka K, Chiba K, Okada A, Kogure Y, Tanaka H, Ogawa S, Miyano S. A comprehensive characterization of cis-acting splicing-associated variants in human cancer. *Genome Research*, 2018 Jul 16. doi: 10.1101/gr.231951.117
2. PCAWG Transcriptome Core Group. Genomic basis for RNA alterations revealed by whole-genome analyses of 27 cancer types. *BioRxiv*, 2018.



Yuichi Shiraishi, PhD

2005-2008	Graduate Student, The Graduate University for Advanced Studies
2008-2010	Postdoctoral Researcher, RIKEN Research Center for Allergy and Immunology
2010-2018	Postdoctoral Researcher (2010-2012), Project Assistant Professor (2012-2014), Assistant Professor (2014-2018), Institute of Medical Science, The University of Tokyo
2018-present	Unit Leader, National Cancer Center

CODING AND NON-CODING CANCER MUTATIONS

Núria López-Bigas
Barcelona Biomedical Genomics Lab
Institute for Research in Biomedicine (IRB Barcelona)
Baldiri Reixac, 10, Barcelona 08012, Spain
(nuria.lopez@irbbarcelona.org)

The genome of cancer cells contains a large number of somatic mutations that are distributed unevenly along its sequence. Several studies have shown that the mutation rate (at the megabase scale) across the genome correlates with the level of chromatin compaction, DNA accessibility and replication timing. However, the genomic features that can influence the mutation rate variation at nucleotide resolution are not yet studied in detail. By analysing the whole-genome somatic mutations of different cancer types, we showed that the mutation rate vary locally around DNA protein-binding sites, such as transcription factor binding sites (TFBS) and nucleosome-covered DNA [1]. In particular, melanomas and lung tumours contain larger number of mutations at TFBS than at their flanks. We further showed that the abnormally high mutation rate at these sites is caused by a decrease of the levels of nucleotide excision repair (NER) activity (Figure 1). We have also discovered that in many tumors exons receive fewer mutations than expected. We

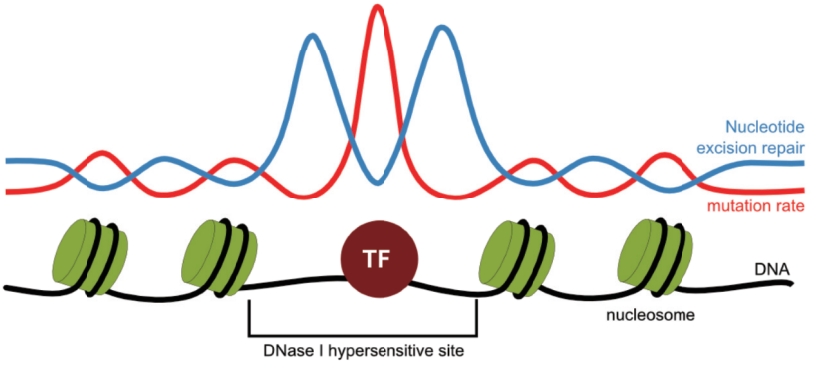


Figure 1

demonstrated that this is not due to negative selection of exonic mutations and instead it is caused by a differential mismatch repair in exons and introns [2]. This is revealed by the fact that tumors deficient of mismatch repair do not show reduced mutation burden in exons (Figure 2). We found evidences that support the notion that differences in chromatin features between exons and introns, more specifically the content of H3K36me3, play a role in this differential mismatch repair function.

We have also observed that somatic mutation rate in melanoma and other cancer types follows a periodicity of around 190 bps, which coincides with the distance between adjacent nucleosome dyads [1,3]. In particular, melanomas, esophageal and gastric carcinomas and others, show a higher rate of somatic mutations at nucleosomes than at linkers. In contrast, in lung adenocarcinomas and lung squamous cell carcinomas, the periodicity is exactly the opposite, in other words, there is a higher rate of mutations at linkers than at nucleosomes (Figure 3).

Furthermore, we found that mutations accumulate with a 10bp periodicity within nucleosome-covered DNA. The phase of the periodicity is different depending on the mutational processes contributing the majority of the mutations in each cancer type (Figure 4). For instance, the somatic mutations in melanomas follow a 10bp periodicity with higher mutation rate at places where the minor groove of the DNA faces away from histones, while somatic mutations in esophageal cancer also show a strong periodic pattern but with maxima at sites where the minor groove faces the histones. We explored whole-genome maps of different types of DNA damage and repair and discovered that the rate of accumulation of some DNA lesions (i.e. UV light damage) and of DNA repair (in particular Nucleotide Excision Repair and Base Excision Repair) are periodic, providing a mechanistic

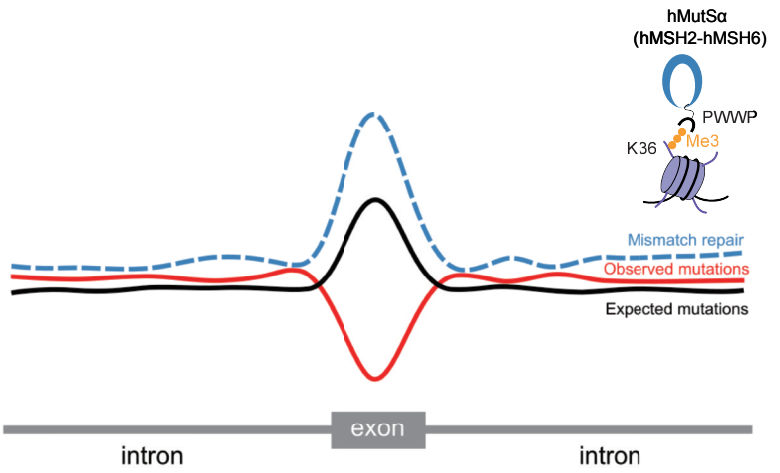


Figure 2

explanation for the periodicity observed for the mutations contributed by certain mutational processes. The periodic pattern was also observed in rare genetic variation across human and Arabidopsis populations, and in the genomic sites that have diverged between these two species and two close relatives with respect to the genome of their last common ancestor.

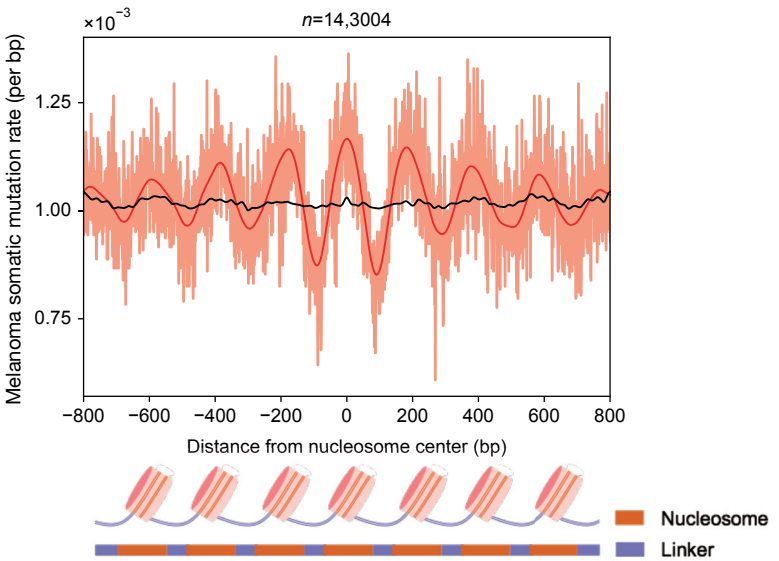


Figure 3

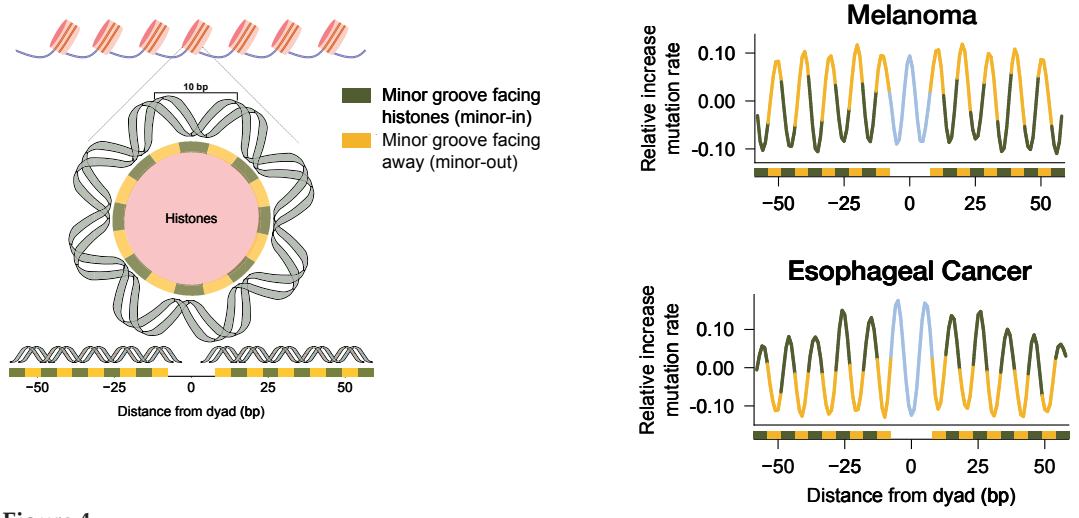


Figure 4

The genomes of eukaryotes and certain archeabacteria exhibit a detectable 10bp sequence periodicity consisting of AT, TT, TA, and AA di-nucleotides (WW periodicity). This has long been associated to the presence of nucleosomes. This WW periodicity has been speculated to have arisen through selection of mutations conducive to sequences stabilizing the DNA bending around nucleosomes. Based on our results, we propose that the interplay between damage and repair would contribute to the generation of the WW periodic pattern in the genomes of eukaryotes (Figure 5) [3].

The findings presented here have strong implications for understanding mutational and repair processes in human DNA, understanding the evolution of eukaryotic genomes, and for the identification of coding and non-coding cancer driver mutations.

Given the evolutionary principles of cancer, one effective way to identify genomic elements involved in cancer is by tracing the signals left by the positive selection of driver mutations across tumors. We analyze thousands of tumor genomes to identify driver mutations in coding and non-coding regions of the genome [4,5]. More specifically, we have analyzed the contribution of genomic alterations in the ubiquitin-mediated proteolysis system in tumorigenesis, including mutations in E3-ligases and in target degrons (Martinez et al., unpublished). The analysis of tumor cohorts provides valuable information to improve the interpretation of individual variants detected in newly sequenced tumors in clinical or research settings. We have also developed CancerGenomeInterpreter.org, a tool designed to identify driver mutations and biomarkers of drug response in individual tumors [6].

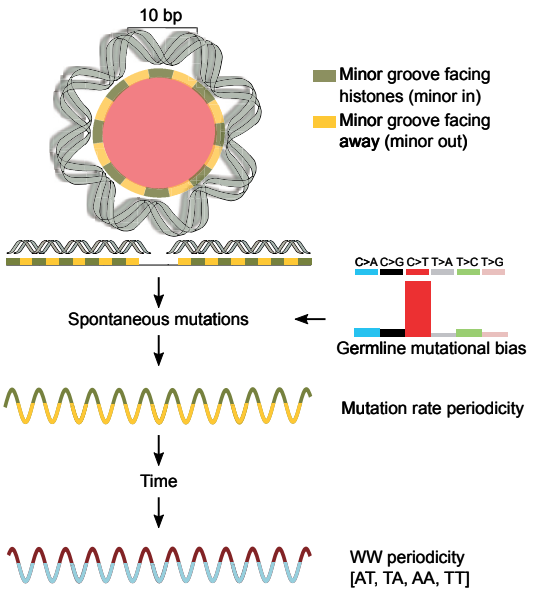


Figure 5

References

1. Sabarinathan R, Mularoni L, Deu-Pons J, Gonzalez-Perez A, Lopez-Bigas N. Nucleotide excision repair is impaired by binding of transcription factors to DNA. Nature, 2016; 532 (7598): 264-7.
2. Frigola J, Sabarinathan R, Mularoni L, Muiños F, Gonzalez-Perez A, López-Bigas N. Reduced mutation rate in exons due to differential mismatch repair. Nature Genetics, 2017 49(12), pp. 1684-1692, doi:10.1038/ng.3991
3. Pich O, Muiños F, Sabarinathan R, Reyes-Salazar I, Gonzalez-Perez A, Lopez-Bigas N, Somatic and Germline Mutation Periodicity Follow the Orientation of the DNA Minor Groove around Nucleosomes. Cell, Vol. 175, no4, 1, pp 1074-1087.e18
4. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, Santos A, Lopez-Bigas N. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods, 2013; 10 (11): 1081-1082
5. Sabarinathan R, Pich O, Martincorena I, Rubio-Perez C, Juul M, Wala J, Schumacher S, Shapira O, Sidiropoulos N, Waszak S, Tamborero D, Mularoni L, Rheinbay E, Hornshoj H, Deu-Pons J, Muinos F, Bertl J, Guo Q, Weischenfeldt J, Korbel JO, Getz G, Campbell PJ, Pedersen JS, Beroukhim R, Perez-Gonzalez A, Lopez-Bigas N, PCAWG Drivers and Functional Interpretation Group, ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Net. The whole-genome panorama of cancer drivers. 2017, Pre-print. doi: 10.1101/190330
6. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, Tusquets I, Albanell J, Rodon J, Tabernero J, de Torres C, Dienstmann R, Gonzalez-Perez A, Lopez-Bigas N. Cancer Genome Interpreter Annotates The Biological and Clinical Relevance of Tumor Alterations. 2017, Genome Medicine. doi.org/ 10.1186/ s13073-018-0531-8

Núria López-Bigas, PhD

1994-1998	B.Sc. Biology, University of Barcelona
1998-2002	PhD. Biology Oncologic Research Institute, University of Barcelona
2002-2005	Postdoctoral Researcher (Human Frontiers Science Program - HFSP) European Bioinformatics Institute (EBI)
2005-2006	Postdoctoral Researcher (Human Frontiers Science Program - HFSP) Centre for Genomic Regulation (CRG)
2006-2011	Group Leader (Ramon y Cajal Researcher) Universitat Pompeu Fabra
2011-2016	Group Leader (ICREA Research Professor), Universitat Pompeu Fabra
2016-present	Group Leader (ICREA Research Professor), Institute for Research in Biomedicine (IRB Barcelona)

MUTATIONS AND THEIR INTERACTIONS IN CANCER AND BEYOND

Ben Lehner
Systems Biology Program, Centre for Genomic Regulation (CRG)
The Barcelona Institute of Science and Technology
Institució Catalana de Recerca i Estudis Avançats (ICREA)
Dr. Aiguader 88, Barcelona 08003, Spain
(lehner.ben@gmail.com)

Our goal is to understand and be able to predict how phenotypes vary amongst individuals.

As part of our interest in stochastic processes, we have been using data from cancer genome sequencing projects to better understand the distribution and spectra of somatic mutations in human cells [1,2]. Many processes can cause the same nucleotide change in a genome, making the identification of the mechanisms causing mutations a difficult challenge. We have used rare clustered mutations as more precise fingerprints of mutagenic processes. We discovered a mutation signature matching the spectrum of translesion DNA polymerase eta (POLH) that is associated with UV-exposure and alcohol consumption and targets the H3K36me3 chromatin of active genes in a mismatch repair (MMR)-dependent manner. These regions normally have low mutation rates because error-free MMR also targets H3K36me3 chromatin. Carcinogens and error-prone repair therefore redistribute mutations to the more important regions of the genome, contributing a substantial mutation load in many tumors, including driver mutations [1].

We are also extremely interested in understanding interactions between mutations. The genetic causes of cancer include both somatic mutations and inherited germline variants. Large-scale tumour sequencing has revolutionized the identification of somatic driver alterations but has had limited impact on the identification of cancer predisposition genes (CPGs). We developed a statistical method, ALFRED, that tests Knudson's two-hit hypothesis to systematically identify CPGs from cancer genome data [3]. Applied to ~10,000 tumour exomes the approach identifies known and putative CPGs - including the chromatin modifier NSD1 - that contribute to cancer through a combination of rare

germline variants and somatic loss-of-heterozygosity (LOH). We estimate that rare germline variants in these genes contribute to a median of 2% of tumours across 17 cancer types.

In parallel, we have been using deep mutagenesis to better understand how mutations interact within [4] and between [5] genes to affect a range of molecular functions and phenotypes. These include the activity of individual proteins and RNAs, protein interactions, splicing, and gene circuits. We find that both pairwise and higher-order interactions between mutations are common, with mutations quite often switching from positive to negative effects, even in closely related genotypes, as illustrated for a yeast tRNA (Figure 1 and reference [4]). These genetic (epistatic) interactions are of two classes: non-specific interactions resulting from non-linearities in genotype-phenotype maps and specific interactions related to 3D structure [5]. Indeed, we have found that quantifying how mutations interact within and between proteins can provide sufficient information to determine their high-resolution structures (Figure 2 and reference [6]). Deep mutagenesis combined with selection and sequencing can therefore be used to solve the structures of macromolecules [6].

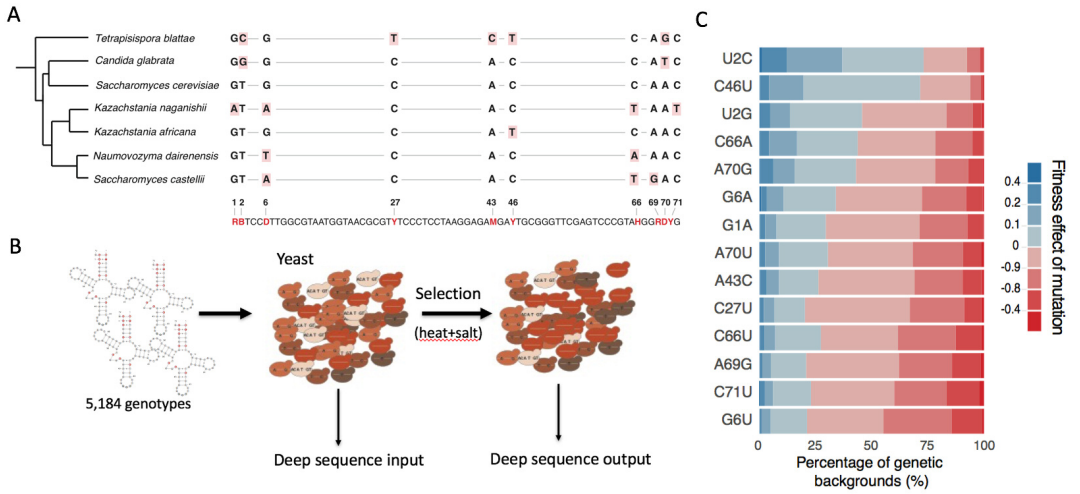


Figure 1 Combinatorial mutagenesis of a tRNA
(A) A library consisting of all combinations of 14 substitutions observed across yeast species in 10 positions consists of >5,000 unique genotypes of a tRNA. (B) The fitness of these genotypes in a restrictive condition was measured using selection and deep sequencing. (C) All 14 mutations have beneficial, detrimental and neutral effects in different genetic backgrounds of the tRNA. The effect of each mutation was quantified in ~1,500 different genotypes.

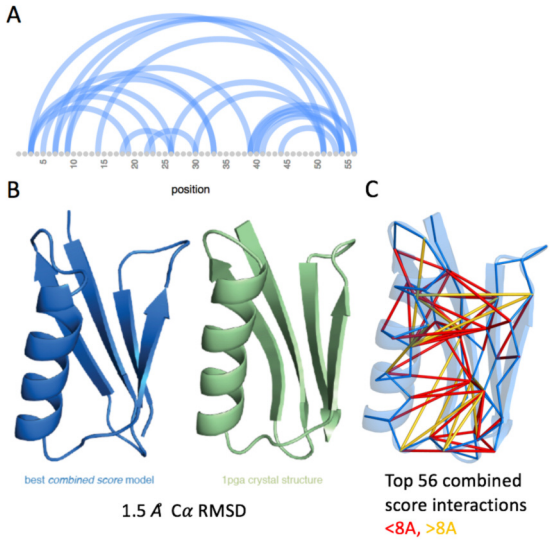


Figure 2 Solving protein structures using deep mutagenesis
(A) Quantifying the genetic interactions between mutations in different positions in a protein domain (protein G B1 domain) by combining two different metrics predicts structural contacts. (B) These predicted contacts are sufficient restraints to determine the backbone 3D structure of the domain. The predicted structure is compared to a reference crystal structure. RMSD – root-mean-square deviation. (C) Top predicted combined score contacts shown on the structure.

Finally, we are using *C. elegans* as a model system to understand how epigenetic inheritance can impact physiology and mutation outcomes in isogenic individuals [7-9]. We discovered an example of transgenerational epigenetic memory of an environmental perturbation and dissected the underlying molecular mechanism [8]. We have also found that impaired DNA replication during the fast cell divisions in an early embryo can drive directional and long lasting changes in chromatin [9]. We speculate that similar effects may occur during tumour development.

References

1. Supek F, Lehner B. Clustered mutation signatures reveal that error-prone DNA repair targets mutations to active genes. *Cell* 170(3):534-547. 2017.
2. Supek F, Lehner B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* 521:81-4. 2015.
3. Park S, Supek F, Lehner B. Systematic discovery of germline cancer predisposition genes through the identification of somatic second hits. *Nature Communications* 4;9(1):2601. 2018.

4. Domingo J, Diss G, Lehner B. Pairwise and higher-order genetic interactions during the evolution of a tRNA. *Nature* 558(7708):117-121. 2018.
5. Diss G, Lehner B. The genetic landscape of a physical interaction. *Elife* 11;7. pii: e32472. 2018.
6. Schmiedel J, Lehner B. Determining protein structures using genetics. bioRxiv. https://www.biorxiv.org/content/early/2018/04/20/303875.
7. Perez MF, Francesconi M, Hidalgo-Carcedo C, Lehner B. Maternal age generates phenotypic variation in *Caenorhabditis elegans*. *Nature* 7;552(7683):106-109. 2017.
8. Klosin A, Casas E, Hidalgo-Carcedo C, Vavouri T, Lehner B. Transgenerational transmission of environmental information in *C. elegans*. *Science* 21;356(6335):320-323. 2017.
9. Klosin A, Reis K, Hidalgo-Carcedo C, Casas E, Vavouri T, Lehner B. Impaired DNA replication derepresses chromatin and generates a transgenerationally inherited epigenetic memory. *Science Advances* 16;3(8):e1701143. 2017.

Ben Lehner, PhD

1997-2000	BA, Natural Sciences, University of Cambridge, UK
2000-2004	PhD, University of Cambridge, UK
2004-2006	Postdoctoral Fellow, Wellcome Trust Sanger Institute, UK
2006-present	Group Leader, CRG, Barcelona
2017-present	Coordinator of Systems Biology Program, CRG

LONG NONCODING RNA AND WNT SIGNALING IN CANCER

Hiroyuki Aburatani
Genome Science Laboratory
Research Center for Advanced Science and Technology
The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
(haburata-tky@umin.ac.jp)

Long noncoding RNAs (lncRNAs) have been found to be involved in cell growth and apoptosis, in part through epigenetic regulation. Wnt signal, which is frequently activated in various cancers, induces the expression of genes that regulate cell cycle and proliferation. While many genes have been identified as the targets of β-catenin, it is not fully understood how activated β-catenin regulates its downstream targets. We identified a β-catenin-target lncRNA, lnc12R, by combination of RNA-seq and ChIP-seq for β-catenin and histone modification. Lnc12R is overexpressed in colorectal tumors with APC mutation in In TCGA dataset. Knockout of lnc12R using CRISPR/Cas9 inhibits tumorigenesis in vivo. Repression of lnc12R resulted in H3K27 deacetylation and repressed β-catenin target genes, such as LGR5, without affecting the binding of β-catenin on their promoter.

Hiroyuki Aburatani, MD, PhD

1980-1981	Internship in Internal Medicine, The University of Tokyo
1983-1988	Clinical Research Fellow, Third Department of Internal Medicine, The University of Tokyo
1988-1991	Assistant Professor, Third Department of Internal Medicine, The University of Tokyo
1988-1994	Visiting Scientist, Center for Cancer Research, MIT, Cambridge, MA, USA
1995-1999	Assistant Professor, Third Department of Internal Medicine, The University of Tokyo
1999-2001	Associate Professor, Genome Science Division, Research Center for Advanced Science and Technology, The University of Tokyo
2001-present	Professor, Genome Science Division, Research Center for Advanced Science and Technology, The University of Tokyo

ANALYSIS OF NON-CODING GENOME ALTERATIONS IN HUMAN CANCER

Matthew Meyerson
Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA
Broad Institute, Cambridge, MA, USA
(Matthew_Meyerson@dfci.harvard.edu)

The advent of high-throughput genome sequencing has enabled systematic discovery of somatic genome alterations through genome, exome, and transcriptome sequencing of paired human cancer and germline nucleic acids. These discoveries have led to the identification of somatic genome alterations in new pathways including epigenetic regulators, RNA splicing factors, and immune response regulators, in addition to previously known signal transduction, cell cycle, and transcriptional pathways. Here, I describe alterations in these pathways by copy number disruption, mutation, and re-arrangement, especially as seen through the lens of recent and ongoing large-scale studies of lung adenocarcinoma and squamous cell lung carcinoma genomes. I connect these genomic analyses with functional studies, and discuss the implications for the understanding of cancer pathogenesis and for the development of targeted and immunomodulatory therapies.

Introduction

This report focuses particularly on the analysis of non-coding alterations in cancer genomes. After a brief introduction, I first discuss copy number alterations in human cancer, with a particular focus on the analysis of aneuploidy [1]. Second, I describe the phenomenon of super-enhancer duplications in the vicinity of oncogenes such as *MYC* [2], *KLF5* [3], and the androgen receptor gene [4]. Third, I discuss the findings of gene body insertions and deletions in lineage-expressed genes such as surfactant protein genes in lung cancer [5]. Then I consider the future of cancer genome sequencing with a concentration on the analysis of linked reads and long reads for complete characterization of genome

structure. This summary does not include a comprehensive description of the scientific field, but is rather limited to the work of my own laboratory group, presented at the 49th International Symposium of the Princess Takamatsu Cancer Research Fund that took place in Tokyo, Japan in November, 2019.

Cancer is a disease of the genome. Cancer is caused by both germline, or inherited, and somatic, or acquired, alterations in the genome. These alterations include single nucleotide substitutions, small insertions, deletions or duplications, chromosomal rearrangements that may occur either within or between chromosomes, and copy number alterations, including focal gains or losses of DNA that may be restricted to particular regions of a chromosome, or losses or gains of entire chromosomes or chromosome arms. Cancer may also be caused by viral or bacterial infection, which could also considered to be the acquisition of novel genomes.

During the time from 2000 to 2015, following the initial sequencing of the human genome and during a time of rapid improvement in genome sequencing technology, most population-level cancer genome studies, such as The Cancer Genome Atlas project, TCGA, in the United States of America, have focused on the sequence of the coding regions of the genome. The sequence of all exons of the coding genome, or the exome, in over 10,000 pairs of human cancer-derived DNA samples, in comparison to the corresponding sequences of their matched normal DNA samples, has permitted the delineation of the most common, recurrent coding mutations across human cancer.

However, the coding genome accounts for only 1% of all of the genome, and the vast majority of somatic genome alterations, as well as germline genome alterations, occur in the 99% of the genome that does not code for proteins. In the last several years, the sequencing of whole genomes has begun to elucidate the alterations in non-coding regions of cancer genomes, such as promoter mutations, and the enhancer duplications described below. In addition, the continued improvement in genome sequencing technology, and the development of new technologies for long reads and linked reads, now enables a more complete description of somatic alterations in the cancer genome, as described below.

Aneuploidy is a universal feature of cancer genomes

Aneuploidy refers to the loss or gain of entire chromosomes or chromosome arms, and is one of the most common features of cancer genomes, being almost universal in carcinomas, or cancers of the epithelium, such as lung, breast, colon, pancreatic, ovarian, and stomach cancers. The recognition of aneuploidy is not new—indeed, using early chromosome staining techniques, Theodor Boveri described aneuploidy as a feature of cancer more than 100 years ago, in 1914. Our recent studies have shown that whole chromosome or chromosome arms or losses are the most common of the currently

detectable features of cancer genomes, more common than *TP53* mutation or focal alterations in copy number [1, 6, 7].

However, even after 100 years of study, we do not yet know the role of aneuploidy in cancer pathogenesis. Does aneuploidy cause human cancer, or is it merely a bystander, occurring because it is tolerated because of the other genome alterations in human cancers?

After completing a survey of the role of aneuploidy in human cancers through The Cancer Genome Atlas, and confirming that chromosome level and chromosome arm level genome alterations are almost universal in epithelial cancers and exhibit recurrent cancer-specific patterns, Alison Taylor in my laboratory developed a method to generate aneuploidy experimental in lung epithelial cells. She used CRISPR methods to cut DNA on chromosome arm 3p near the centromere together with introduction of a plasmid that has homology to the cut site and contains an artificial telomere (a gift from Professor Mitsuo Oshimura at Tottori Universty), to mimic the 3p loss that is so common in squamous cell lung cancers. In so doing, Dr. Taylor generated cells that had lost chromosome arm 3p. Interestingly, these cells grew more slowly than parental cells, and were eventually over-grown by cells that gained an extra copy of chromosome arm 3p, becoming functionally cells with a gain of chromosome arm 3q [1]. This work provides an experimental method by which to study aneuploidy but it still leaves unresolved the key question: does aneuploidy promote cancer? Further studies will be required to answer this question.

Duplication of super-enhancer elements near oncogenes

Until recently, we believed that focal amplifications of oncogenes were always targeting the coding gene body itself together with the promoter elements of the gene, causing gene over-expression through increased number of coding gene copies. One of the biggest surprises in whole genome sequencing of cancer is that this is not always the case.

Instead, the amplification of non-coding elements near oncogenes has now been found to be a common type of activation of oncogene expression. Dr. Xiaoyang Zhang in my laboratory first discovered such alterations near the *MYC* gene. Different enhancer regions, both showing lineage specific histone H3 lysine acetylation which is a marker of enhancers, were found to be amplified in endometrial cancer or in lung adenocarcinoma. In both cases, the amplifications were only in the enhancer regions, a few hundred kilobases from the *MYC* gene body, and not in the gene itself. Working with Drs. Peter Choi and Joshua Francis in my laboratory, Dr. Zhang went on to study the impact of enhancer deletion or repression in a lung cancer cell line bearing *MYC* enhancer duplication. Using either CRISPR-mediated deletion of an enhancer sub-region, or KRAB-mediated repression of this region, Dr. Zhang and colleagues showed that the enhancer activity promotes *MYC* expression and cancer cell growth [2].

Analysis of enhancer duplication also helped enable us to find another oncogene, the *KLF5* gene. We discovered *KLF5* enhancer duplication in a variety of squamous cell cancer types including lung, esophageal and cervical squamous carcinomas. *KLF5* is also activated in other specimens of these same cancer types, by point mutations either in a DNA-binding domain or in a degradation domain. Here, Dr. Zhang again went on to show an oncogenic phenotype for enhancer duplication that was reversed by CRISP-mediated detion [3].

Finally, in our most recent discovery on this topic, Drs. Srinivas Viswanathan, Gavin Ha and Andreas Hoff performed linked-read sequencing of metastatic castrate-resistant prostate cancers and searched for novel regions of copy number alteration. Here, they found duplication of a region several hundred kilobases from the gene body of the androgen receptor gene, *AR*. *AR* enhancer duplications occurred especially in the population of prostate cancers resistant to nuclear hormone receptor inhibitors, and were associated with increased levels of *AR* expression [4].

One possible gene for discovering additional such super-enhancer duplications is the *SOX2* gene subjected to focal amplification in squamous carcinomas of the lung, esophagus and cervix. As shown in Figure 1, copy number analysis of data from The Cancer Genome Atlas reveals multiple regions of amplification near the *SOX2* gene body, suggestive of potential enhancer duplication. The *NKX2-1* gene that is amplified in lung adenocarcinoma also shows evidence of non-coding duplications (data not shown).

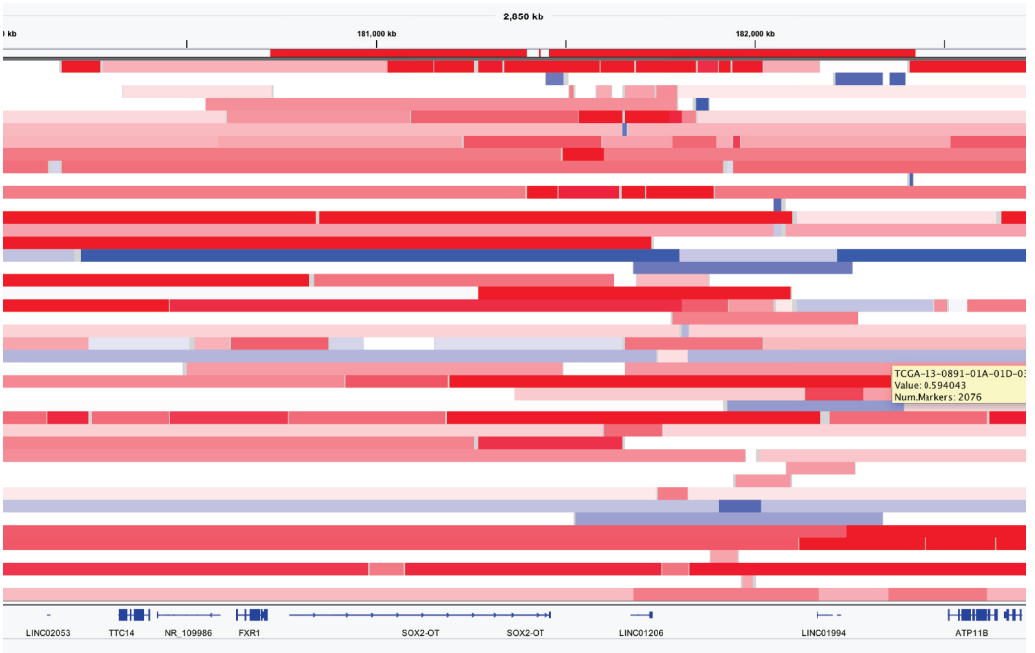


Figure 1 Copy number analysis of the region near the *SOX2* gene across human cancers shows multiple non-coding deletions in the region (red and pink bars)

Together, these findings suggest that enhancer duplication is a common mechanism of oncogene activation and could often be a mechanism by which alterations in the cancer genome induce expression of these oncogenes. This implies that whole genome sequencing, which can reveal such duplications, may be key for genome-based cancer diagnosis.

Deletions within or near genes expressed at high levels in a lineage-unique manner

Whole genome sequencing of cancer can reveal surprises. One such surprise was our recent discovery of deletions within the 3′ untranslated regions, or sometimes the coding gene bodies, of highly expressed genes within specific cancer lineages. The genes subjected to lineage specific deletion mutations include the surfactant protein genes *SFTPA, SFTPB,* and *SFTPC* in lung adenocarcinoma, the albumin gene *ALB* in hepatocellular carcinoma, the lipase gene *LIPF* in gastric carcinoma, and the thyroglobulin gene *TG* in thyroid carcinoma [5]. These deletions do not appear to affect gene expression. One possibility is that these alterations are actually passenger mutations that are a consequence of the extremely high transcriptional level in these tissues, but this question remains to be resolved.

Towards the future—complete cancer genome sequencing with long reads or linked reads

As described above, we are beginning to learn that cancer genomes harbor critical alterations beyond the coding genome. Even though our current technology represents a revolutionary leap from previous generations of genome sequencing technology, we do not yet have the capability to find all cancer genome alterations at once. The advent of new technologies such as linked-read sequencing methods and long read sequencing technologies such as nanopore sequencing, now gives us the opportunity to describe the absolutely complete sequence and structure of cancer genomes. Once we make this complete description, we will find far more causative alterations in cancer genomes, which in turn will help us to develop new cancer therapies.

References

1. Taylor AM, Shih J, Ha G, Gao GF, Zhang X, Berger AC, Schumacher SE, Wang C, Hu H, Liu J, Lazar AJ; Cancer Genome Atlas Research Network, Cherniack AD, Beroukhim R, Meyerson M. Genomic and functional approaches to understanding cancer aneuploidy. Cancer Cell. 2018 Apr 9;33(4):676-89.

2. Zhang X, Choi PS, Francis JM, Imielinski M, Watanabe H, Cherniack AD, Meyerson M. Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. Nat Genet. 2016 Feb;48(2):176-82.
3. Zhang X, Choi PS, Francis JM, Gao GF, Campbell JD, Ramachandran A, Mitsuishi Y, Ha G, Shih J, Vazquez F, Tsherniak A, Taylor AM, Zhou J, Wu Z, Berger AC, Giannakis M, Hahn WC, Cherniack AD, Meyerson M. Somatic superenhancer duplications and hotspot mutations lead to oncogenic activation of the KLF5 transcription factor. Cancer Discov. 2018 Jan 8(1):108-25.
4. Viswanathan SR, Ha G, Hoff AM, Meyerson M. Structural alterations driving castration-resistant prostate cancer revealed by linked-read genome sequencing. Cell 2018 Jul 12;174(2):433-447.
5. Imielinski M, Guo G, Meyerson M. Insertions and deletions target lineage-defining genes in human cancers. Cell. 2017 Jan 26;168(3):460-472.
6. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, Mc Henry KT, Pinchback RM, Ligon AH, Cho YJ, Haery L, Greulich H, Reich M, Winckler W, Lawrence MS, Weir BA, Tanaka KE, Chiang DY, Bass AJ, Loo A, Hoffman C, Prensner J, Liefeld T, Gao Q, Yecies D, Signoretti S, Maher E, Kaye FJ, Sasaki H, Tepper JE, Fletcher JA, Tabernero J, Baselga J, Tsao MS, Demichelis F, Rubin MA, Janne PA, Daly MJ, Nucera C, Levine RL, Ebert BL, Gabriel S, Rustgi AK, Antonescu CR, Ladanyi M, Letai A, Garraway LA, Loda M, Beer DG, True LD, Okamoto A, Pomeroy SL, Singer S, Golub TR, Lander ES, Getz G, Sellers WR, Meyerson M. The landscape of somatic copy-number alteration across human cancers. Nature. 2010 Feb 18;463(7283):899-905.
7. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, Lawrence MS, Zhang CZ, Wala J, Mermel CH, Sougnez C, Gabriel SB, Hernandez B, Shen H, Laird PW, Getz G, Meyerson M, Beroukhim R. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013 Sep 26;45(10):1134-40.

Matthew Meyerson, MD, PhD

1993-1996	Resident, Massachusetts General Hospital
1995-1998	Research Fellow, Whitehead Institute
1998-2005	Asst. Professor, Harvard Medical School
2005-2009	Assoc. Professor, Harvard Medical School
2009-present	Professor, Harvard Medical School
2017-present	Director of Cancer Genomics, Dana-Farber Cancer Institute and Broad Institute

A UNIQUE MECHANISM OF CANCER IMMUNE EVASION VIA THE DISRUPTION OF PD-L1 3′-UTR

Seishi Ogawa
Department of Pathology and Tumor Biology, Graduate School of Medicine
Kyoto University, Japan
Department of Molecular Hematology, Karolinska Institute, Sweden
Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
(sogawa-tky@umin.ac.jp)

Successful treatment of many advanced cancer patients using antibodies against programmed cell death 1 (PD-1) and its ligand (PD-L1) has highlighted the critical importance of PD-1/PD-L1-mediated immune escape in cancer development. However, the genetic basis for the immune escape has not been fully elucidated, with the exception of elevated *PD-L1* expression by gene amplification and utilization of an ectopic promoter by translocation, as reported in Hodgkin and other B-cell lymphomas, as well as stomach adenocarcinoma. Here we show a unique genetic mechanism of immune escape caused by structural variations (SVs) commonly disrupting the 3′ region of the PD-L1 gene. Widely affecting multiple common human cancer types, including adult T-cell leukemia/lymphoma (27%), diffuse large B-cell lymphoma (8%), and stomach adenocarcinoma (2%), and other solid cancers (Figure 1), these SVs invariably lead to a

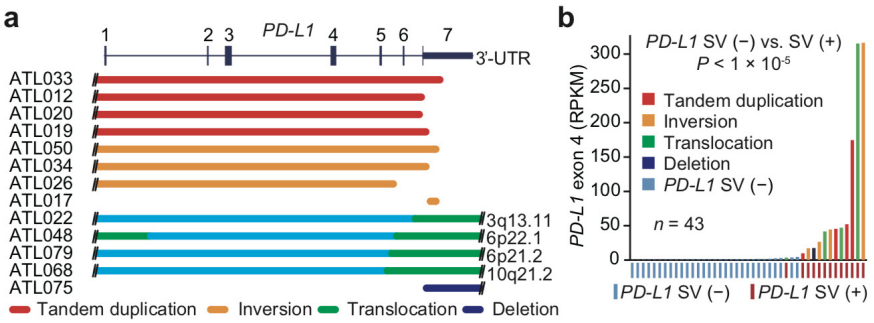


Figure 1 Aberrant expression of PD-L1 in ATL samples with SVs involving *PD-L1* 3′-UTR
a, Different types of SVs commonly affecting 3′ part of *PD-L1* are shown by indicated colors. b, *PD-L1* exon 4 expression (reads per kb of exon per million mapped reads, RPKM) in 43 ATL samples, colored by *PD-L1* SV status. Welch's t-test.

marked elevation of aberrant *PD-L1* transcripts that are stabilized by truncation of the 3′-untranslated region (UTR) (Figure 2). We further interrogated PD-L1/PD-L2-involving somatic aberrations in 387 samples from various lymphoma subtypes using high-throughput sequencing, particularly focusing on virus-associated lymphomas. A high frequency of *PD-L1*-involving genetic aberrations was confirmed in this large set of lymphoma samples (Figure 3). In addition, we found recurrent SVs affecting 3′-UTR of PD-L2. PD-L1/PD-L2-involving abnormalities, including SVs and focal amplification, was especially common in EBV-positive lymphomas [36 (24%) of 151 cases], including EBV-positive diffuse large B-cell lymphoma (DLBCL, 19%) and peripheral T-cell lymphoma-not otherwise specified (15%), and mature NK/T-cell neoplasms (17-57%) (Figure 4). In particular, PD-L1-involving alterations represented the most prevalent somatic lesions in extranodal NK/T-cell lymphoma. By contrast, the frequency was much lower in EBV-negative lymphomas regardless of histology type [12 (5%) of 236 cases]. The majority of these abnormalities were explained by truncations of their 3′-untranslated regions, which are thought to result in overexpression of the involved gene, contributing to tumor immune evasion. Disruption of Pd-l1 3′-UTR in mice enables immune evasion of EG7-OVA tumor cells with elevated Pd-l1 expression in vivo, which is effectively inhibited by Pd-1/Pd-l1 blockade, supporting the role of relevant SVs in clonal selection through immune evasion. Our findings not only unmask a novel regulatory mechanism of expression of these PD-1 ligands, but also suggest that PD-L1/PD-L2 3′-UTR disruption could serve as a genetic marker to identify cancers that actively evade anti-tumor immunity through PD-L1 overexpression.

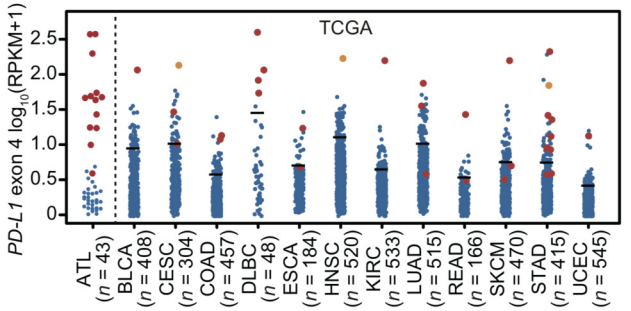


Figure 2 *PD-L1* SVs associated with overexpression of aberrant *PD-L1* transcripts in multiple cancers *PD-L1* expression in each TCGA cancer type containing *PD-L1* SV cases. Each bar represents the 10th percentile. BLCA, bladder urothelial carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal squamous cell carcinoma; KIRC, kidney renal clear cell carcinoma; LUAD, lung adenocarcinoma; READ, rectal adenocarcinoma; SKCM, skin cutaneous melanoma; UCEC, uterine corpus endometrioid carcinoma.

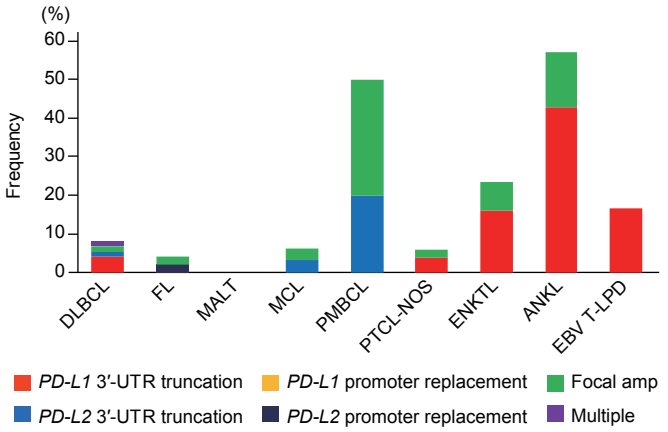


Figure 3 Genetic alterations involving PD-1 ligands in various subtypes of lymphomas
Frequency of genetic alterations involving *PD-L1* and/or *PD-L2* in each lymphoma subtype. Type of alterations is indicated by color. Cases harboring both SV and focal CNA affecting *PD-L1* and/or *PD-L2* are combined into the corresponding SV group. Multiple represents cases harboring both *PD-L1* and *PD-L2* SVs. DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; MALT, mucosa-associated lymphoid tissue lymphoma; MCL, mantle cell lymphoma; PMBCL, primary mediastinal B-cell lymphoma; PTCL-NOS, peripheral T-cell lymphoma-not otherwise specified; ENKTL, extranodal NK/T-cell lymphoma; ANKL, aggressive NK-cell leukemia; EBV T-LPD, Systemic EBV-positive T-cell lymphoproliferative disorder; Amp, amplification.

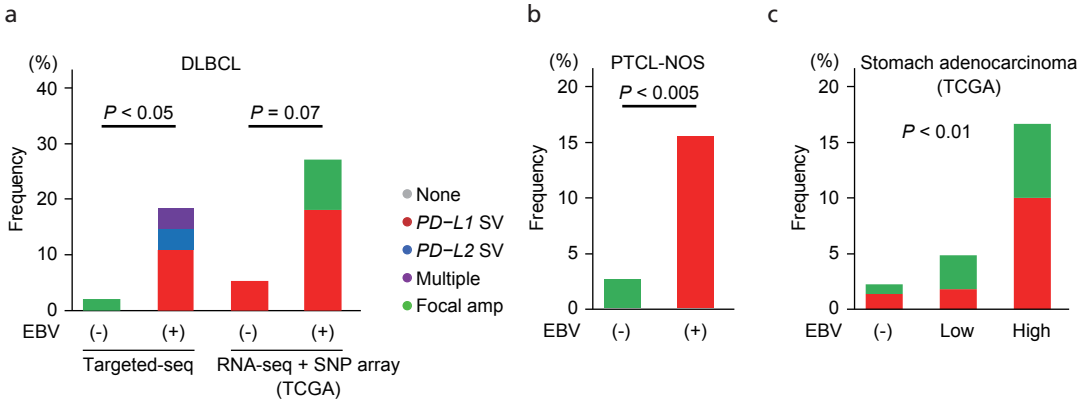


Figure 4 *PD-L1*/*PD-L2* genetic alterations associated with EBV infection
Frequency of genetic alterations involving *PD-L1*/*PD-L2* according to EBV status in DLBCL (a), PTCL-NOS (b), and stomach adenocarcinoma (c). Fisher's exact test is used for DLBCL and PTCL-NOS, and Cochran-Armitage trend test is applied to stomach adenocarcinoma. EBV high group corresponds to the EBV-positive group in the TCGA classification of stomach adenocarcinoma

References

1. Kataoka K, Shiraishi Y, Takeda Y, Sakata S, Matsumoto M, Nagano S, Maeda T, Nagata Y, Kitanaka A, Mizuno S, Tanaka H, Chiba K, Ito S, Watatani Y, Kakiuchi N, Suzuki H, Yoshizato T, Yoshida K, Sanada M, Itonaga H, Imaizumi Y, Totoki Y, Munakata W, Nakamura H, Hama N, Shide K, Kubuki Y, Hidaka T, Kameda T, Masuda K, Minato N, Kashiwase K, Izutsu K, Takaori-Kondo A, Miyazaki Y, Takahashi S, Shibata T, Kawamoto H, Akatsuka Y, Shimoda K, Takeuchi K, Seya T, Miyano S, Ogawa S. Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. Nature. 2016;534(7607):402-406.
2. Kataoka K, Miyoshi H, Sakata S, Dobashi A, Couronne L, Kogure Y, Sato Y, Nishida K, Gion Y, Shiraishi Y, Tanaka H, Chiba K, Watatani Y, Kakiuchi N, Shiozawa Y, Yoshizato T, Yoshida K, Makishima H, Sanada M, Onozawa M, Teshima T, Yoshiki Y, Ishida T, Suzuki K, Shimada K, Tomita A, Kato M, Ota Y, Izutsu K, Demachi-Okamura A, Akatsuka Y, Miyano S, Yoshino T, Gaulard P, Hermine O, Takeuchi K, Ohshima K, Ogawa S. Frequent structural variations involving programmed death ligands in Epstein-Barr virus-associated lymphomas. Leukemia. 2019.

Seishi Ogawa, MD, PhD

1989-1993	MD, PhD Program, The University of Tokyo
1983-1988	Undergraduate/MD Program, School of Medicine, The University of Tokyo
1988-1989	Postgraduate Clinical Training in Internal Medicine
1994-1995	Clinical Associate, The University of Tokyo
1995-1996	Research Fellow
1996-2002	Assistant Professor of Medicine, The University of Tokyo
2002-2013	Associate Professor, The University of Tokyo
2013-present	Professor, Department of Pathology and Tumor Biology, Kyoto University
2017-present	Guest Professor, Department of Molecular Hematology, Karolinska Institute

GENOMIC PATHOLOGY OF DIFFUSE-TYPE GASTRIC CARCINOMA

Shumpei Ishikawa
Department of Molecular Preventive Medicine
Graduate School of Medicine, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
(isikawas-tky@umin.org)

Diffuse-type gastric carcinoma (DGC) is one of the major subtypes of stomach cancers and has unique biological & clinical features such as single cell infiltration, strong fibrotic stroma, and massive metastasis/dissemination potential. Except for CDH1 (E-cadherin), DGC-specific drivers that explain these unique features and thus, could be molecular targets, had not been discovered so far. Recent comprehensive cancer genomics has revealed its genomic stability, lack of popular cancer drivers, and unique RHOA gene driver mutations [1,2,3]. DGC has low mutation burdens, and unlike other chromosomal instable-type gastric cancers, harbors low frequencies of receptor tyrosine kinase amplifications including HER2 gene aberrations. Therefore, DGC is not a good target of currently available molecular drugs like tyrosine kinase inhibitors. The newly discovered RHOA mutations did not show a random distribution but exhibited several hot spots, suggesting that it is a cancer driver. The driver nature of mutant RHOA was supported by several experiments, where the survivals of RHOA-mutant cell lines were strongly dependent on their mutant RHOAs. Although the exact biochemistry of mutant RHOAs would be complex, one of its mechanisms is suggested to be their dominant negative nature of mutant proteins, which inactivate ROCK pathway to promote cell survival [4]. Histologically, DGCs with mutant RHOA are composed of the major component of poorly cohesive carcinoma with limited tubular differentiation, and show penetrative growth patterns at the edges of the mucosal area [5]. Likewise, mutant RHOA is macroscopically linked with Borrman type3 appearance and is not frequent among linitis-plastica-like cancers showing wide spread submucosal infiltration.

While global genomic analysis so far has revealed heterogeneity and subgroups of gastric cancer, tumor immune microenvironment is also very heterogeneous among gastric cancers. Applying recently developed deconvolution program for large scale transcriptome data showed a characteristic immune microenvironment of DGCs, where T-cell frequently tended to be inactive and B-cell infiltration was relatively increased. DGC was suggested to be categorized into "immunologically cold tumors", which is compatible with its low tumor mutation burden and a recent report showing its low responsiveness to immune checkpoint inhibitors in a clinical trial. Based on these observation, comprehensive profiling of immune repertoires of T-cell and B-cell in clinical gastric cancer tissues was performed [7]. It revealed different patterns of responses of T-cell and B-cell repertoires in the cancer environment and clonal expansion of tumor-specific B-cells were frequently observed. The CDR (Complementarity Determining Region) of immunoglobulins were substantially variable and it was difficult to find common features of cancer-specific immunoglobulin sequences. In order to investigate what kinds of antigens were recognized by such tumor-specific dominant clones we discoverd, human IgG antibodies were reconstructed based on the sequences of tumor-specific immunoglobulin clones. While many of the tested clones were shown to recognise autoantigens which were also reported in autoimmune diseases, several clones commonly recognized a unique subgroup of carbohydrates, sulfated glycosaminoglycan (s-GAG). By reconstituting more numbers of immunoglobulin clone, it was revealed that around 30 to 40% of the most dominantly infiltrated B-cell clones in individual cases recognized s-GAG; therefore, s-GAG was suggested to be the major humoral cancer antigens among DGC. The anti-s-GAG antibodies synthesized based on the sequences of dominant clones exhibited tumor cell binding properties, and moreover, some of which also showed growth suppressive functions. These clones were well incorporated into cancer cells and induced cell death by using antibody-drug conjugate technologies.

These results have unraveled the unique genomic and immunogenomic features of DGC, which have also provided us with new concepts for designing molecular targeted drugs as well as the possibility of another-type of cancer immunotherapy by modulating the humoral cancer immune system for this refractory cancer.

References

1. Kakiuchi M, Nishizawa T, Ueda H, Gotoh K, Tanaka A, Hayashi A, Yamamoto S, Tatsuno K, Katoh H, Watanabe Y, Ichimura T, Ushiku T, Funahashi S, Tateishi K, Wada I, Shimizu N, Nomura S, Koike K, Seto Y, Fukayama M, Aburatani H, Ishikawa S*: Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat. Genet.* 46(6):583-7,2014 (* Correspondence)

2. Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S, Chan KH, Chan AS, Tsui WY, Ho SL, Chan AK, Man JL, Foglizzo V, Ng MK, Chan AS, Ching YP, Cheng GH, Xie T, Fernandez J, Li VS, Clevers H, Rejto PA, Mao M, Leung SY. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. Nat Genet. 46(6):573-82, 2014
3. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 513(7517):202-9, 2014
4. Nishizawa T, Nakano K, Harada A, Kakiuchi M, Funahashi SI, Suzuki M, Ishikawa S, Aburatani H. DGC-specific RHOA mutations maintained cancer cell survival and promoted cell migration via ROCK inactivation. Oncotarget. 9(33):23198-23207,2018
5. Ushiku T, Ishikawa S*, Kakiuchi M, Tanaka A, Katoh H, Aburatani H, Lauwers GY, Fukayama M: RHOA mutation in diffuse-type gastric cancer: A comparative clinicopathological analysis of 87 cases. *Gastric Cancer* 19(2):403-11,2016 (* Correspondence)
6. Katoh H, Komura D, Konishi H, Suzuki R, Yamamoto A, Kakiuchi M, Sato R, Ushiku T, Yamamoto S, Tatsuno K, Oshima T, Nomura S, Seto Y, Fukayama M, Aburatani H, Ishikawa S*: Immunogenetic Profiling for Gastric Cancers Identifies Sulfated Glycosaminoglycans as Major and Functional B Cell Antigens in Human Malignancies. *Cell Rep.* 1;20(5):1073-1087,2017 (* Correspondence)

Shumpei Ishikawa, MD, PhD

2000	Graduated from Faculty of Medicine, The University of Tokyo
2004	Graduated from Graduate School of Medicine, The University of Tokyo
2004-2007	Assistant Professor, Genomescience Division, Research Center for Advanced Science and Technology, The University of Tokyo
2007-2013	Assistant and Associate Professor, Department of Pathology, Graduate School of Medicine, The University of Tokyo
2013-2018	Professor, Department of Genomic Pathology, Medical Research Institute, Tokyo Medical and Dental University
2018-present	Professor, Department of Molecular Preventive Medicine, Graduate School of Medicine, The University of Tokyo

CANCER EVOLUTION AND IMMUNE ESCAPE: TRACERx

Charles Swanton
Translational Cancer Therapeutics Lab
The Francis Crick Institute
1 Midland Road, London NW1 1AT, UK
(charles.swanton@crick.ac.uk)

Introduction

Over the last decade, there has been an increasing mismatch between cost and benefit of cancer drugs. In an overview of 12 years of drug approvals by the FDA, of 71 anticancer drugs approved, the median overall survival improvement was only 2.1 months [1]. The cost of cancer drugs has doubled in the last 10 years, now estimated to be over $10,000 per month. Taken together with the modest improvement in overall survival associated with these therapies, it is now estimated that the cost per life year saved has reached $2.7million [2]. Such increments in cost that are not matched by survival improvements or cures in the metastatic setting are unsustainable within socialised health economies in Europe.

Cancer genome instability from the single nucleotide to the whole chromosomal level results in cell-to-cell variation and provides a substrate for selection, ultimately contributing to intratumour heterogeneity and cancer evolution. Intratumour heterogeneity is now considered to be one of the major hurdles in cancer drug development and results in the acquisition of cancer drug resistance, treatment failure and death, contributing to the failure of new therapies to deliver meaningful survival benefits and the excessive health economic cost of new therapies [3-5].

Through advances in tumour evolutionary genome analysis and longitudinal clinical cancer evolutionary studies (TRACERx and PEACE autopsy studies) combined with detailed analysis of the tumour microenvironment through mass cytometry, the laboratory focussed on firstly, deciphering drivers of cancer genome instability and intercellular heterogeneity that form the substrate upon which selection can act and secondly, understanding how the propagation of unstable cancer genomes is initiated, tolerated and

sensed by the immune system. We are endeavouring to understand cancer evolution by combining our genomic analyses with functional studies to understand the roles of driver events in cancer evolution (both clonal and subclonal) utilising genome editing techniques, live cell microscopy and the generation of animal models that more faithfully recapitulate cancer evolutionary processes. By building on our tumour phylogenetics work through bespoke circulating free DNA monitoring, we are attempting to identify and treat disease earlier when tumour heterogeneity is at a minimum and leverage the host immune response, with a deeper understanding of immune escape pathways, in order to constrain tumour adaptation and maximise tumour cell attrition by targeting multiple clonal neo-antigens.

The Clinical Challenges of Cancer Cell Diversity and Selection

Our cancer evolutionary interests have contributed to defining the extent of branched evolution and selection across solid tumours including renal [6-9], colorectal [10, 11], lung [12-15], glioblastoma [16] and oesophageal cancers [17]. Our work has contributed to the elucidation of branched tumour evolution across tumour types and the prevalence of the APOBEC mutational process contributing to subclonal evolution [18]. We have developed tools to understand mechanisms of immune evasion, that have demonstrated HLA Loss of heterozygosity [15] in 40% of early non-small cell lung cancers. We have found evidence for the spatial and temporal separation of driver and passenger genetic events and on-going subclonal expansions that may confound the success of targeted therapies [8, 9, 15, 19]. We have found extensive evidence for parallel evolution with the same gene, signal transduction pathway or protein complex subject to distinct alterations in different subclones of the same tumour [6, 7, 12, 13, 18]. These data provide evidence for constraints to tumour evolution and suggest the presence of "evolutionary rule books" that might govern the sequence of somatic events present in individual tumours that might be exploitable therapeutically [20, 21].

We have begun to elucidate the impact of intratumour heterogeneity on drug resistance with evidence of heterogeneous tumour responses driven by subclonal driver events, spatially separated within distinct subclones of the same tumour [6]. From a clinical perspective, tumours may harbour both good and poor prognostic signatures in the same tumour, illustrating major challenges to biomarker discovery and qualification for clinical use [22]. In this regard, us and others have found that low frequency events in a tumour that may not be detectable at diagnosis may influence clinical outcome later in the disease course as subclonal events begin to dominate the disease at recurrence [7, 14, 16].

Tracking Cancer Evolution: The TRACERx and PEACE studies

In order to derive deeper insights into cancer evolutionary constraints, the mechanistic basis for cancer genomic instability and the impact of tumour heterogeneity on the host immune response, drug resistance and clinical outcome, we have initiated the TRACERx (**TRA**cking **C**ancer **E**volution through Therapy/**Rx**) national lung and renal cancer evolution studies. Lung TRACERx is a national study that has recruited approximately 700 of the planned 842 patients, employing multi-region and longitudinal deep exome sequencing approaches to decipher the timing of genomic events in relation to distinct genomic instability processes and derive tumour phylogenies for each tumour [19]. The study is collecting prospective clinical data over 10 years (2014-2024) that will be integrated with tumour phylogenetic, immunological and microenvironment analysis to attempt to identify therapeutic and environmental selection pressures that constrain cancer evolution, the origins of the lethal metastatic subclones and mechanisms of immune escape and metastasis. The primary endpoint of TRACERx is to establish whether there is a relationship between genetic intratumour heterogeneity and clinical outcome.

Analysis of the first 100 patients in TRACERx lung has revealed that chromosomal instability (CIN) assessed by somatic copy number heterogeneity is a major determinant of clinical outcome independent of tumour stage [13]. We have developed an approach to quantify chromosome instability from cancer exome TRACERx datasets through haplotype and allelic imbalance analysis at subclonal resolution that is revealing the role of CIN later in tumour evolution.

Using these tools, TRACERx has revealed four findings that may contribute to the poor prognosis associated with CIN. Firstly, through the LOHHLA software tool developed in our laboratory, we have found that CIN results in copy loss of class I HLA alleles and permits branched evolution through the expansion of subclones harbouring neo-antigens predicted to have bound to the lost HLA allele [15] and loss of clonal neoantigens (Rosenthal manuscript in revision). Secondly through mirrored subclonal allelic imbalance (MSAI), and parallel evolution of driver amplification events occurring recurrently in distinct subclones of the same tumour[13]. Thirdly, renal TRACERx studies have revealed that CIN results in DNA copy number loss of chromosomes 9p and 14q, high risk events selected during metastatic dissemination [8, 9, 23]. Finally, we have found that genome doubling, that occurs as an early event in >70% of patients with primary NSCLC [13], accelerates cancer evolution and the propagation of aneuploid cancer cell populations, accelerating cancer cell diversity [11].

Developing Insights into the Primary-Metastatic Transition

Prompted by our preliminary findings of contingencies and constraints to renal cancer evolution converging upon PI3K pathway activation [6, 20], we initiated the Renal TRACERx study. In this study we have been investigating tumour evolution from primary to metastatic sites through extensive multi-region sequencing analysis of 100 prospectively recruited patients with clear cell carcinoma of the kidney and two validation cohorts of primary-metastases pairs from HUC (Hospital Unversitario Cruces) and MSKCC [8, 9, 23]. A principle of the TRACERx studies is that deep analysis of the primary tumour and metastatic sites will enable the metastatic competent subclone to be distinguished from metastatic incompetent subclones, that may shed light on the cell intrinsic processes inherent to tumour metastasis and latency.

We found no evidence for the selection of single mutational driver events in the metastatic competent subclone. However, cancer cell ploidy, chromosomal complexity (wGII score: weighted genome instability index; the average proportion of the genome with aberrant copy number weighted on each of the 22 autosomal chromosomes) and proliferation indices were enriched in metastatic competent compared to incompetent subclones. Loss of chromosome 9p encoding CDKN2A/B was the most frequently selected event at the metastatic transition. Comparing the clinical phenotype to genomic patterns of evolution in the TRACERx cohort, we find that tumours dominated by high chromosomal complexity selected early in tumour evolution in the most recent common ancestor and low driver event heterogeneity, are associated with more rapid metastatic dissemination to multiple organ sites. In contrast, tumours with late onset chromosomal complexity (i.e. present subclonally) and with higher intratumour heterogeneity (assessed by the ratio of subclonal drivers:clonal drivers) were associated with a more protracted clinical course and dissemination to single sites of disease over years [8, 9]. Therefore, both TRACERx studies in renal and NSCLC have revealed the importance of chromosomal instability rather than point mutational driver diversity as the major factor associated with poor metastasis free survival. Importantly, these data suggest that a detailed understanding of cancer evolution may shed light on distinct patterns of future clinical behaviour and outcome.

Clonal Neoantigens: relevance to tumour control and checkpoint inhibitor response

An increased understanding of cancer evolution has led to an appreciation of the importance of targeting clonal events that are present in every tumour cell [18, 24], and the requirements for new cancer biomarkers to account for tumour heterogeneity. Early results from the TRACERx program indicate that therapeutically relevant activating mutations in EGFR and rearrangements in EML4-ALK are always clonal founder events, present in every tumour cell [13]. Despite progress in targeted therapy development in non-small cell

lung cancer, and the ability to target single driver founder events effectively, progression of disease is inevitable in almost all patients due to selection of subclones harbouring resistance somatic events (eg T790M gatekeeper mutation following EGFR tyrosine kinase inhibitor resistance [25]). These clinical data indicate that targeting a single genetic driver of disease is unlikely to prove sufficient to maintain long-term disease control in a patient with advanced lung cancer. An ideal therapeutic would target multiple clonal events in an individual tumour, present and specific to every tumour cell, whilst minimising normal tissue toxicity.

Our TRACERx work over the last 3 years indicates that the high mutational burden present in a tobacco-associated non-small cell lung cancer (accounting for >40,000 deaths in the UK per year) may represent a specific tumour vulnerability that may be taken advantage of by augmenting the host immune response. We have found evidence that the burden of clonal mutations encoding tumour HLA Class I putative neo-antigens, is associated with improved survival outcome in the absence of immune-therapy intervention in both a TCGA discovery cohort [26] and the TRACERx validation cohort; patients with tumours harbouring a high clonal neo-antigen burden (upper quartile) rarely suffer metastatic relapse following surgery [26] and (Rosenthal et al manuscript in revision).

We have found that insertion-deletion frameshift mutations (fs-indels) likely comprise a potent class of neo-antigen for therapeutic exploitation and the burden of such mutations appears to correlate with checkpoint inhibitor response [27]. Intriguingly, renal cell carcinoma has a particularly high burden of fs-indels that may explain the unique sensitivity of this tumour type to immunotherapy, despite its overall low mutational burden.

These data suggest that the clonal or subclonal status of a tumour neo-antigen is important, with clonal neo-antigens present in every tumour cell representing a major cancer cell vulnerability subject to immune surveillance. Similarly, patients with relatively homogeneous tumours with a large clonal burden appear to benefit from anti-PD1 or anti-CTLA4 immune checkpoint inhibitor treatment [26], recently independently validated by the Van Allen laboratory [28]. The importance of targeting clonal neo-antigens has also been emphasised by animal studies from the Scheinberg laboratory demonstrating that the clonal fraction of cells expressing a neo-antigen is important for optimal immune-mediated control; immune rejection is impaired when the fraction of cells harbouring a neo-antigen is low [29]. These data suggest that optimal immune surveillance may not occur until a cancer population size is reached, sufficient to present a "dose" of neo-antigen required for immune activation. This argument, taken to its logical conclusion, implies that immune surveillance early in tumour evolution may be sub-optimal due to low neo-antigen dosage for early control of tumour initiation.

In a collaboration with the Quezada and Hadrup laboratories, through multi-region tumour sequencing and phylogenetic and MHC tetramer analysis, we have detected CD8+ T cells that recognise clonal neo-antigens present in every tumour cell [26, 30]. Analysis of patients from TRACERx, classified by smoking status revealed that tobacco associated NSCLC harbour 5-10 fold more clonal non-synonymous mutations compared to non-Tobacco associated lung cancer suggesting this group of patients may be ideal candidates for adoptive cell therapy or vaccine programs targeting clonal neo-antigens.

This work has resulted in the founding of a Francis Crick Institute biotechnology company, Achilles Therapeutics, which through the Wellcome Trust and CRUK is attempting to transform our laboratory findings into clinical practice and patient benefit. In this approach, multiple, unique clonal events in lung cancers are targeted through immunotherapy (vaccine and T cell therapy) with the hope of limiting the evolution of resistance and avoiding tumour growth. We intend to expand multiple distinct T cell clones reactive against clonal neoantigens, in a patient-specific manner ex vivo, and reinfuse them within two first-in-human clinical trials in melanoma and NSCLC sponsored by Achilles Therapeutics in 2019

Immune Editing of Neo-antigen presentation during Tumour Evolution

Given emerging evidence of the importance of clonal neo-antigens in tumour evolution and immune surveillance, we have endeavoured to understand mechanisms of immune escape during NSCLC and renal cancer evolution. Loss of human leukocyte antigens has long been known to occur in cancer [31]. However, its prevalence, timing, and relationships with clonal expansions in solid tumours was unclear. The polymorphic nature of the HLA locus had prevented accurate HLA allele specific copy number assessment. With this in mind, Mcgranahan and Rosenthal in the laboratory developed a software tool- LOHHLA- which uses the patient's own germline reference reads on chromosome 6p to infer allele specific copy number of class I HLA from tumour exome data. Through this approach we found that HLA class I loss occurs in up to 40% of early stage, untreated NSCLC [15], approximately 10 times more frequently than B2M mutations. Intriguingly, HLA class I LOH commonly occurs as a subclonal event, consistent with the hypothesis that an optimal cancer cell population size driving increased antigen dosage is required before an effective T cell response can be mounted to limit tumour growth [29]. We find that subclones harbouring HLA LOH appear to have a higher non-synonymous mutational burden than closely related subclones in the tumour that have all 6 alleles intact, consistent with the permissive role of HLA LOH in ongoing branched evolution. Our future TRACERx work will explore how clonal neo-antigens are recognised, repressed or deleted during the disease course, in order to inform clonal neo-antigen directed approaches to limit tumour growth.

References

1. Fojo, T., S. Mailankody, and A. Lo, Unintended consequences of expensive cancer therapeutics-the pursuit of marginal indications and a me-too mentality that stifles innovation and creativity: the John Conley Lecture. JAMA Otolaryngol Head Neck Surg, 2014. 140(12): p. 1225-36.
2. Kantarjian, H.M., et al., Cancer drugs in the United States: Justum Pretium--the just price. J Clin Oncol, 2013. 31(28): p. 3600-4.
3. Yates, L.R. and P.J. Campbell, Evolution of the cancer genome. Nat Rev Genet, 2012. 13(11): p. 795-806.
4. Burrell, R.A., et al., The causes and consequences of genetic heterogeneity in cancer evolution. Nature, 2013. 501(7467): p. 338-45.
5. Siravegna, G., et al., Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. Nat Med, 2015. 21(7): p. 827.
6. Gerlinger, M., et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med, 2012. 366(10): p. 883-92.
7. Gerlinger, M., et al., Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet, 2014. 46(3): p. 225-33.
8. Turajlic, S., et al., Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. Cell, 2018. 173(3): p. 581-594 e12.
9. Turajlic, S., et al., Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. Cell, 2018. 173(3): p. 595-610 e11.
10. Burrell, R.A., et al., Replication stress links structural and numerical cancer chromosomal instability. Nature, 2013. 494(7438): p. 492-6.
11. Dewhurst, S.M., et al., Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. Cancer Discov, 2014. 4(2): p. 175-85.
12. de Bruin, E.C., et al., Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science, 2014. 346(6206): p. 251-6.
13. Jamal-Hanjani, M., et al., Tracking the Evolution of Non-Small-Cell Lung Cancer. N Engl J Med, 2017. 376(22): p. 2109-2121.
14. Abbosh, C., et al., Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature, 2017. 545(7655): p. 446-451.
15. McGranahan, N., et al., Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell, 2017. 171(6): p. 1259-1271 e11.
16. Favero, F., et al., Glioblastoma adaptation traced through decline of an IDH1 clonal driver and macroevolution of a double minute chromosome. Ann Oncol, 2015.
17. Murugaesu, N., et al., Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov, 2015.

18. McGranahan, N., et al., Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. Sci Transl Med, 2015. 7(283): p. 283ra54.
19. Jamal-Hanjani, M., et al., Tracking Genomic Cancer Evolution for Precision Medicine: The Lung TRACERx Study. PLoS Biol, 2014. 12(7): p. e1001906.
20. Fisher, R., et al., Development of synchronous VHL syndrome tumors reveals contingencies and constraints to tumor evolution. Genome Biol, 2014. 15(8): p. 433.
21. McGranahan, N. and C. Swanton, Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. Cancer Cell, 2015. 27(1): p. 15-26.
22. Gulati, S., et al., Systematic Evaluation of the Prognostic Impact and Intratumour Heterogeneity of Clear Cell Renal Cell Carcinoma Biomarkers. Eur Urol, 2014.
23. Mitchell, T.J., et al., Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell, 2018. 173(3): p. 611-623 e17.
24. Yap, T., et al., Intratumour Heterogeneity: Seeing the wood for the trees. Science Translational Medicine, 2012. In press 2012.
25. Piotrowska, Z., et al., Heterogeneity Underlies the Emergence of EGFRT790 Wild-Type Clones Following Treatment of T790M-Positive Cancers with a Third-Generation EGFR Inhibitor. Cancer Discov, 2015. 5(7): p. 713-22.
26. McGranahan, N., et al., Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science, 2016. 351(6280): p. 1463-9.
27. Turajlic, S., et al., Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. Lancet Oncol, 2017. 18(8): p. 1009-1021.
28. Miao, D., et al., Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat Genet, 2018. 50(9): p. 1271-1281.
29. Gejman, R.S., et al., Rejection of immunogenic tumor clones is limited by clonal fraction. Elife, 2018. 7.
30. Bentzen, A.K., et al., Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. Nat Biotechnol, 2016. 34(10): p. 1037-1045.
31. Campoli, M. and S. Ferrone, HLA antigen changes in malignant cells: epigenetic mechanisms and biologic significance. Oncogene, 2008. 27(45): p. 5869-85.

Charles Swanton, MD, PhD, FMedSci, FRS

1999	Completed MD, PhD Training, Imperial Cancer Research Fund, UK (this became Cancer Research UK in 2002)
2008	Cancer Research UK Clinician Scientist/Medical Oncology Training Established Lab at the London Research Institute, Cancer Research UK
2011	Appointed Chair of Personalised Cancer Medicine at UCL Hospitals
2015	Group Leader, the Francis Crick Institute, London
2016	Appointed Royal Society Napier Professor of Cancer Medicine

QUANTIFYING AND FORECASTING TUMOR PROGRESSION

Christina Curtis
Departments of Medicine and Genetics, Stanford University School of Medicine
265 Campus Drive, Suite G2120C, Stanford, CA 94305, USA
(cncurtis@stanford.edu)

Cancer results from the acquisition of somatic alterations in an evolutionary process that typically occurs over many years, much of which is occult. Understanding the evolutionary dynamics that are operative at different stages of progression in individual tumors might inform the earlier detection, diagnosis, and treatment of cancer. Although direct observations of human tumor evolution are impractical, the spatiotemporal patterns of somatic alterations amongst cells within a tumor faithfully encode their evolutionary histories.

Whereas it has traditionally been assumed that tumor progression results from ongoing sequential selection for 'driver' mutations that confer a stringent fitness advantage, unexpectedly, we recently found that after transformation, some tumors grow in the absence of stringent selection, compatible with *effectively neutral* evolution. This led to our description of a *Big Bang* model of colorectal tumor growth where the neoplasm grows as a terminal expansion propagated by numerous heterogeneous—and effectively equally fit subclones [1] (Figure 1). This new model explains the early origins of intra-tumor heterogeneity (Figure 2) and the dynamics of tumor growth with implications for earlier detection, treatment resistance and metastasis. For example, we demonstrate that subclonal variants are largely undetectable using current sequencing strategies, and yet may provide a rich substrate for the emergence of resistance under treatment selective pressure. Moreover, the data suggest that some tumors are *born to be* bad, wherein malignant potential is specified early.

Big Bang dynamics have since been reported in other tumor types, implying that *effective neutrality* is relatively common and that at least two dominant 'modes' of evolution

are operative in established solid tumors. Despite the implications of tumor dynamics for precision medicine, methods to infer the strength of selection from cancer genomic data are lacking. We therefore developed a population genetics framework to simulate spatial tumor growth and infer the mode of evolution from patient genomic data based on patterns of genetic divergence [2, 3] (Figure 3). By applying this approach to diverse solid tumors, we demonstrate variability in the mode of evolution *within* and *between* tumor types and at different stages of disease (Figure 4). Further, we illustrate the utility of this classification scheme for delineating the true 'drivers' of tumor progression.

Building on our established tumor evolutionary dynamic framework, we subsequently developed a theoretical and computational approach to 'time' metastasis from primary tumors and paired metastases. This new method leverages a 3-dimensional computational model to simulate the spatial growth of realistically sized tumors (composed of ~10^9 cells) and a robust statistical inference framework to measure evolutionary parameters from genomic data in a patient-specific manner [4]. Application of this quantitative method to paired primary colorectal cancers and distant metastases enables the systematic analysis of the rates and routes of metastasis and reveals fundamental insights into the drivers of this lethal process with implications for the approach to systemic therapy and earlier detection.

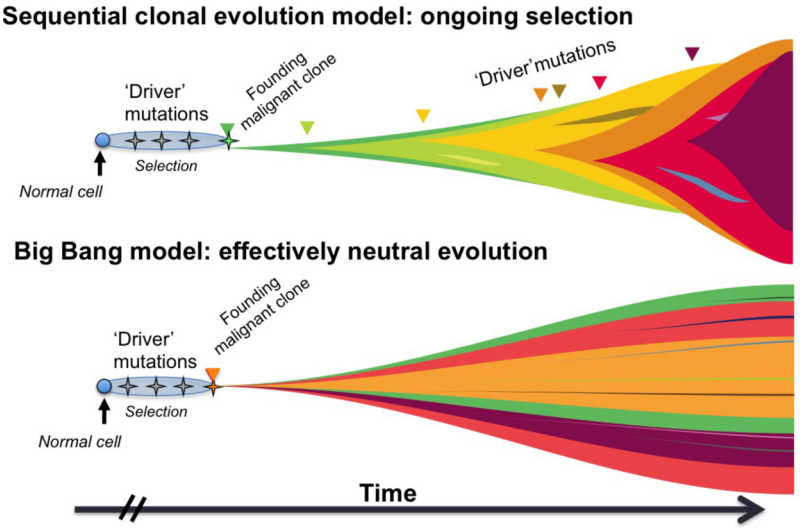


Figure 1 Schematic illustration of the sequential clonal evolution model versus Big Bang model of effectively neutral tumor evolution
In both models driver mutations are strongly selected for during tumor initiation ultimately leading to transformation of the founding malignant clone. In the sequential clonal evolution model, ongoing subclonal selection and clonal expansion occurs *after* transformation as a result of additional 'driver' alterations that confer a fitness advantage. In contrast, in the Big Bang model, after transformation additional mutations accrue but do not result in detectable subclonal expansions because the background population is already highly fit (adapted from Sottoriva et al, *Nature Genetics*, 2015).

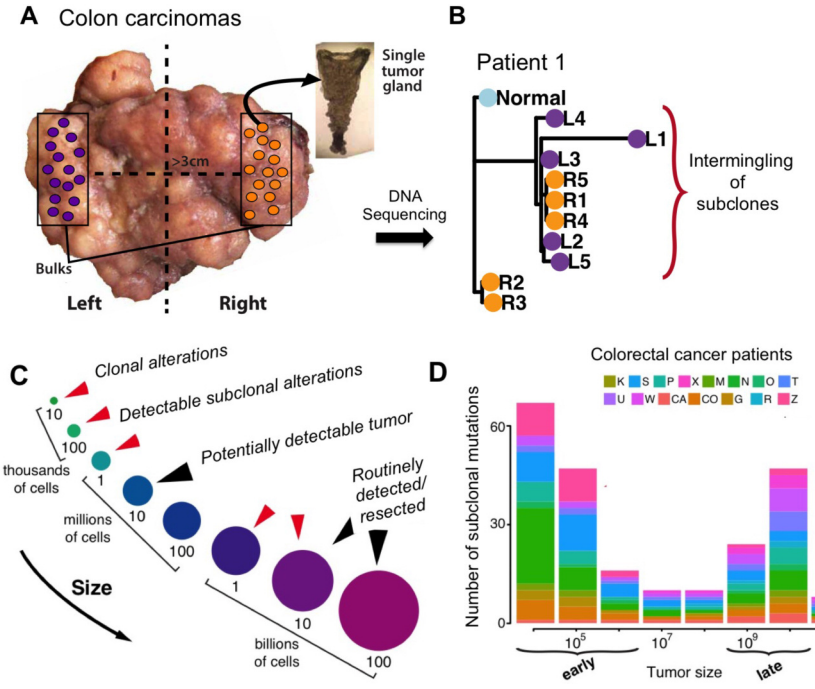


Figure 2 Multi-region profiling of colorectal cancers and statistical inference on the genomic data verify the predictions of the Big Bang model
A. Multi-region sampling and integrated genomic profiling. B. Single gland copy number profiles reveal extensive ITH and subclone mixing. Collectively, the genomic data indicate that recent clonal expansions are rare and subclone mixing is common in CRC. C. Inference of the mutational timeline for classes of somatic alterations indicates that the majority of *detectable* subclonal alterations occur early during growth (data from Sottoriva et al. *Nature Genetics,* 2015).

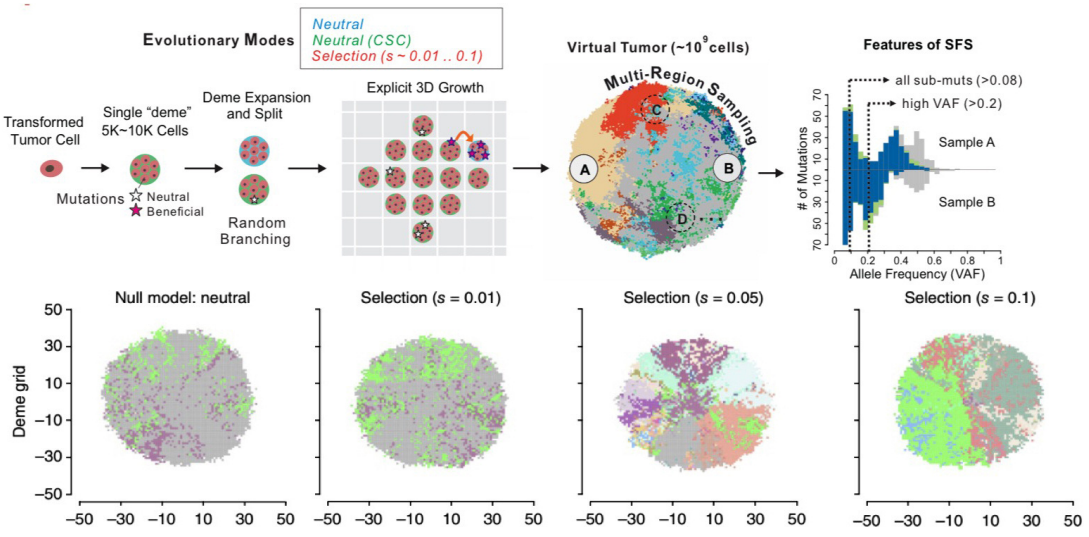


Figure 3 Schematic overview of spatial tumor growth model (from Sun et al. *Nature Genetics,* 2017)

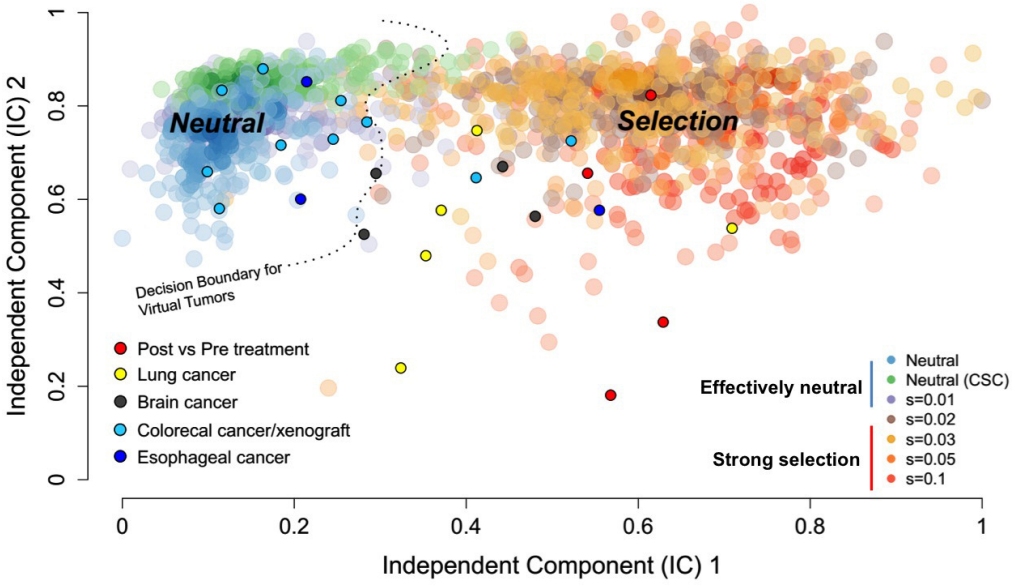


Figure 4 Classification of the 'mode' of tumor evolution for diverse solid tumors
Independent component analysis of virtual tumors (transparent circles) and patient tumors (small filled circles) based on multiple measures of ITH distinguishes tumors that are evolving in an effectively neutral fashion, from those that are evolving under stringent subclonal selection (from Sun et al. *Nature Genetics*, 2017).

References

1. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon M, Zhao J, Marjoram P, Siegmund K, Press MF, Shibata D and Curtis C. *A Big Bang model of human colorectal tumor growth.* Nature Genetics, 2015, 47(3)209-216.
2. Sun R, Hu Z, Sottoriva A, Graham TA, Ma Z, Fischer JM, Shibata D and Curtis C., *Between-region genetic divergence reflects the mode and tempo of tumor evolution.* Nature Genetics, 2017, 49(7):1015-1024.
3. Hu, Z., R. Sun, and C. Curtis, *A population genetics perspective on the determinants of intra-tumor heterogeneity.* Biochim Biophys Acta – Reviews on Cancer, 2017, pii: S0304-419X(17)30046-X.
4. Hu Z, Ding J, Ma Z, Sun R, Seoane JA, Suarez CJ, Berghoff AS, Cremolini C, Falcone A, Loupakis, Birner P, Preusser M, Lenz HJ and Curtis C. Quantitative evidence for early metastatic seeding in colorectal cancer. *In revision*

Christina Curtis, PhD, MSc

2003-2007	Graduate Student, University of Southern California
2007-2010	Postdoctoral Fellow, University of Cambridge
2010-2014	Assistant Professor, University of Southern California, Keck School of Medicine
2014-present	Assistant Professor, Stanford University School of Medicine
2014-present	Co-director, Molecular Tumor Board, Stanford Cancer Institute
2018	NIH Director's Pioneer Award Recipient

NOVEL EVOLUTIONARY PRINCIPLE SHAPING INTRATUMOR HETEROGENEITY IN COLORECTAL CANCER

Koshi Mimori
Department of Surgery
Kyushu University Beppu Hospital
4546 Tsurumihara, Beppu 874-0838, Japan
(kmimori@beppu.kyushu-u.ac.jp)

Cancer evolution and intratumor heterogeneity (ITH) have attracted increasing attention in the cancer research field because ITH generated during cancer evolution presumably contributes to the therapeutic and diagnostic difficulties of cancer. The recent technological innovations enabled us to conduct the multiregion sequencing approach, which has been popularly used to understand ITH.

In the view point of the development of colorectal cancer (CRC), adenoma first forms a polyp and then partially progress to early carcinoma, which subsequently grows beyond the muscularis mucosa to invade surrounding tissues. To examine ITH in advanced CRC (ACRC), we previously implemented multiregion sequencing of nine locally advanced or metastatic tumors [1]. While most of the known driver events represented by APC and KRAS mutations were observed as ubiquitous mutations, branched or parallel evolution was rarely observed in evolutionary histories of ACRC. By additionally performing a computational simulation of cancer evolution, we demonstrated possibility that ITH in ACRC could be generated by neutral evolution [2].

Subsequent to our previous study, we perform multiregion whole-exome sequencing on 10 precancerous lesions of colorectal cancers (PCRCs), which contained adenoma and carcinoma in situ [3]. By comparing with sequencing data from ACRC, we show that the early tumors accumulate a higher proportion of subclonal driver mutations than the advanced tumors, which is highlighted by subclonal mutations in KRAS and APC. We also demonstrate that variant allele frequencies of subclonal mutations tend to be higher in early tumors, suggesting that the subclonal mutations are subject to selective sweep in early tumorigenesis while neutral evolution is dominant in advanced ones. This study

establishes that the evolutionary principle underlying intratumor heterogeneity shifts from Darwinian to neutral evolution during colorectal tumor progression.

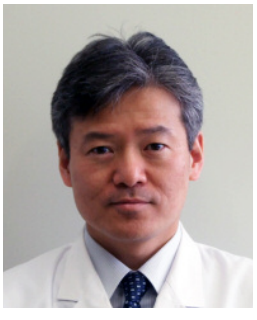
In addition to the SNV study, copy number aberrations were progressively increased from adenoma through early carcinoma to ACRC. The increase in ubiquitous CNAs was especially prominent in ACRC. Therefore, CNAs play more critical roles in the progression from PCRC to ACRC.

Multi-regional sequencing analysis of solid clinical samples provides a major breakthrough in disclosing ITH. The level of uniformity depends on the type of cancer, and the causes of diversity vary among cancers. The actual application of the findings in our studies for clinical diagnosis and treatment might require furthermore time to save patients of intractable cancers eventually.

In terms of "cancer evolution", I have proposed an alternative evolution pathway which was engendered by depressed type tumor other than adeno-carcinoma sequence. At last, I have introduced "therapy driven pathway" which was observed in recurrence tumor after the chemoradiotherapy in esophageal cancer cases. The shape of evolution looks like a punctuated manner.

References

1. Uchi R, Takahashi Y, Niida A, Shimamura T, Hirata H, Sugimachi K, Sawada G, Iwaya T, Kurashige J, Shinden Y, Iguchi T, Eguchi H, Chiba K, Shiraishi Y, Nagae G, Yoshida K, Nagata Y, Haeno H, Yamamoto H, Ishii H, Doki Y, Iinuma H, Sasaki S, Nagayama S, Yamada K, Yachida S, Kato M, Shibata T, Oki E, Saeki H, Shirabe K, Oda Y, Maehara Y, Komune S, Mori M, Suzuki Y, Yamamoto K, Aburatani H, Ogawa S, Miyano S, Mimori K. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. *PLoS Genet.* 2016 12(2):e1005778..
2. Niida A, Nagayama S, Miyano S, Mimori K. Understanding intratumor heterogeneity by combining genome analysis and mathematical modeling. *Cancer Sci.* 2018 Apr;109(4): 884-892.
3. Saito T, Niida A, Uchi R, Hirata H, Komatsu H, Sakimura S, Hayashi S, Nambara S, Kuroda Y, Ito S, Eguchi H, Masuda T, Sugimachi K, Tobo T, Nishida H, Daa T, Chiba K, Shiraishi Y, Yoshizato T, Kodama M, Okimoto T, Mizukami K, Ogawa R, Okamoto K, Shuto M, Fukuda K, Matsui Y, Shimamura T, Hasegawa T, Doki Y, Nagayama S, Yamada K, Kato M, Shibata T, Mori M, Aburatani H, Murakami K, Suzuki Y, Ogawa S, Miyano S, Mimori K. A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. *Nat Commun.* 2018 Jul 23;9(1):2884.



Koshi Mimori, MD, PhD

1991	Graduated from the Medical College of Oita
1994-1997	Resident, Department of Surgery, Medical Institute of Bioregulation, Kyushu University
1997-2000	Postdoctoral Fellow, Kimmel Cancer Institute, Thomas Jefferson University in Philadelphia
2000-2006	Research Associate, Department of Surgery, Medical Institute of Bioregulation, Kyushu University
2010-2011	Associate Professor, Department of Surgery, Kyushu University Beppu Hospital
2012-present	Professor, Department of Surgery, Kyushu University Beppu Hospital

WHOLE GENOME SEQUENCING AND IMMUNO-GENOMIC ANALYSIS FOR LIVER CANCER

Hidewaki Nakagawa
Laboratory for Cancer Genomics, RIKEN Center for Integrative Medical Sciences
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
(hidewaki@ims.u-tokyo.ac.jp)

Cancer is essentially a "disease of the genome" which develops and evolves with the accumulation of a variety of mutations, based on the background of its genomic instability, and some driver mutations were successfully targeted for treatment. Cancer also has been proved to have a feature of "immune reaction" and have been affected by immune editing in carcinogenic steps. Now immune therapies are a real in most types of cancer. To explore whole genomic pictures and immuno-genomic features of cancer, we have been addressing cancer whole genome sequencing (WGS) analysis for liver cancer and other types of cancer. These approaches combined with mathematical analysis and other -omics analysis can clarify the underlying carcinogenesis and cancer immunology and achieve molecular sub-classification of cancer, which facilitates discovery of genomic biomarkers and personalized cancer medicine. I here present discovery from cancer WGS and immuno-genomic analysis of liver cancers, and discuss its utility and limitation of an analysis platform and mutation interpretation for cancer genomics and cancer immunology.

(A) Whole genome sequencing analysis of liver cancer

As one of the Japanese ICGC projects, we sequenced whole genomes of 300 liver cancer, which were manly affected by virus infection [1]. The median number of somatic mutation of liver cancer was approximately 10,000. We identified several mutated driver genes and pathways in liver cancer, including *TERT*, *TP53* pathway, Wnt/*CTNNB1* pathway, and *ARID1A*/*ARID2*, which is summarized in Figure 1. We found several non-coding mutational clusters, such as *TERT* promoter, *NEAT1*/*MALAT* lincRNA, and *WDR74* promoter, and genome-wide mutational clustering analysis also found several mutational

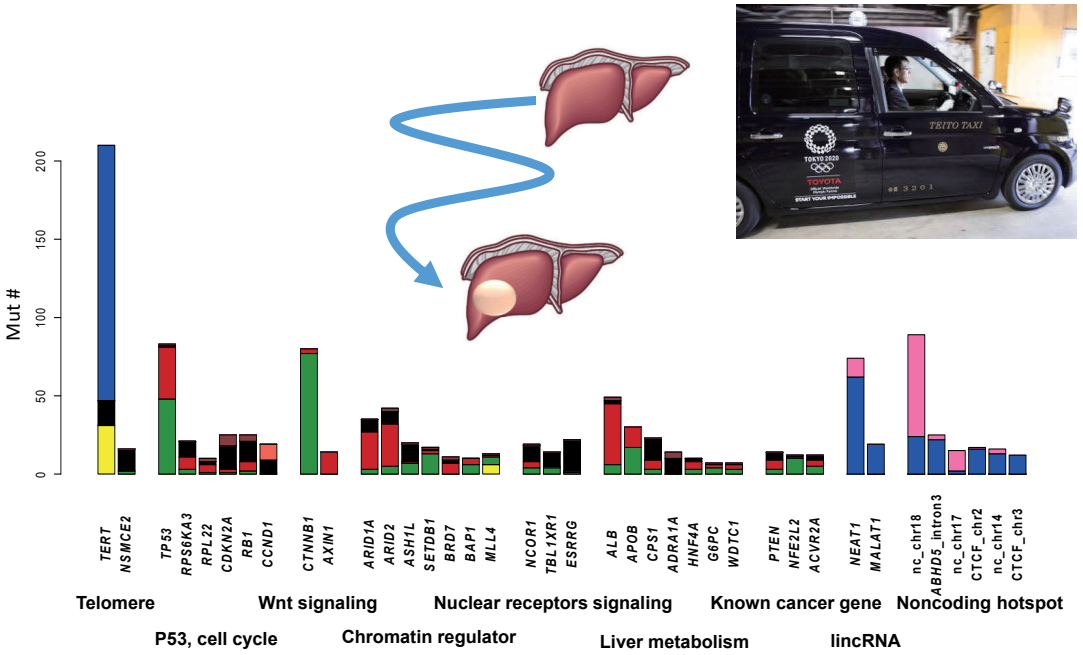


Figure 1 Summary of liver cancer Driver genes
Fujimoto, et al. Nat Genet 48:500-509, 2016

clusters in CTCF-binding sites. WGS and RNA-seq detected virus integrations of HBV and new virus, AAV, in liver cancer genome. Additional deep sequencing analysis targeting HBV also showed 1,684 HBV integration sites in cancer and liver tissues, which preferentially occurred in the open chromatin regions and mitochondria genome in mouse HBV infection model [2]. As one of the drivers, *TERT* promoter regions were frequently affected by the hotspot mutations, SV, and virus integrations in liver cancer. From WGS data, we extracted several mutational signature such as smoking signature (Sig4) and alcohol signature (Sig16), and estimated cell-of-origin through whole genome mutation distribution, which indicated some of intrahepatic cholangiocarcinomas are likely to be originated from hepatocyte, as well as HCC [3].

(B) Pan-cancer immuno-genomic analysis (PCAWG-15)

The PCAWG (PanCancer Analysis of Whole Genomes) study is an international collaboration to identify common patterns of mutation in more than 2,800 cancer whole genomes from the International Cancer Genome Consortium. As PCAWG-15 group, we characterized the immunological feature of theses PCAWG samples by analyzing HLA genotype, immuno-signature, deconvolution of immune cells, neo-antigen prediction using WGS and RNA-Seq data (Figure 2). We developed the algorism to precisely call HLA

genotype from WGS and found somatic mutations and copy number change of HLA (chr6), which frequently occurred in colorectal cancer, lymphoma, and pancreatic cancer. SV affected 10 immune-related genes, including *PD-L1* and *PD-L2*, and lead to their overexpression in some types of tumor. To estimate the immuno-editing history of each cancer genome, we defined the immuno-editing index (IEI) by comparing the number of neo-antigens in pseudogenes with that in exonic regions. IEI analysis indicated that immuno-edited tumor (IEI-low) was enriched in MSI-positive CRC, and IEI was statistically related to overall survival for lung cancer and pancreatic cancer [4].

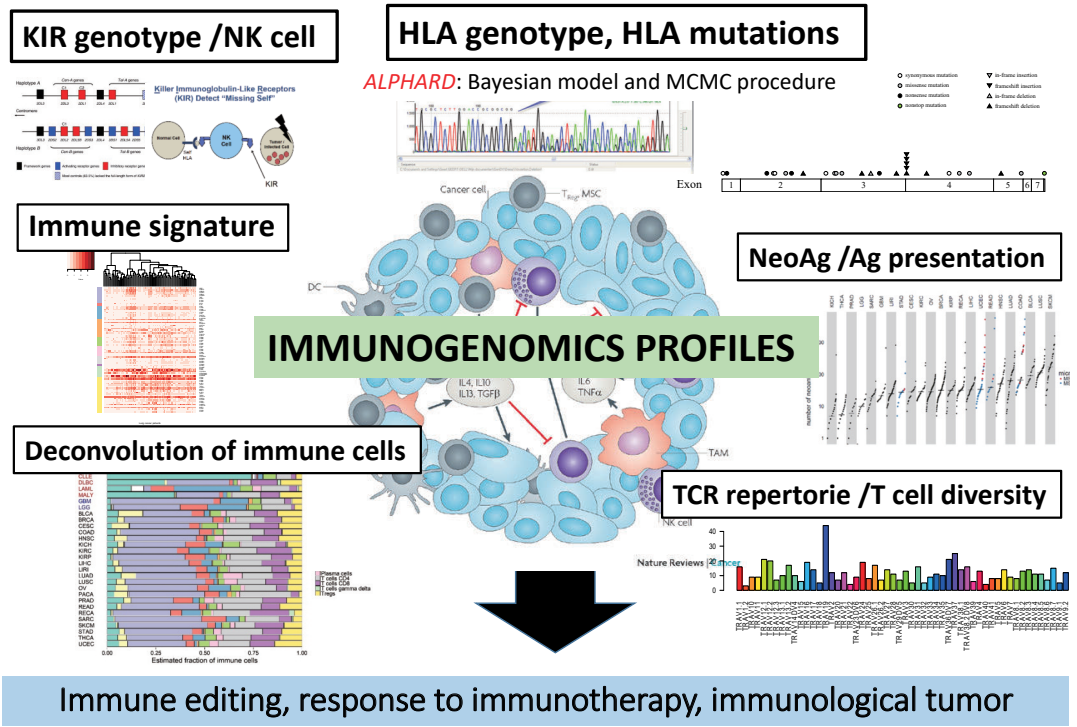


Figure 2 PCAWG-15 HLA/Immunogenomics Group

(C) Immuno-genomic analysis of liver cancer

Liver cancer develops in chronic hepatitis where various types of immune cells are activated and suppressed. Although the background liver is highly inflamed, liver cancer is generally considered as immune suppressive. We analyzed RNA and WGS of 234 liver cancers and matched non-tumorous livers with chronic hepatitis, and characterized their immunological feature by comparing the immune profiles in liver cancers and hepatitis livers. Anti-tumor immunity was associated with significantly better prognosis. Tumor had

lower expression levels of immune genes than adjacent hepatitis liver, indicating predominant immune suppression in tumor. Gene signature for Treg were overexpressed in tumor. Three tumor subclasses of proliferation, CTNNB1, interferon were immunologically characterized by high Treg signature, less immune infiltrates, and deficient immune suppression, respectively.

References

1. Fujimoto A, Furuta M, Totoki Y, et al. Whole genome mutational landscape and characterization of non-coding and structural mutations in liver cancer. Nat Genet 2016;48:500-509.
2. Furuta M, et al. Characterization of HBV integration patterns and timing in liver cancer and HBV-infected liver. Oncotarget 2018;9:25075-25088.
3. Wardell CP, Fujita M, et al. Genomic characterization of biliary tract cancers identifies their driver genes and predisposing mutations. J Hepatol 2018;68:959-969,
4. Mizuno S, Yamaguchi R, Hasegawa T, Hayashi S, Fujita M, Zhang F, Koh Y, Lee SY, Yoon SS, Shimizu E, Komura M, Fujimoto A, Nagai M, Kato M, Liang H, Miyano S, Zhang Z, Nakagawa H, Imoto S, on behalf of the PCAWG Mitochondrial Genome and Immunogenomics Working Group and The PCAWG Network. Immuno-genomic Pan-Cancer landscape reveals diverse immune escape mechanisms and immuno-editing histories. *bioRxiv* 285338 (2018). doi:10.1101/285338

Hidewaki Nakagawa, MD, PhD

1991	MD, Osaka University, School of Medicine
1991-1992	Resident, Osaka University Hospital, General Surgery
1992-1993	Resident, Osaka University Hospital, ICU
1993-1996	Resident, National Osaka Hospital, GI and Breast Surgey
1996-1999	PhD, Osaka University, Graduate School of Medicine
1999-2003	Postdoc Fellow, The Ohio State University, Human Cancer Genetics Program
2003-2008	Assitant Professor, Institute of Medical Science, The University of Tokyo
2008-2013	Team Leader, RIKEN Center for Genomic Medicine
2013-present	Team Leader, RIKEN Center for Integrative Medical Sciences

GENETIC ALTERATIONS IN T-CELL LYMPHOMAGENESIS

Keisuke Kataoka
Division of Molecular Oncology
National Cancer Center Research Institute
5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(kekataok@ncc.go.jp)

Peripheral T-cell lymphomas (PTCLs) represent a clinically, histologically, and molecularly heterogeneous group of non-Hodgkin lymphomas derived from mature post-thymic T cells. Among them, the most common entity in Japan is adult T-cell leukemia/lymphoma (ATL), which is an aggressive peripheral T-cell lymphoma associated with human T-cell leukemia virus type-1 (HTLV-1) infection. To delineate the genetic landscape of ATL, we carried out an integrated molecular study, in which whole genome, exome, transcriptome, and targeted resequencing, as well as array-based copy number and methylation analyses were performed in more than 400 ATL cases [1]. We found recurrent genetic alterations in T-cell receptor/NF-κB signaling, T-cell trafficking, and other T-cell-related pathways as well as immunosurveillance. A conspicuous feature of ATL genome is the predominance of gain-of-function alterations, including activating mutations (in *PLCG1*, *PRKCB*, *CARD11*, *VAV1*, *IRF4*, *CCR4,* and *CCR7*) and gene fusions (*CTLA4-/ICOS-CD28*). We also discovered frequent intragenic deletions involving *IKZF2*, *CARD11* and *TP73* and mutations in *GATA3*, *HNRNPA2B1*, *GPR183*, *CSNK2A1*, *CSNK2B* and *CSNK1A1*. Besides these alterations, we identified a unique genetic mechanism of immune evasion caused by structural variations (SV) disrupting 3'-untranslated region (UTR) of the *PD-L1* gene [2]. These SVs were also found in multiple common human cancer types, such as diffuse large B-cell lymphoma and stomach adenocarcinoma, and caused a marked elevation of aberrant *PD-L1* transcripts, leading to immune evasion of tumor cells *in vivo*.

Aggressive (acute/lymphoma) subtypes were associated with an increased burden of genetic and epigenetic alterations, higher frequencies of *TP53* and *IRF4* mutations, and many copy number alterations (CNAs), including *PD-L1* amplifications and *CDKN2A*

deletions, compared with indolent (chronic/smoldering) subtypes [3]. By contrast, *STAT3* mutations were more characteristic of indolent ATL. Higher numbers of somatic mutations and CNAs significantly correlated with worse survival. In addition, ATL subtypes are further classified into molecularly distinct subsets with different prognosis by genetic profiling. Particularly, somatic alterations characterizing aggressive diseases predict worse prognosis in indolent ATL, among which *PD-L1* amplifications are a strong genetic predictor in both aggressive and indolent ATL.

Among PTCLs, the most common in Western countries is PTCL, not otherwise specified (NOS), accounting for approximately 30% of all PTCLs. Combining whole-exome and deep targeted-capture sequencing of 133 cases, we delineated the entire picture of genetic alterations in PTCL, NOS. Of note is the identification of a previously undescribed molecular subtype characterized by *TP53* and/or *CDKN2A* mutations and deletions in PTCL, NOS without showing a T follicular helper cell phenotype. This subtype exhibited different prognosis and unique genetic features, including extensive chromosomal instability, which preferentially affected molecules involved in immune escape and transcriptional regulation.

Taken together, our findings provide novel insights into genetic and molecular heterogeneity in PTCLs, which should help to devise a novel molecular classification and to exploit a new therapeutic strategy for these malignancies.

References

1. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, et al. Integrated molecular analysis of adult T-cell leukemia/lymphoma. *Nat Genet.* 2015; 47(11):1304-15.
2. Kataoka K, Shiraishi Y, Takeda Y, Sakata S, Matsumoto M, Nagano S, et al. Aberrant PD-L1 expression through 3′ -UTR disruption in multiple cancers. *Nature.* 2016; 534(7607):402-6
3. Kataoka K, Iwanaga M, Yasunaga J, Nagata Y, Kitanaka A, Kameda T, et al. Prognostic relevance of integrated genetic profiling in adult T-cell leukemia/lymphoma. *Blood.* 2018;131(2):215-225.

Keisuke Kataoka, MD, PhD

2005-2007	Resident, Toranomon Hospital
2007-2008	Medical Staff, The University of Tokyo Hospital
2008-2012	Graduated from The University of Tokyo, Graduate School of Medicine
2009-2012	Japan Society for the Promotion of Science Research Fellow
2012-2013	Project Assistant Professor, The University of Tokyo
2013-2017	Project Assistant Professor, Kyoto University
2017-present	Chief, National Cancer Center Research Institute

A SINGLE CELL LENS INTO TUMOR HETEROGENEITY, METASTASIS AND TUMOR ECOSYSTEMS

Dana Pe'er
Program in Computational and Systems Biology
Sloan Kettering Institute
Memorial Sloan Kettering Cancer Center
417 E 68th Street, New York, NY 10027, USA
(peerster@gmail.com)

Single-cell RNA-sequencing is fast becoming a major technology that is revolutionizing the studies of tumor heterogeneity, tumor immune ecosystem and metastasis. Studies have shown that cancer cell dysregulate regeneration pathways, creating new phenotypes and malignant capabilities. We combine single cell RNA-seq with computational analysis to build a map tumor heterogeneity in AML and lung adenocarcinoma, connecting between cancer and healthy phenotypes.

Acute myeloid leukemia (AML) evolution is a multistep process in which cells evolve from hematopoietic stem and progenitor cells (HSPCs) that acquire genetic anomalies leading to to expansions of stem and progenitor cells with myeloid lineage bias. Many recurring mutations in AML involve known and putative epigenetic and transcriptional regulators, including TET2, IHD1/2 and NPM1. We profiled multiple AML samples with single-cell RNA sequencing creating an atlas that demonstrates demonstrating large scale inter and intra patient heterogeneity. While the patients show a large degree of inter-patient heterogeneity, we find a cluster shared by 11/12 samples that contains cells with a HSPC-like phenotype. We developed a manifold based algorithm to map trajectories of differentiating cells, to map and compare altered differentiation trajectories, starting from the shared HSPC-like cluster, across the patients. We found NPM1 to be a late driver in this disease with sub-clonal appearance, hence we could find trajectories with and without NPM1 within the same patient, to better understand its role. Interestingly, these mutations were detected at various frequencies in erythroid cells, suggesting that NPM1 mutations are acquired in cells with different lineage commitment in different patients. Most notably, the HSPC-like also contained a subpopulation of cells that have acquired NPM1 mutations

and are transcriptionally different from wild-type cells.

The understanding of cancer as a chronic, non-healing, wound invokes a central role for developmental pathways that regulate tissue regeneration and repair. How these pathways relate to tumor progression and metastasis remains poorly understood. Here we employ single-cell RNA sequencing to primary and metastatic patient lung adenocarcinomas. We show that primary human tumors demonstrate a striking degree of developmental plasticity, regenerating most normal lung epithelial lineages of the alveolar and bronchial airway and expressing key embryonic lineage-determining transcription factors, SOX2 and SOX9. Conversely, metastases exhibit a significant reduction in lineage diversity and are predominantly restricted to a Wnt-responsive, SOX9high epithelial progenitor. To dissect potential mechanisms underlying this lineage-specific escape, cancer cells from a mouse model of delayed lung cancer metastasis were transcriptionally profiled at different stages of disease progression. Lineage constraint was alleviated upon natural killer (NK)-cell depletion, demonstrating a dynamic interplay between developmental plasticity and immune-mediated pruning during lung cancer progression and the evolution of metastasis.

Dana Pe'er, PhD

1999-2003	PhD, Computer Science, Hebrew University of Jerusalem
2003-2006	Postdoctoral Fellow, Genetics, Harvard Medical School
2006-2016	Professor, Biological Sciences, Columbia University
2016-present	Chair, Computational and Systems Biology, Sloan Kettering Institute
2016-present	Director, Gerry Center for Metastasis and Tumor Ecosystems

WHOLE-ORGANISM CLONE-TRACING USING SINGLE-CELL SEQUENCING

Alexander van Oudenaarden
Hubrecht Institute
Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
(a.vanoudenaarden@hubrecht.eu)

Embryonic development is one of the most crucial periods in the life of a multicellular organism. A limited set of embryonic progenitors gives rise to all cells in the adult body. Determining which fate these progenitors acquire in adult tissue is a major challenge and requires the simultaneous measurement of clonal history and cell-type at single-cell resolution. Clonal history has traditionally been quantified by microscopically tracking cells during development, monitoring the heritable expression of genetically encoded fluorescent proteins and, most recently, by utilizing next generation sequencing technology exploiting somatic mutations, transposon tagging, viral barcoding, and CRISPR/Cas9 genome editing strategies. Single-cell transcriptomics on the other hand, provides a powerful technology platform for cell-type classification in an unbiased manner. However, integrating both measurements for many single cells has been a major hurdle. Here, we present ScarTrace, a single-cell sequencing strategy that allows us to simultaneously quantify information on clonal history and cell type for thousands of single cells obtained from different organs from adult zebrafish. Using this approach we show that all blood cells in the kidney marrow arise from a small set of multipotent embryonic progenitors that give rise to all blood cell types. In contrast, we find that cells in the eyes, brain, and caudal tail fin arise from many embryonic progenitors, which are more restricted and produce specific cell types in the adult tissue. Next we use ScarTrace to explore when embryonic cells commit to forming either left or right organs using the eyes and brain as a model system. Lastly we monitor regeneration of the caudal tail fin and identify a subpopulation of resident macrophages that have a clonal origin that is distinct from other blood cell types. We envision that ScarTrace will have major applications in other experimental model

systems to match embryonic clonal origin to adult cell-type to ultimately reconstruct how the adult body was built from a single cell.

References

1. Alemany A, Florescu M, Baron CS, Peterson-Maduro J, van Oudenaarden A. Whole-organism clone tracing using single-cell sequencing. *Nature*. Apr 5;556 (7699):108-112. doi: 10.1038/nature25969. Epub 2018 Mar 28.

Alexander van Oudenaarden, PhD

1998	Delft University of Technology, PhD Physics
1998-1999	Postdoctoral Research, Stanford
2000-2004	Assistant Professor, MIT, Cambridge
2004-2008	Associate Professor, MIT, Cambridge
2008-2014	Professor, MIT, Cambridge
2013-present	Professor, Utrecht University and UMC Utrecht
2012-present	Director Hubrecht Institute (KNAW), Utrecht

IN VIVO MECHANISMS OF RESISTANCE TO GENOTOXIC CHEMOTHERAPY

Michael T. Hemann
Koch Institute for Integrative Cancer Research at MIT
Cambridge, MA, USA
(hemann@mit.edu)

Many human cancers fail to respond to DNA damaging chemotherapy, and cancers that initially respond frequently acquire drug resistance and relapse. This process of tumor relapse is particularly confounding, as patients can be in remission for years following treatment prior to the reemergence of a cancer. Additionally, tumors that relapse have generally acquired resistance to the initial treatment. While conventional anti-cancer therapies have been in clinical use for decades, little is known about the mechanisms by which a tumor cell can survive treatment and persist in a patient for extended period of time. In fact, a major cause of cancer deaths is the inability to eradicate small sets of surviving tumor cells, termed "minimal residual disease" or "MRD". Work from our laboratory has found that chemotherapy paradoxically elicits a pro-survival response in certain anatomical cites [1]. Specifically, normal cells proximal to tumor cells secrete factors that counter the effects of the chemotherapy. This survival response likely serves to protect normal progenitor and stem cells to allow for tissue regeneration following damage, but it is coopted by tumor cells that find themselves in these specialized sites.

Microenvironment mediated chemoprotection

There is now considerable data supporting the idea that the response of cancer cells to chemotherapy fundamentally differs in cell culture versus the native tumor microenvironment. Indeed, we can quite easily eradicate cells from culture plates at drug doses that are unable to clear persistent disease in mice or humans. While the mechanisms long thought to underlie this differential response relate to drug target interaction or drug pharmacokinetics, there has been significant recent evidence that tumor

microenvironments can protect tumors from the consequences of drug action. These correlative data have prompted us to investigate the basic mechanisms of paracrine survival signaling that exist in specific microenvironments, during normal development and physiological stress responses, and whether these strategies are coopted by tumor cells to evade cancer therapy.

Our first work in the area focused on stress-induced signaling in the endothelial compartment [2-4]. In response to genotoxic damage, specific sets of endothelial cells are induced to engage secretory response that involves the release of pro-survival factors, including IL-6 and Timp-1. These factors act through the Stat/Jak pathway tumor cells lying immediately adjacent to the tumor vasculature, a process that leads to induction of the anti-apoptotic Bcl2 family member Bclxl. Importantly, the protective microenvironment that contains endothelial cells does not function by blocking DNA damage induced by chemotherapy, rather it counteracts the pro-apoptotic signals activated by DNA damage. This mechanism of chemoprotection has significant therapeutic implications, as blocking IL-6 activity with neutralizing antibodies eliminates the protective effect of endothelial cells and ablates the persistent MRD following chemotherapy.

We have gone on to further characterize the basic molecular mechanisms of this therapy related secretory response. Specifically, doxorubicin induces endothelial IL-6 mRNA stabilization and release via reactive oxygen species (ROS)-induced p38 signaling. This release is accompanied in endothelial cells by many of the hallmarks of cellular senescence, an irreversible cell cycle arrest that is accompanied by a robust secretory response. However, this therapy related secretory response is an atypical senescence-associated secretory response that is characterized by an acute, transient release of paracrine factors that does not produce a systemic inflammatory response. Thus, we identified a specialized chemotherapy-induced secretory response that protects subsets of persisting tumor cells. This data also show how protective secretory responses – the kinds of responses that protect repopulating cells in regenerative organs and persistent disease in cancer – can occur in the absence of systemic inflammatory responses (Figure 1). We believe that this regulation represents a fundamental mechanism of "shielding" paracrine pro-survival processes that underlie tissue regrowth in metazoans.

In addition to conventional chemotherapy, we have recently shown that select microenvironments can underlie resistance to antibody-based therapy. Using a humanized model of treatment-refractory B-cell leukemia, we have identified the bone marrow as a site of persistent MRD [5]. Specifically, we find that infiltration of leukemia cells into the bone marrow rewires the tumor microenvironment to inhibit engulfment of antibody-targeted tumor cells. Resistance to macrophage-mediated killing can be overcome by combination regimens involving therapeutic antibodies and genotoxic chemotherapy. These data

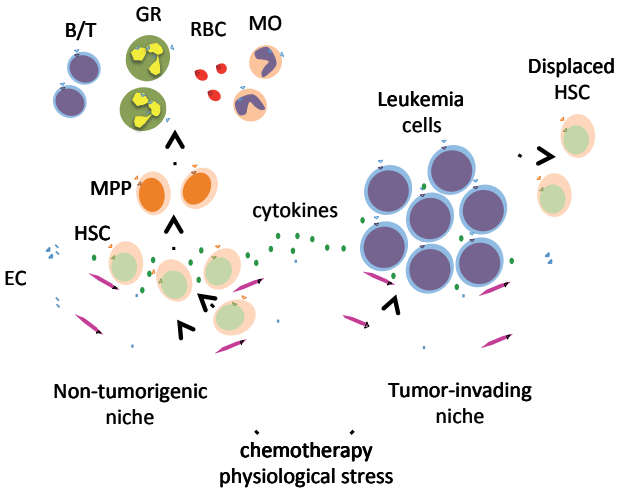


Figure 1 Development of protective cancer niches. Following the induction of exogenous stress, proximal cells – including endothelial cells – release factors that recruit stem and progenitor cells to sites of damage, promote the expansion of these cells and the subsequent regeneration of the hematopoietic system. Lymphoma and leukemia cells can displace normal stem from protective microenvironments, coopt pro-survival paracrine signaling and persist following chemotherapy. Abbreviations: EC are endothelial cells, HSC is hematopoietic stem cells, MPP is multi-potent progenitor, B/T are B and T cells, GR is granulocyte, RBC is red blood cell and MO is monocyte.

suggest that effective cancer therapy may not only involve targeting cancer cells but also may require inhibiting survival signals emanating from surrounding cells. By generating multiple models of MRD, we have not only identified mechanisms of tumor cell persistence, but also strategies to sensitize chemorefractory tumors to both front-line and targeted chemotherapy.

Mechanisms of action of genotoxic chemotherapy

The data described above reveal that genotoxic chemotherapies that are presumed to exert their effect primarily through the induction of DNA breaks may show profound mechanistic distinctions. In the case of doxorubicin, genotoxic damage promotes cytokine mRNA stabilization in release. In the case of cyclophosphamide, treated cells undergo systemic changes that promote innate immune activation. To more closely examine the precise mechanism of diverse genotoxic chemotherapies, we have developed an RNA interference (RNAi)-based approach to characterize chemotherapeutic function in mammalian cells [6-7]. Briefly, have used RNAi-induced phenotypic "signatures" to characterize drug action or the effects of distinct tumor microenvironments on drug action with unprecedented resolution. These robust RNAi-based "signatures" can be used to cluster unknown drugs into functional categories and define mechanisms of action for

uncharacterized cytotoxic agents. We have subsequently used this signature approach for applications that were previously not feasible using other drug characterization metrics.

A major focus of these efforts has been the study of platinum compounds that are commonly used as components of chemotherapeutic drug regimens [8]. Strikingly, we found that oxaliplatin and cisplatin, two agents that are thought to exert their effects via the same mechanism of action, are mechanistically quite distinct. While cisplatin promotes cell death via inducing DNA crosslinks and a DNA damage response, oxaliplatin kills cells by promoting inhibiting ribosome synthesis and translation. Thus, platinum drugs may not function interchangeably with their derivatives in cancer regimens. This phenomenon may explain a lack of efficacy for oxaliplatin in the treatment of malignancies conventionally treated by cisplatin, and the initially unanticipated value of oxaliplatin as a treatment for colorectal cancer. This work also highlights how changes in small molecule structure can alter the genetic determinants of chemotherapeutic response. For example, instead of cisplatin, oxaliplatin has begun to be used as a front-line treatment for pancreas cancer, and recent sequencing data show that greater than 20% of these tumors show defects in homologous recombination. Our data predict that these mutant tumors may be better treated by cisplatin than oxaliplatin. Thus, tumor mutations may represent important biomarkers underlying the choice between related platinum drugs (Figure 2). Additionally, whereas conventional chemotherapeutics are commonly thought of as

Platinum chemotherapy can be used in a targeted manner

	% H.R. mutant	% translation addicted
Ovarian Cancer	<50%	<5%
Colon Cancer	<5%	>90%
Breast Cancer	~30%	20-40%
Pancreatic Cancer	~25%	~5%
	Cisplatin	Oxaliplatin

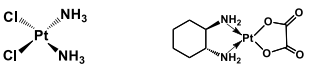


Figure 2 Platinum-containing compounds can function as tumor vulnerability directed targeted therapeutics. Cisplatin shows particular efficacy in tumors, like ovarian cancer, that show frequent defects in Homologous Repair (HR). Conversely, oxaliplatin promotes ribosomal dysfunction and shows efficacy in cancers, like colon cancer, with high translation requirements. Tumors like breast or pancreas cancer that may have HR or translation vulnerabilities should be treated in a targeted fashion with platinum compounds targeted towards specific pathway defects.

"generic" cytotoxic agents, our data highlight fundamental mechanistic distinctions that exist between highly similar molecules and support a personalized and genetically directed use of these compounds.

The convergence of resistance and mechanism

The challenges in identifying mechanisms of action and mechanisms of resistance for genotoxic chemotherapy are intimately linked (Figure 3). Indeed, in the context of targeted therapy, much of what we know about drug action is inferred from drug resistance mechanisms. For example, the identification of resistance mutations that alter putative sites of drug target interaction provide strong evidence that the mutant protein is indeed the relevant drug target. Moreover, understanding a drug's putative mechanism of action allows one to perform a focused search for mutations or compensatory gene expression changes in specific signaling pathways. These kinds of analyses have not been informative in the context of conventional genotoxic chemotherapy for two major reasons. First, as described above, resistance to genotoxic chemotherapy is generally not conferred by cell intrinsic, genetically encoded alterations. In fact, numerous large sequencing studies have failed to identify clear resistance mutations in relapsed tumors following genotoxic chemotherapy. Thus, resistance mutations cannot be used to define drug mechanism. Second, the obligate use of combination drug regimens has impeded the analysis of single component drugs within these regimens. These barriers have significantly blocked efforts to identify biomarkers that can predict the efficacy of genotoxic chemotherapy or strategies to rationally combine agents into multi-drug regimens.

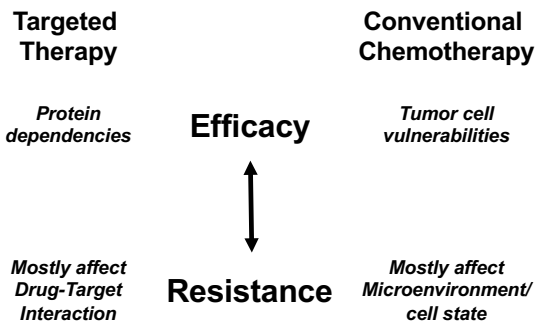


Figure 3 The efficacy and resistance landscapes of targeted versus conventional (frequently termed genotoxic) chemotherapy. Targeted therapies show efficacy in tumors expressing a specific target mutation, and resistance is generally driven by mutations that impact drug target interaction or eliminate the target. Conversely sensitivity to conventional chemotherapy is likely driven by general tumor cell vulnerabilities, ie defects in DNA damage repair. Resistance to these agents is conferred by protective tumor microenvironments or cell state changes that globally alter tumor cell sensitivities.

These challenges involved in understanding resistance mechanisms to conventional chemotherapy also, paradoxically, highlight the unique utility and tremendous potential of these compounds. The lack of clear resistance-driving mutations in relapsed tumors suggests that, unlike targeted therapy, individual mutations in tumor cells may be insufficient to promote chemoresistance. Thus, tumor cells have difficulty evading the action of genotoxic agents when present at sufficient doses. Additionally, the fact that a substantial component of drug resistance is microenvironmental provides a strategy for enhancing the effective doses of these agents. Specifically, by introducing agents that block paracrine pro-survival signaling in select microenvironments, one can effectively deliver potent dose of chemotherapy selectively to tumors in a manner that does not promote systemic toxicity.

References

1. Gilbert LA, Hemann MT. (2010) DNA damage-mediated induction of a chemoresistant niche. Cell, 143(3), 355-366.
2. Bent EH, Gilbert LA, Hemann MT. (2016). A senescence secretory switch mediated by PI3K/AKT/mTOR activation controls chemoprotective endothelial secretory responses. *Genes Dev*. 30(16):1811–21
3. Gilbert LA, Hemann MT. (2011). Chemotherapeutic resistance: surviving stressful situations. *Cancer Res*. 71(15):5062–66
4. Corre I, Paris F, Huot J. (2017). The p38 pathway, a major pleiotropic cascade that transduces stress and metastatic signals in endothelial cells. *Oncotarget* 8(33):55684–714
5. Pallasch C, Leskov I, Braun C, Vorholt D, Drake A, Soto-Feliciano Y, Bent E, Schwamb J, Iliopoulou B, Kutsch N, van Rooijen N, Frenzel L, Wendtner C, Heukamp L, Kreuzer K, Hallek M, Chen J, Hemann M. (2014) Sensitizing Protective Tumor Microenvironments to Antibody-Mediated Therapy. Cell. 156(3):590-602. PMCID: PMC3975171
6. Jiang H, Pritchard JR, Williams RT, Lauffenburger DA, Hemann MT. (2011) A mammalian functional-genetic approach to characterizing cancer therapeutics. Nat. Chem. Biol., 7(2), 92-100. PMCID: PMC3070540
7. Pritchard JR, Bruno PM, Gilbert LA, Capron KL, Lauffenburger DA, Hemann MT. (2013) Defining principles of combination drug mechanisms of action. Proc Natl Acad Sci. 110(2):E170-9.
8. Bruno PM, Liu Y, Park Y, Murai J, Koch CE, Eisen TJ, Pritchard JR, Pommier Y, Lippard SJ and Michael T Hemann. (2017) A subset of platinum-containing chemotherapeutic agents kills cells by inducing ribosome biogenesis stress. Nature Medicine. 23(4):461-471

Michael T. Hemann, PhD

1996-2001	PhD from Johns Hopkins School of Medicine
2001-2006	Postdoctoral Fellow at Cold Spring Harbor Laboratory
2006-2010	Assistant Professor of Biology at MIT
2013-present	Tenured Associate Professor of Biology at MIT

SIGNATURES OF MUTATIONAL PROCESSES IN CANCER

Michael Stratton
Wellcome Sanger Institute
Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
(mrs@sanger.ac.uk)

Somatic mutations are believed to occur in all cells and all cancers are thought to arise due to somatic driver mutations that occur in cancer genes. However, the repertoire of mutational processes that contributes to the somatic mutation burden in normal and cancer cells has been poorly understood. Furthermore, although it has been recognised for decades that exposures such as tobacco smoking or ultraviolet light contribute to the development of lung and skin cancers respectively, probably through increasing the somatic mutation burden in exposed cells, the causes of many other cancers, for example breast, glioma or pancreatic, and their relationships to mutation loads have remained obscure.

Early studies of mutations in the TP53 gene in the 1990s, following the discovery of frequent driver mutations in this gene in many human cancers, showed that patterns of somatic base substitution mutation differ according to the environmental or lifestyle exposures that contribute to a particular cancer type. Notably, the predominant pattern of base substitution mutations in lung cancers from smokers (C>A) is different from the predominant pattern of substitutions in ultraviolet light induced skin cancers (C>T) and both of these mutation types were consistent with the nature of the DNA damage known to be induced by the exposures. These studies therefore indicated that characterisation of patterns of somatic mutation, now called “mutational signatures”, present in cancers and in normal cell genomes could in principle provide insight into the repertoire of mutational processes that have been operative in them and potentially provide clues to endogenous and exogenous exposures contributing to cancer.

Although the TP53 gene provided transformative insights of principle, the small number of mutations from each cancer genome in TP53 (usually just a single mutation), provided limited ability to resolve different mutational signatures. The advent of next generation sequencing, in particular the large and comprehensive catalogues of somatic mutation from whole cancer genome sequences, changed the landscape in this regard with thousands to millions of mutations commonly detected in individual cancer genomes. In order to optimally use this information, however, two further major analytic developments required introduction. First, new subclassifications of base substitutions, doublet base substitutions, indels, rearrangements and copy number changes in order to provide sufficient resolution between closely related mutational signatures. For example the previous classification of base substitutions used for TP53 into six subtypes (C>A, C>G, C>T, T>A, T>C, T>G) has commonly now been elaborated into a classification based on 96 subtypes in which the base immediately 5′ and 3′ to the mutated base is also taken into account (generating 16 X 6 = 96 subtypes). Second, on the assumption that multiple mutational processes (and therefore mutational signatures) have contributed to the somatic mutations found in a single cell or cancer, methods were needed that could separate each mutational signature from the others present. These methods work on the assumption that the contribution of each signature to each cell/cancer can differ. A commonly used approach employs non-negative matrix factorisation (NMF) [1], although other methods have been developed. Further developments of these methods also allow estimation of the mutation burden of each signature in each sample.

The first application of these approaches to large numbers of cancer genomes of multiple types revealed 21 base substitution mutational signatures [2] and, in a subsequent update, 30 signatures [3]. Recently, as part of the Pan Cancer Analysis of Whole Cancer Genomes (PCAWG), which includes ~4,000 whole cancer genome and ~15,000 exome sequences, this list of signatures has been extended and now incorporates 49 base substitution signatures [4]. These include signatures of likely endogenous origin that are ubiquitous in normal and cancer cells and which accumulate in a clocklike, linear manner with age of the individual (single base substitution signatures (SBS) SBS1 and SBS5); signatures due to known exogenous exposures (including tobacco smoke (SBS4), ultraviolet light (SBS7a, SBS7b, SBS7c and SBS7d); signatures due to previous chemotherapeutic treatments (including temozolomide (SBS11), a cocktail of chemotherapeutics (SBS25), cisplatin (SBS31 and SBS35), azothiaprine (SBS32)); signatures of defective DNA repair (defective DNA mismatch repair (SBS6, SBS14, SBS15, SBS20, SBS21, SBS26, SBS40), defective base excision repair (SBS30, SBS36), defective homologous recombination based repair (SBS3)); signatures of overactive DNA editing by enzymes of the APOBEC family (SBS2, SBS13 and SBS84); signatures of defective proof reading by replicative polymerases

(SBS10a, SBS10b and SBS28); and many signatures of unknown origins (for example, SBS17 which is found predominantly in oesophageal adenocarcinoma (although also in other cancers)) and many of speculated origin (eg SBS18 which may be due to reactive oxygen species). In the course of this more highly powered analysis many signatures have been better separated from each other and some have split into multiple constituents. For example, SBS7, attributable to ultraviolet light exposure, has split into four components (SBS7a, SBS7b, SBS7c and SBS7d) suggesting that there are four distinct, but closely correlated, mutational processes operating. Many signatures show evidence of transcriptional strand bias indicating that some may have been initiated by bulky adducts binding to DNA, with transcription coupled nucleotide excision repair being recruited.

Although single base substitutions account for most somatic mutations there are other classes of mutation which require incorporation into the framework in order to render mutational signature analysis comprehensive. These include doublet base substitutions, indels, rearrangements and copy number changes. All these types of somatic mutation are present at much lower frequencies than single base substitutions and therefore offer less power for resolution of related signatures. In the recent PCAWG analyses doublet base substitutions and indels were analysed, either as separate classes or as part of composite signatures including all the different types (and rearrangements/structural variation were analysed by another group). Seventeen indel and 11 doublet base substitution signatures were found. The mutation loads attributable to some of these signatures closely correlate with the mutation loads of certain base substitution signatures suggesting that they are caused by the same or closely related mutational processes (eg DBS1, characterised predominantly by CC>TT mutations, correlates with SBS7a due to ultraviolet light exposure; and DBS2, characterised predominantly by CC>AA mutations, correlates with SBS4 due to tobacco smoke exposure).

The sets of somatic mutation present in a normal cell or cancer represent the aggregate accumulation of mutations over time since the fertilised egg of that individual was constituted. However, these mutations could, in principle, have been accumulated continuously during the complete lifetime of the individual or, alternatively, all-at-one-go, with any intermediate formulation between these two extremes being a possibility. It is often not possible to reconstruct well these temporal patterns of activity from mutation catalogues in cancer or normal cell genomes. Therefore, to explore temporal patterns of mutation accumulation of different signatures further, we have recently employed immortal cancer cell lines which are derived from naturally occurring human cancers [5].

Using previously generated exome somatic mutational catalogues from ~1000 human cancer cell lines we first characterised the mutational signatures present in each stock cell line. Many of these signatures will have been generated during the part of the life history of

Acknowledgements: many have contributed to this work over the years but I would particularly like to thank colleagues on the PCAWG mutational signatures working group, Ludmil Alexandrov, Jaegil Kim, Nick Haradvala, Gaddy Getz and Steve Rozen and Mia Petljak for leading the APOBEC episodic mutagenesis study.

References

1. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 2013 Jan 31;3(1):246-59. doi: 10.1016/j.celrep.2012.12.008. Epub 2013 Jan 10. PubMed PMID: 23318258; PubMed Central PMCID: PMC3588146.
2. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jäger N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, López-Otín C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdés-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature. 2013 Aug 22;500(7463):415-21. doi: 10.1038/nature12477. Epub 2013 Aug 14. Erratum in: Nature. 2013 Oct 10;502(7470):258. Imielinsk, Marcin [corrected to Imielinski, Marcin]. PubMed PMID: 23945592; PubMed Central PMCID: PMC3776390.
3. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, Stratton MR. Clock-like mutational processes in human somatic cells. Nat Genet. 2015 Dec;47(12):1402-7. doi: 10.1038/ng.3441. Epub 2015 Nov 9. PubMed PMID: 26551669; PubMed Central PMCID: PMC4783858.
4. Alexandrov LB, Kim J, Haradhvala NJ, Huang M, Ng AWT, Boot A, Covington KR, Gordenin DA, Bergstrom E, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V, Getz G, Rozen SG, Stratton MR, PCAWG Mutational Signatures Working Group, ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Network. The Repertoire of Mutational Signatures in Human Cancer doi: https://doi.org/10.1101/322859
5. Cell (in press 2019).

Michael Stratton, FMedSir FRS

1996	Leader in the Molecular Genetics of Cancer, Institute of Cancer Research
1997	Professor of Cancer Genetics and Chair of the newly constitute section of Cancer Genetics at the Institute of Cancer Research
1999	Fellow, Academy of Medical Sciences
2000	Member, the Sanger Institute's Board of Management Initiated the Cancer Genome Project and became the Head of this Programme
2007	Deputy Director of the Sanger Institute
2008	Fellow, the Royal Society
2009	Member of the European Molecular Biology Organization
2010	Director of the Sanger Institute
2012	Chief Executive Officer of the Wellcome Genome Campus
2013	Knighted by the Queen

IN VIVO AND *EX VIVO* ENGINEERING OF CANCER-ASSOCIATED CHROMOSOMAL VIA SOMATIC GENOME EDITING

Andrea Ventura
Cancer Biology and Genetics Program
Memorial Sloan Kettering Cancer Center
1275 York Avenue, Box 524, New York, NY 10065, USA
(venturaa@mskcc.org)

Introduction

The advent of high throughput sequencing methods has profoundly changed our understanding of cancer, revealing its genetic complexity, paving the way to novel therapeutic strategies, and ushering the era of precision medicine. This technological revolution, however, has also uncovered limitations in the experimental approaches available to model human cancers and characterize gene function *in vivo*. Although conventional gene targeting approaches based on homologous recombination in murine embryonic stem cells allow the generation of sophisticated genetically engineered mice, these technologies are not easily scalable, being costly and time-consuming. Furthermore, chromosomal rearrangements, a common class of cancer associated mutations that often result in the formation of therapeutically actionable gene fusions, have proven challenging to model in mice by homologous recombination.

To overcome this major limitation, we have developed a general CRISPR-Cas9-based strategy that allows the rapid generation of complex chromosomal rearrangements (deletions, inversions, duplications, and translocations) directly in somatic cells of live organisms [1, 2]. The approach is based on the concomitant introduction of Cas9 and two guide RNAs (gRNAs) designed to generate double stranded breaks (DSBs) at the desired chromosomal breakpoints [3]. The idea is that in response to the two DSBs, the endogenous DNA repair machinery will occasionally erroneously repair the two breaks generating the desired chromosomal rearrangement s [4].

Modeling EML4-ALK driven lung cancers

As initial proof of concept, we used this approach to model a chromosomal inversion of approximately 11 Mbp that results in the formation of a gene fusion between EML4 and the ALK kinase (EML4-ALK) and is detected in 3-5% of non-small cell lung cancers patients (NSCLCs). The EML4-ALK fusion protein is a known driver of NSCLCs and confers resistwith a recombinant adenovirus encoding Cas9 and two gRNAs designed to generate the most common EML4-ALK variant. By 8 weeks post-infection, all injected mice presented multiple bilateral lung lesions that upon examinations were diagnosed as lung adenoma. These lesions progressed over the next 8-10 weeks to lung adenocarcinomas. Importantly, each tumor examined displayed signs of constitutive activation of signaling pathways downstream of ALK (MAPK, AKT and STAT3), expressed the Eml4-Alk cDNA and was positive for the Eml4-Alk inversion (Figure 1).

To further confirm that the resulting tumors were in fact driven by the Eml4-Alk fusion protein, a cohort of animals was infected with the recombinant virus and imaged by microCT 12 weeks later to determine tumor load. Tumor bearing animals were then treated daily with Crizotinib or vehicle and reimaged 2 weeks later. All crizotinib treated animal displayed complete or near complete tumor regression, while all vehicle-treated mice showed marked disease progression (Figure 2).

These results demonstrate that *in vivo* somatic CRISPR-Cas9-mediated chromosomal engineering can be used to model lung adenocarcinomas driven by a recurrent chromosomal inversion.

Modeling chromosomal rearrangements in human brain cancers

In principle CRISPR-based somatic chromosomal engineering could be used to rapidly test the functional relevance of newly identified gene fusions and generate relevant preclinical models to test new therapeutic approaches. For this new set of studies, we chose to focus on human brain cancers because they represent a major clinical challenge with limited therapeutic options and because a large number of rare gene fusions of unclear functional significance have been reported in these tumors. Among the chromosomal rearrangements sporadically observed in human gliomas we first attempted to model a ~600 kb deletion resulting in the fusion between BCAN and NTRK1. This gene fusion has been reported in only a few patients but its pathogenic role was unclear. To model this rearrangement in mice, we chose two complementary approaches. First, we isolated adult neuronal stem cells (aNSCs) from the brain of p53-null mice (the p53 pathway is nearly always inactivated in human glioblastomas) and generated the BCAN-NTRK1 rearrangement by transfecting them with plasmids encoding the two gRNAs and Cas9. We then orthotopically injected the transfected cells with into the sub-ventricular zone (SVZ) of

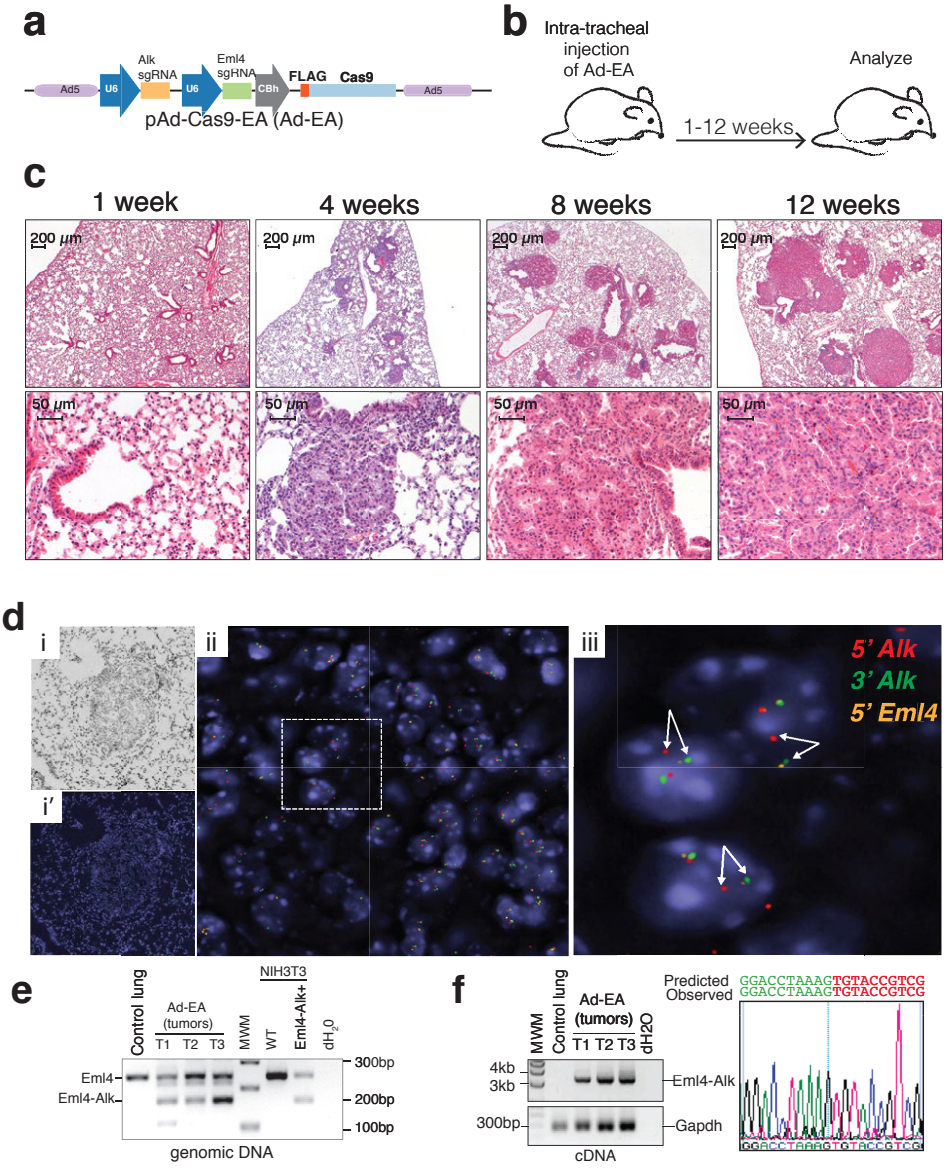


Figure 1 (a) Schematic of the Adenoviral vector used to deliver Cas9 and sgRNAs targeting Eml4 and Alk to the lung of adult mice. (b) Schematic of the experiment. Mice are injected intratracheally with 1.5x10^8 PFU of recombinant adenoviruses and euthanized at increasing intervals between 1 and 12 weeks post-infeciton. (c) Hematoxylin-eo-sin staining of lungs from mice at the indicated times after intratracheal instillation of Ad-EA. (d) Break-apart interphase FISH showing the presence of the Eml4-Alk inversion in a tumor from an Ad-EA-infected mouse (8 weeks post-infection). (i) Bright Field. (i', ii, iii increasing magnifications of merged fluorescent channels). (e) Detection of the wild type Eml4 locus and Eml4-Alk inversion in micro-dissected tumors from Ad-EA-infected mice using a three-primer PCR strategy. (f) RT-PCR detection (left) of the full-length Eml4-Alk cDNA (~3.2Kb) in the tumors shown in (e). The full-length PCR products were sequenced on both strands. A chromatogram of the Eml4-Alk junction is shown (right).

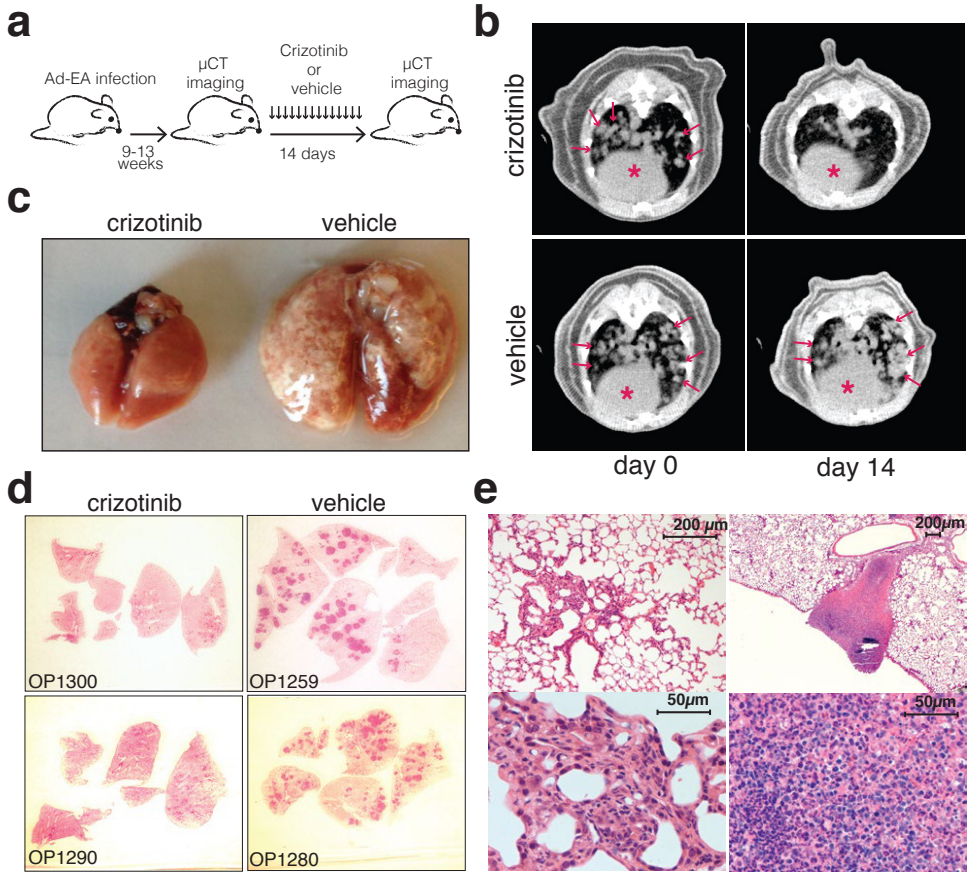


Figure 2 (a) Schematic of the experiment. (b) Representative μCT of the lungs of mice treated with crizotinib or vehicle at day 0 and after 2 weeks of treatement. Lung tumors are indicated by arrows. Red asterisks mark the hearts. (c) Macroscopic appearance of the lungs after 2 weeks of treatment. (d) Low magnification of lung sections from two crizotinib- and 2 vehicle-treated mice (hematoxylin eosin). (e) Higher magnification of representative hematoxylin-eosin stained lung sections from crizotinib-treated mice showing residual atrophic foci of tumor cells (left) or necrotic-inflammatory debris (right).

adult immune-deficient mice to determine whether cells harboring the BCAN-NTRK1 rearrangement would be capable of forming brain tumors *in vivo*. 3-4 months later, the majority of animals injected with the transfected aNSCs, but not control animals injected with control p53-null aNSC, developed neurological signs consistent with the presence of an intracranial mass. At necropsy, each mouse showed infiltrating masses that upon histologically resulted to be high grade gliomas. Importantly, genomic analysis revealed the presence, in each lesion examined, of the BCAN-NTRK1 rearrangements (Figure 3).

In parallel to the *ex-vivo* somatic chromosomal engineering approach just described, we also generated autochthonous BCAN-NTRK1 driven gliomas by *in vivo* chromosome engineering. To do so, we injected $p53^{fl/fl}$ mice intracranially with recombinant adenoviruses expressing the Cre recombinase (AdCre, to inactivate p53) and with adenoviruses expressing Cas9 and the two gRNAs designed to engineer the BCAN-NTRK1 deletion. This approach further emphasizes the flexibility of CRISPR-based chromosomal engineering, as it can be easily combined with the large number of already available genetically engineered mouse models (GEMMs).

As expected, within 4 months, most of the injected animals had developed high grade gliomas that were histologically indistinguishable from the ones generated using the ex vivo approach, invariably harbored the BCAN-NTRK1 rearrangements, and showed recombination of both floxed p53 alleles (Figure 4).

Finally, we tested whether BCAN-NTRK1 gliomas generated by chromosomal engineering would be sensitive to a novel experimental TRKA inhibitor (Entrectinib, Ignyta). Cell based experiments indicated that BCAN-NTRK1-positive murine gliomas are markedly sensitive to this drug in vitro at nanomolar concentrations. A cohort of mice harboring the BCAN-NTRK1 driven gliomas was therefore treated daily with Entrectinib or vehicle. During the treatment window, all vehicle control animals succumbed to the disease, while all Entrectinib treated mice were still alive at the end of the treatment, showing wither disease regression or stable disease (Figure 5).

In summary, we have developed a simple, flexible, and rapid strategy to generate a wide range of chromosomal rearrangements in cells and in mice. We are currently employing this strategy to model and investigate additional cancer types, including pediatric brain tumors driven by BRAF fusions. We expect that this approach will accelerate the development of better therapeutic strategies, will provide new insights into the pathogenesis of a wide range of human cancers driven by chromosomal rearrangements, and will facilitate the functional characterization of gene rearrangements involving the non-coding fraction of the human genome.

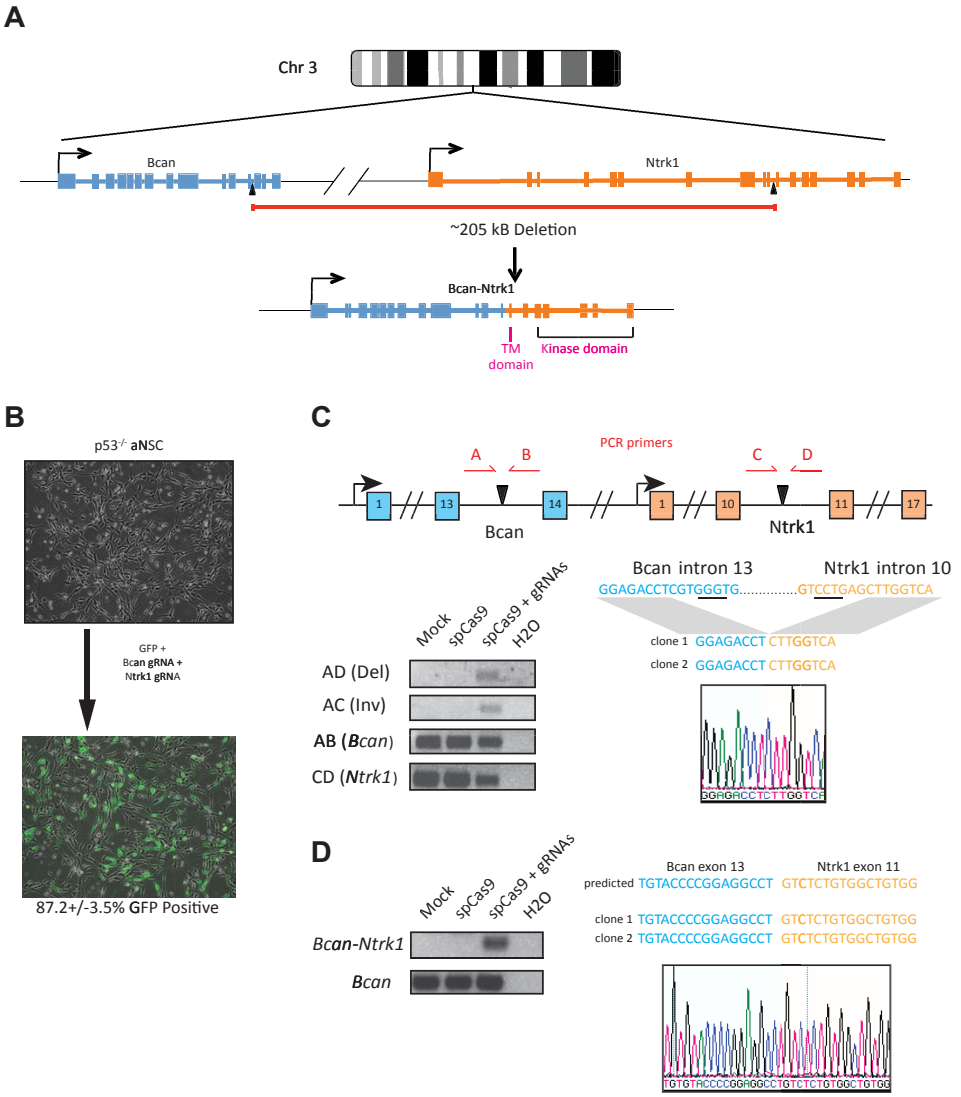


Figure 3 Induction of Bcan-Ntrk1 deletion in murine cells using CRISPR-Cas9
(A) Schematic representation of the deletion on mouse chromosome 3 generating in the Bcan-Ntrk1 fusion allele. gRNA cut sites are indicated by black arrowheads. (B) Representative GFP fluorescence images of p53-/- mouse adult neural stem cells (aNSC) before and 24h after nucleofection with indicated expression plasmids. FACS analysis was used to determine the percent of GFP-positive cells (mean +/-SEM, n=3). (C) Top, schematic of the targeted intronic region in Bcan and Ntrk1. Arrowheads indicate clieavage sites by the gRNAs. Orange arrows indicate PCR primers (A-D) designed to detect the rearrangement. PCRs were performed on genomic DNA extracted from aNSCs nucleofected with the indicated pX330 constructs (bottom left). The PCR bands were sub-cloned and the sequences of two independent clones and a representative chromatogram are shown in the lower right panel. (D) RT-PCR (left panel) on total RNA extracted from aNSCs nucleofected with the indicated constructs using using primer designed to detected the wild type Bcan cDNAs or the Bcan-Ntrk1 fusion transcript. The band corresponding to the fusion transcript was sub-cloned and sequenced (right panel).

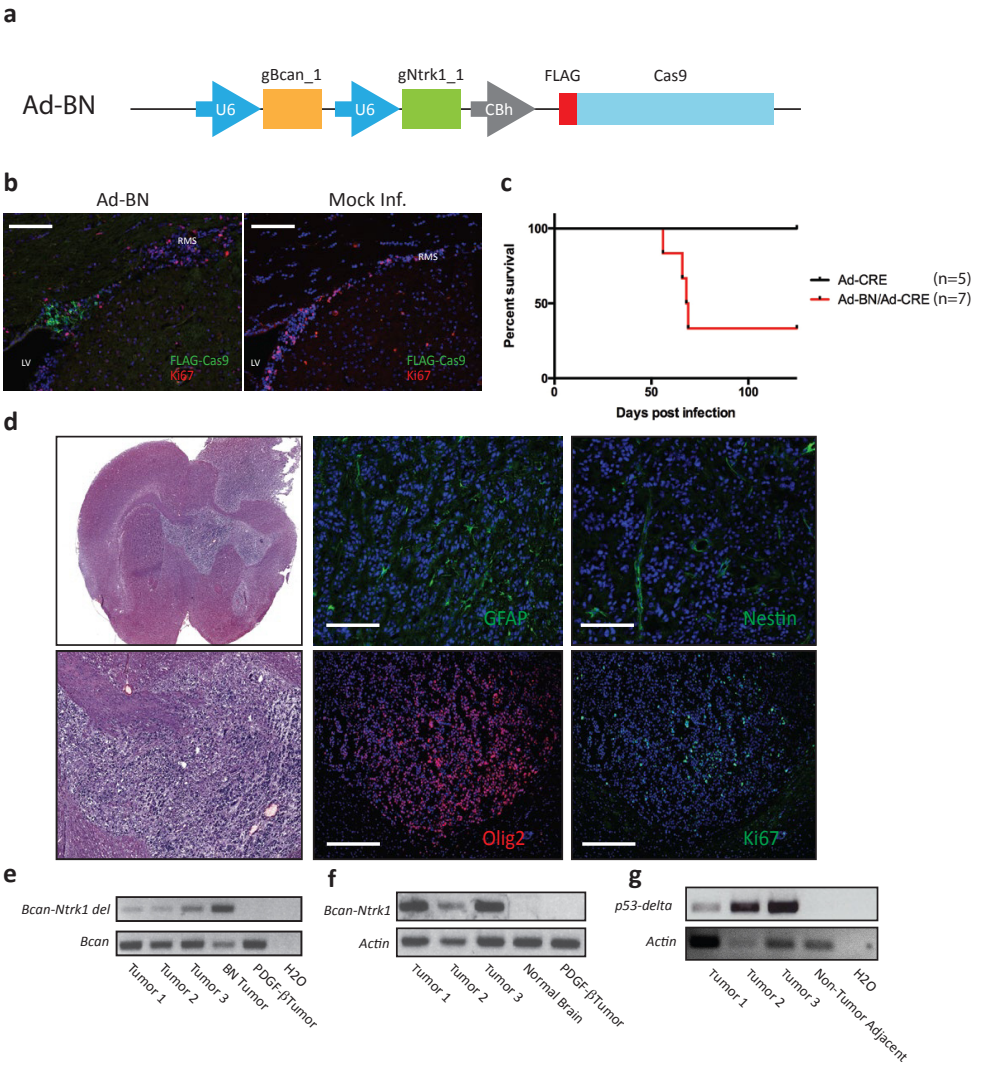


Figure 4 Generation of Bcan-Ntrk1-positive high-grade gliomas by in vivo somatic chromosomal engineering

a. Schematic of the recombinant adenoviral vector (Ad-BN) expressing the BN1 gRNA pair and FLAG-Cas9. b. Immunostaining with anti FLAG and anti-Ki67 antibodies of brain sections from adult wild type mice 48 hours after stereotactical intracranial injection with Ad-BN (scale bar: 0.1mm). c. Survival curves of p53fl/fl mice injected intracranially with a 1:1 mixture of Ad-BN and Ad-CRE (n=7) or with Ad-CRE alone (n=4). d. H&E staining and immunostaining with the indicated antibodies of a representative tumor observed in mice infected with the Ad-BN virus (scale bar: 0.5mm). e. PCR analysis of genomic DNA purified from three representative gliomas observed in the Ad-BN/Ad-Cre-infected mice (Tumor 1-3), a Bcan-Ntrk1-positive glioma generated by orthotopic implantation of aNSC harboring the Bcan-Ntrk1 rearrangement (BN Tumor) and a PDGF-beta-driven mouse glioma were included as controls. Primers to detect the Bcan-Ntrk1 deletion or the wilde type Bcan allele were used. f. RT-PCR using primers designed to detect the Bcan-Ntrk1 fusion transcript was performed on total RNAs extracted from the indicated tumors. g. Detection of the recombined p53 allele by genomic PCR on the indicated tumors and control tissues.

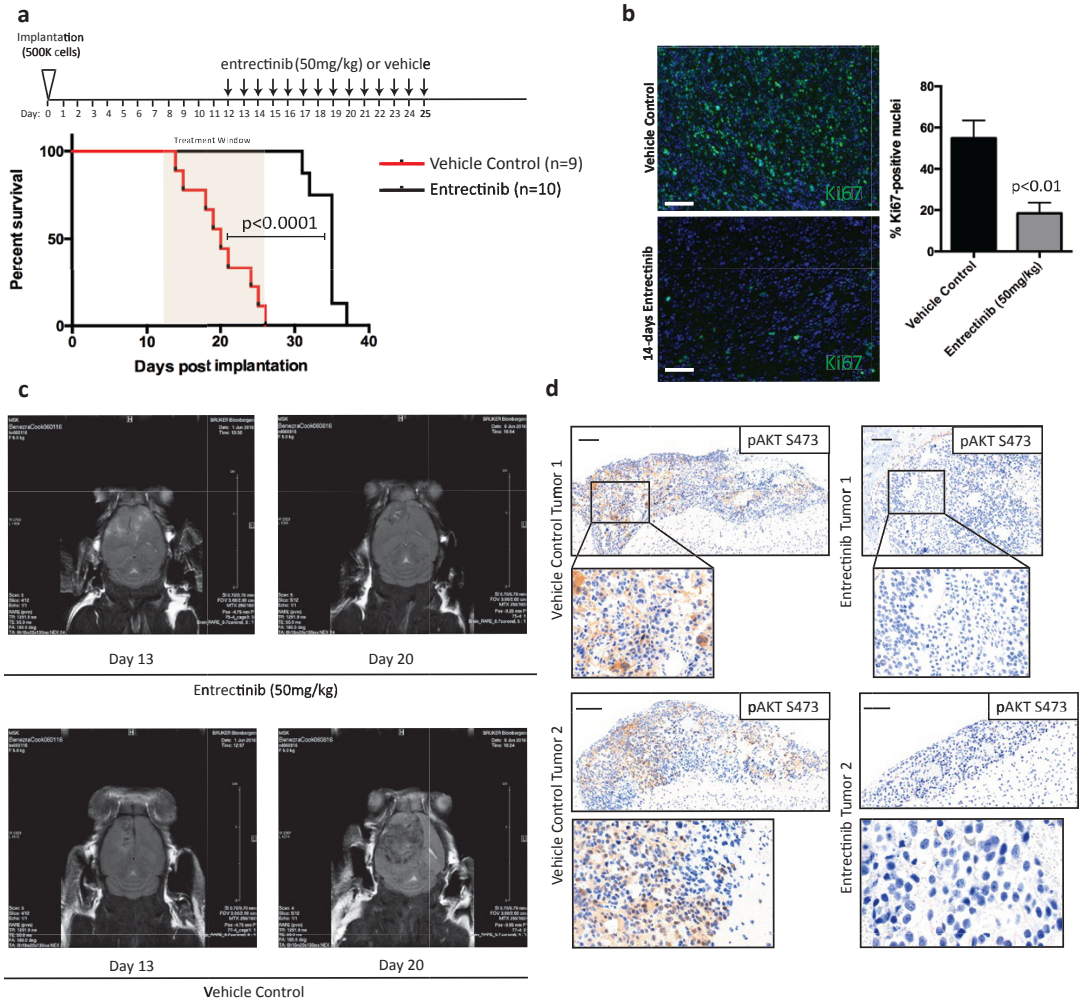


Figure 5 Entrectinib inhibits Bcan-Ntrk1-positive gliomas growth in vivo

a. Upper panel: schematic of the experiment. Mice were injected intracranially with cells derived from a glioma harboring the Bcan-Ntrk1 rearrangement at day 0. Daily treatment with entrectinib or vehicle was initiated at day 12 and continued for 14 days. Lower panel: Kaplan-Meier curve of mice receiving entrectinib or vehicle (p-value=log-rank test). b. Immunofluorescence staining for Ki67 (left panel) and quantification of Ki67+ cells (right panel) of treated and control tumors harvested on day 26 post-implantation. Scale bar=0.1mm. p-value=two tailed t-test Error bars represent mean +/-S.E.M. c. representative MRI analysis performed on entrectinib- or vehicle-treated mice on day 15 and 22 post-implantation. d. Histological analysis of animals receiving either entrectinib or vehicle for two days (day 14 post tumor cell implantation). Immunostaining for pAkt1 (S473) of tumors treated with entrectinib of vehicle for two days. Scale bar=0.1mm.

References

1. Maddalo, D., E. Manchado, C.P. Concepcion, C. Bonetti, J.A. Vidigal, Y.C. Han, P. Ogrodowski, A. Crippa, N. Rekhtman, E. de Stanchina, S.W. Lowe, and A. Ventura, *In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system.* Nature, 2014. 516(7531): p. 423-7.
2. Cook, P.J., R. Thomas, R. Kannan, E.S. de Leon, A. Drilon, M.K. Rosenblum, M. Scaltriti, R. Benezra, and A. Ventura, *Somatic chromosomal engineering identifies BCAN-NTRK1 as a potent glioma driver and therapeutic target.* Nat Commun, 2017. 8: p. 15987.
3. Choi, P.S. and M. Meyerson, *Targeted genomic rearrangements using CRISPR/Cas technology.* Nat Commun, 2014. 5: p. 3728.
4. Maddalo, D. and A. Ventura, *Somatic Engineering of Oncogenic Chromosomal Rearrangements: A Perspective.* Cancer Res, 2016. 76(17): p. 4918-23.

Andrea Ventura, MD, PhD

1997	MD, Catholic University Medical School, Rome
2003	PhD in Molecular and Cellular Biology, European Institute of Oncology, Milan
2003-2008	Postdoctoral Fellow, MIT in the Laboratory of Dr. Tyler Jacks
2008-2017	Assistant Member, Dept. of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center (MSKCC)
2014-present	Associate Member, Dept. of Cancer Biology & Genetics

SYSTEMATIC MAPPING OF GENETIC VULNERABILITIES IN CANCER

Prashant Mali
Department of Bioengineering, University of California San Diego
9500 Gilman Drive 0412, La Jolla, CA 92093, USA
(pmali@ucsd.edu)

A vast number of mutations contribute to cancer, but the observed non-random combinations of these leading to transformation highlight the importance of hallmark pathways and networks in cancer progression. While many pathways have been implicated in cancer, tumor heterogeneity stemming from different mutagens, tissue of origin, degree of progression, etc. leads to each case exhibiting a unique subset of altered pathways. Taken together, this diversity among cancer types and their origins, compounded by the huge genetic and epigenetic heterogeneity of tumors has complicated the development of targeted cancer treatments. The central premise of our studies is that cancer mutations converge into genetic interaction networks, and these networks bring together mutations of all varieties, including genes with low frequency of oncogenic mutations, and tumor suppressor profiles [1,2]. We thus hypothesize that by systematically mapping these networks in both cell lines and in clinically relevant *ex vivo* and *in vivo* models, new molecular targets for cancer therapy can be identified. Correspondingly, we develop and use state-of-the-art high-throughput epistasis mapping via CRISPR-Cas [3,4] and coupled single-cell analyses to enable systematic interrogation of the functions of individual genes and combinations while also assaying the impact of tumor heterogeneity. Coupled with *in vitro* and *in vivo* drug and targeted genetic validations in preclinical models, we anticipate this integrated experimental and computational framework will result in new insights into the underlying tumor biology, as well as unraveling of clinically actionable genetic vulnerabilities to advance the practice of precision oncology.

To enable systematic mapping of genetic interaction networks, we have developed a CRISPR-Cas9 based screening methodology for targeting single and pairs of genes in high-

throughput (*Nature Methods, 2017; Molecular Cell 2018*). Below I reproduce and summarize some of the textual descriptions, experiments, and results from these studies.

In the CRISPR-Cas9 system, a guide-RNA (gRNA), in complex with the Cas9 protein, targets genomic sequences homologous to the gRNA. Targeting new genomic elements entails modifying the gRNA sequence, thus enabling many targeted genome editing and regulation capabilities. Notably, Cas9 also enables facile multiplex targeting *via* delivery of multiple gRNAs per cell. Here, we combined multiplex targeting with array-based oligonucleotide synthesis to create dual-gRNA libraries covering up to 10^5 defined gene pairs (Figure 1a). In these libraries, each construct bears two gRNAs, with each gRNA designed to target either a gene or a scrambled non-targeting sequence absent from the genome. Thus, all combinations of gene-gene (double gene perturbation) and gene-scramble (single gene perturbation) are exhaustively assayed for effects on cell growth. Notably, both spacers for a dual-gRNA construct are directly specified during oligonucleotide synthesis thereby enabling the library constituents to be exactly defined to facilitate custom gRNA-pairing. By enabling determination and comparison of single-gene and dual-gene perturbation effects in the same assay, this approach allows for the systematic quantification of genetic interactions in human cells.

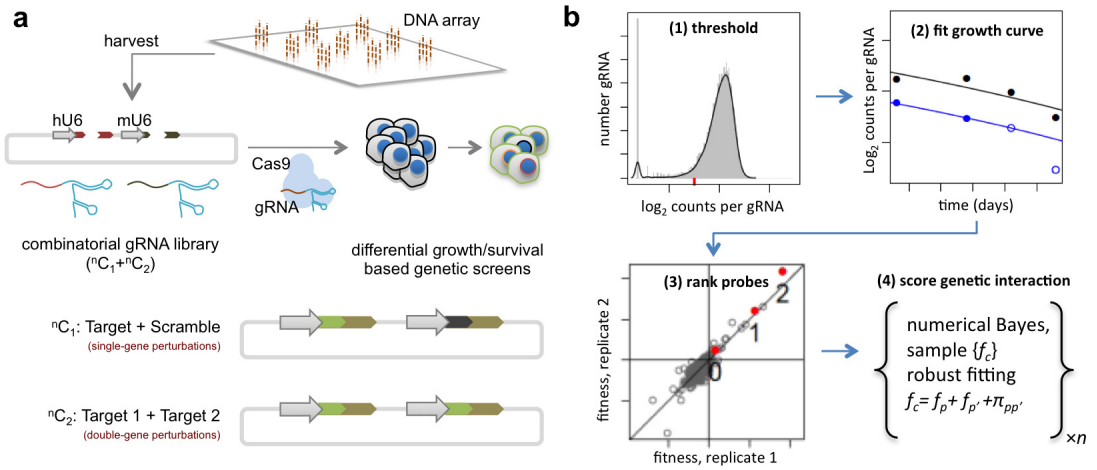


Figure 1 (a) Combinatorial genetic screening experimental approach to de novo map functional genetic interactions. (b) Overall analysis workflow: CRISPR screens are runs as two independent replicate experiments, with cells harvested at 4 time points and gRNA frequencies determined by next-generation sequencing (NGS). All gRNAs below a threshold (red dash) are excluded from further analysis. A fit of log relative abundance vs. time represents fitness, and probes are subsequently ranked by absolute fitness and weighted. Finally, a numerical Bayesian method is used to test for presence of a genetic interaction. (*Figure reproduced from Nature Methods 2017*).

We conduct genetic interaction screens by transducing the dual-gRNA lentiviral library into a population of cells stably expressing Cas9, maintaining these cells in exponential growth over the course of four weeks, and then sampling the relative changes in gRNAs at days 3, 14, 21 and 28 post-transduction. To robustly quantify gene fitness and genetic interactions, we have developed a computational analysis framework that integrates all samples across the multiple days of the experiment. This method: one, detects and removes gRNA constructs with insufficient read coverage; two, fits growth curves to the measured $\log_2$ abundances of each construct over time, the slopes of which reflect fitness; three, integrates data from the multiple gRNA constructs to derive a robust fitness value for disruption of each gene, f_g, and gene pair, $f_{g,g'}$; and four, computes the genetic interaction score $\pi_{gg'}$, as the difference of the observed from the expected fitness of the double gene knockout (Figure 1b). Significant departures from expected ($\pi < 3\sigma$ or $\pi > 3\sigma$) are called as negative or positive genetic interactions, respectively. A negative interaction indicates slower-than-expected growth, suggesting synthetic sickness or lethality, while a positive interaction indicates faster-than-expected growth, suggesting epistasis.

As an exemplar, using this method, we evaluated all pairwise gene knockout combinations among a panel of 73 genes divided between tumor-suppressor genes (TSG) and cancer-relevant drug targets (DT), a subset of which were also verified oncogenes (Figure 2). Experiments were performed in three somatic cell lines: HeLa, a cervical cancer cell line driven by Human Papilloma Virus (HPV); and A549, a lung cancer cell line driven by KRAS G12S mutation. With nine gRNA pairs per combination, the library comprised 23,652 double gene knockout constructs and 657 single gene constructs; testing two replicates in each cell line yielded a total of 94,608 unique tests of interaction. Measurements of gene fitness (f_g) were well correlated between biological replicates in the same cell line, as were the π scores for significant genetic interactions. Moreover, we observed a significant correlation between the number of genetic interactions identified for a gene and its single gene fitness suggesting that network 'hubs' may have increased functional importance to cells relative to genes with fewer interactions; such a relationship has been previously observed in model organisms but not before in humans. Comparing the two cell lines, we found lower but significant correlation of the single gene fitness scores between HeLa and A549. Differences in these fitness scores recapitulated known biological differences between cell lines, including the large positive growth effect of *TP53* knockout in A549 but not HeLa, in which TP53 is already inactivated by HPV. Knockout of DTs was generally more deleterious to growth than knockout of TSGs, suggesting that cancer cell lines are particularly dependent on the signaling pathways for which targeted drugs have been successfully developed thus far. Furthermore, we found that the genetic interactions identified from these data were remarkably different between cell lines. ~20%

of the interactions we observed had been previously identified, including the therapeutically relevant interactions EGFR-BRAF and PTEN-MTOR. Finally, the differences in genetic interaction across cell lines seen by systematic CRISPR could be largely reproduced as drug-drug interactions in small-scale assays.

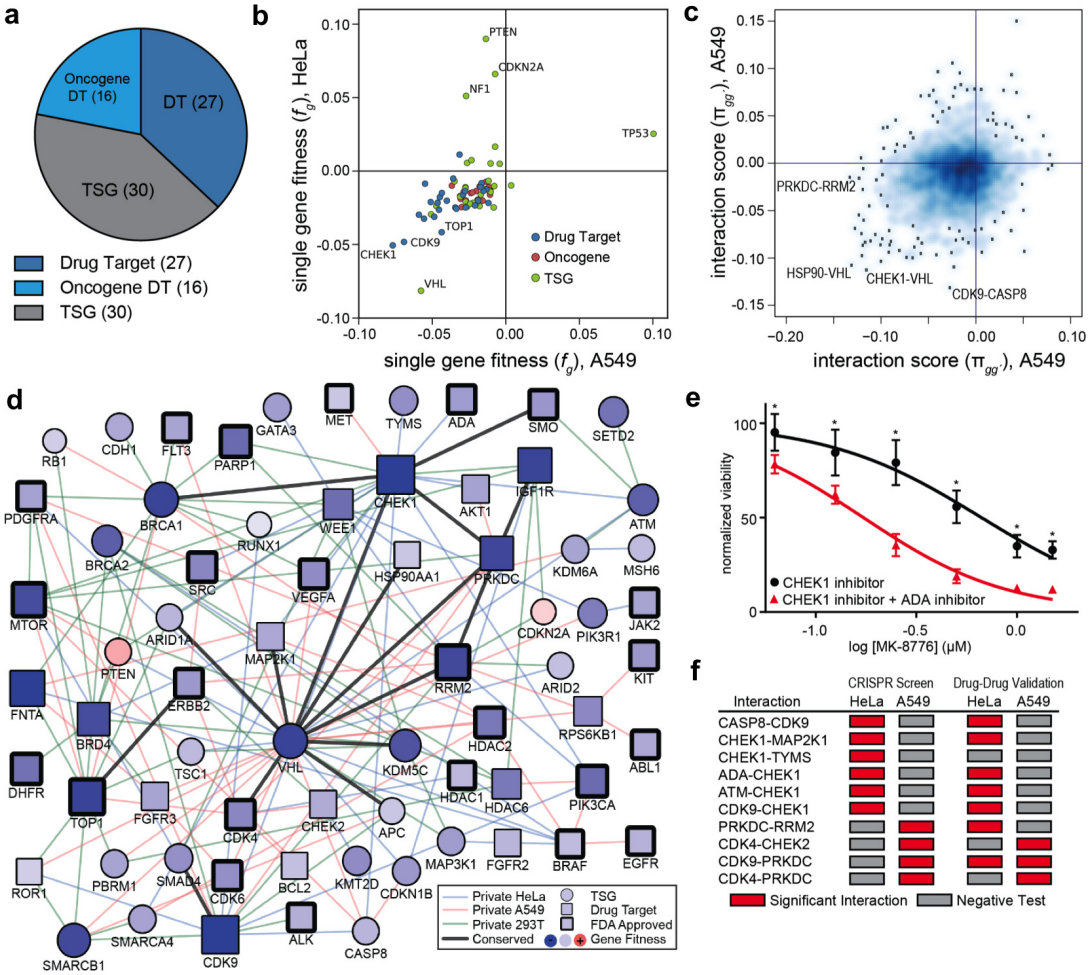


Figure 2 (a) A panel of 73 oncogenes, tumor suppressors, and cancer relevant drug targets was screened. (b) Correlation of the single-gene fitness, and (c) pi-scores between HeLa and A549 is depicted. (d) Combined synthetic-lethal network for three cell lines. Circles indicate TSG, squares DTs. Node colors indicate single gene knockout fitness effect, red: positive fitness effect, blue: negative fitness effect. Black edge around node indicates that the protein product of the gene is the target of an FDA approved drug. Color of edge indicates the cell line in which the interaction was identified, blue: HeLa, red: A549, green: 293T. Black edges were identified in multiple cell lines. (e) Example dose-response curve: adding the ADA inhibitor Fludarabine shifts the dose-response curve of CHEK1 inhibitor MK-8776 to the left. (f) Summary of the eight synthetic lethal interactions validated in this pilot study. (*Figure reproduced from Nature Methods 2017*).

As a second exemplar, we have also developed an integrated platform (Figure 3) combining genetic, transcriptomic, and flux measurements, to improve elucidation of metabolic network alterations, and guide precision targeting of metabolic vulnerabilities based on tumor genetics. Specifically, we conducted a screen of isozymes involved in carbohydrate metabolism in different non-small cell lung cancer cells, we characterized how loss or mutation of the *KEAP1* tumor suppressor influences the susceptibility of cells to knockout of oxidative pentose phosphate pathway enzymes.

In summary, we have successfully established a combinatorial CRISPR-Cas9 genetic interaction mapping technology that successfully identifies genetic interactions in human cancer cells and shows the great importance of cellular context on the architecture of the genetic interaction network.

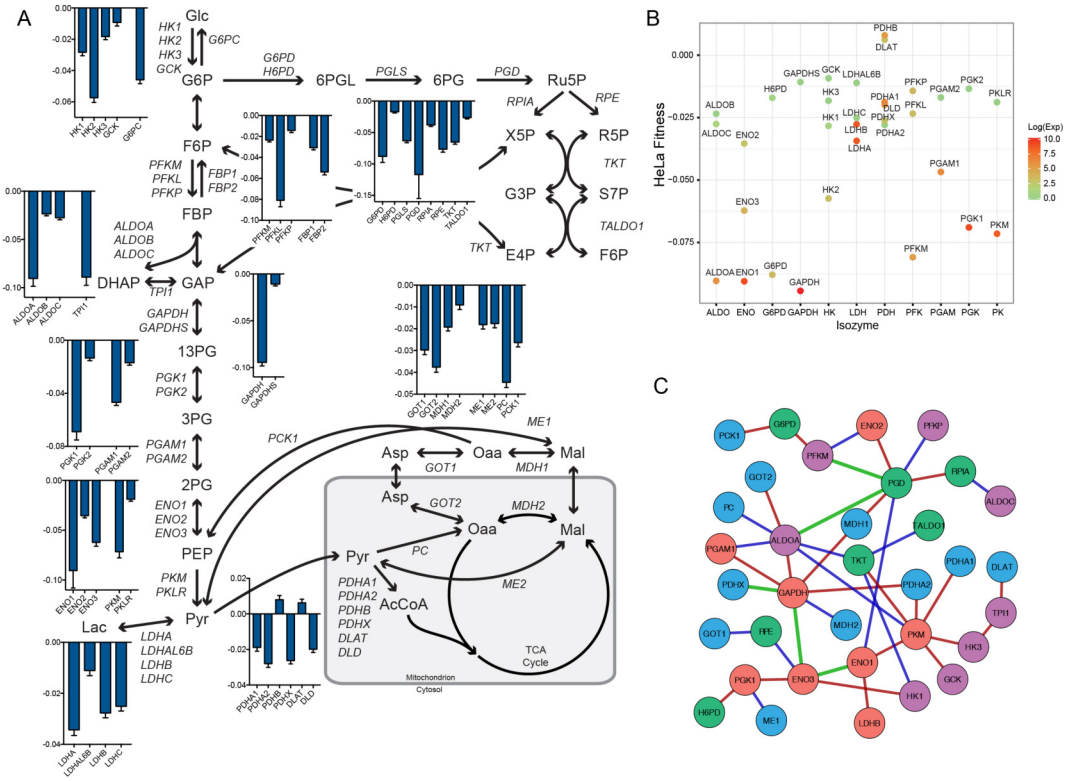


Figure 3 Combinatorial CRISPR screens reveal metabolic network dependencies. (a) SKO fitness scores for HeLa cells, plotted as f_g (day^{-1}), with a more negative score representing a loss in fitness with SKO. Plotted as mean ± SD. (b) Comparison of fitness scores versus gene expression across multi-isoform families in HeLa. (c) Genetic interaction map computed based on the combinatorial knockouts. (*Figure reproduced from Molecular Cell 2018*).

References

1. Dongxin Zhao, Mehmet G. Badur, Jens Luebeck, Jose H. Magana, Amanda Birmingham, Roman Sasik, Christopher S. Ahn, Trey Ideker, Christian Metallo, Prashant Mali, "Combinatorial CRISPR-Cas9 metabolic screens reveal critical redox control points dependent on the KEAP1-NRF2 regulatory axis," *Molecular Cell* 69, pp. 699-708 (2018).
2. John Paul Shen, Dongxin Zhao, Roman Sasik, Jens Luebeck, Amanda Birmingham, Ana Bojorquez-Gomez, Katherine Licon, Kristin Klepper, Daniel Pekin, Alex Beckett, Kyle Sanchez, Alex Thomas, Chih-Chung Kuo, Dan Du, Assen Roguev, Nathan E. Lewis, Aaron N. Chang, Jason F. Kreisberg, Nevan Krogan, Lei Qi, Trey Ideker, Prashant Mali, "Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions," *Nature Methods* 14, pp. 573-576 (2017).
3. Prashant Mali, John Aach, P. Benjamin Stranges, Kevin M. Esvelt, Mark Moosburner, Sriram Kosuri, Luhan Yang, George M. Church, "CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," *Nature Biotechnology* 31, pp. 833-838 (2013).
4. Prashant Mali, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, George M. Church, "RNA-guided human genome engineering via Cas9," *Science* 339, pp. 823-826 (2013).

Prashant Mali, PhD

2004-2011	PhD, Biomedical Engineering, Johns Hopkins University, School of Medicine
2011-2014	Postdoctoral Fellow Department of Genetics, Harvard Medical School
2014-present	Assistant Professor Bioengineering Department, University of California, San Diego

MATCHED TUMOR-NORMAL MUTATION PROFILING TO IDENTIFY THERAPEUTIC BIOMARKERS AND GUIDE CLINICAL CARE

Michael F. Berger
Department of Pathology and Center for Molecular Oncology
Memorial Sloan Kettering Cancer Center
1275 York Avenue, New York, NY 10065, USA
(bergerm1@mskcc.org)

Tumor molecular profiling is a fundamental component of precision oncology, enabling the identification of oncogenomic mutations that can be targeted therapeutically. To accelerate enrollment to clinical trials of molecularly targeted agents and guide treatment selection, we have established an enterprise-wide, prospective clinical sequencing program at Memorial Sloan Kettering Cancer Center using a custom, paired tumor-normal sequencing assay (MSK-IMPACT). MSK-IMPACT is a deep-coverage, targeted sequencing panel designed to identify somatic mutations, copy number alterations, and recurrent structural rearrangements in 468 cancer-associated genes with high sensitivity, and is authorized for clinical use by the US Food and Drug Administration [1]. As of November 2018, we have prospectively profiled the tumors of more than 28,000 active patients at our institution to inform their clinical care. More than 40% of patients harbor at least one somatic alteration that can be targeted by an approved or investigational therapy.

Matched tumor-normal sequencing has allowed us to characterize not only individual somatic alterations but also complex mutational signatures, germline cancer susceptibility alleles, and clonal hematopoiesis, all of which may have important clinical implications (Figure 1). We have detected microsatellite instability, a signature that predicts response to immune checkpoint inhibitors, in many unexpected cancer types, often associated with underlying Lynch Syndrome-associated germline mutations [2]. Other pathogenic germline mutations are also observed more frequently than previously thought, often undergoing tumor-specific zygosity alterations, and in genes and tumor types not encompassed by current clinical guidelines [3]. Mutations associated with clonal hematopoiesis in patients with solid tumors are associated with smoking history, prior therapy, shorter survival, and

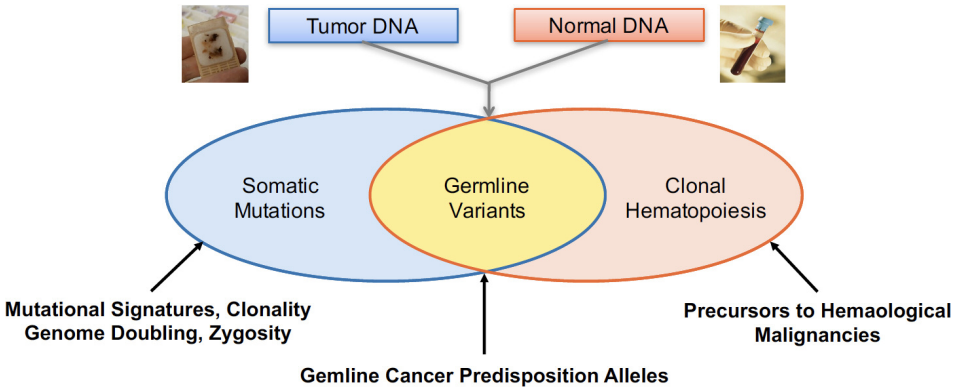


Figure 1 Different classes of alterations detected by matched tumor/normal MSK-IMPACT sequencing

the development of secondary hematological malignancies [4]. Moreover, our clinical sequencing strategy has produced a comprehensive genomic and clinical dataset that can be mined to discover molecular biomarkers that predict therapeutic response or outcome within and across cancer types, such as whole genome doubling [5]. The high depth of MSK-IMPACT enables precise estimates of mutant allele fraction and the inference of zygosity, revealing oncogenic mutations that undergo allelic imbalance. Our analysis showed that imbalance typically favors enrichment of the oncogenic allele and can serve as a predictive biomarker [6]. By providing broad access to genomic data and accompanying clinical annotation to researchers through initiatives such as the American Association for Cancer Research (AACR) Genomics Evidence Neoplasia Information Exchange (GENIE) consortium [7], we hope to accelerate translational research programs throughout the scientific community.

Comprehensive molecular profiling across all cancer types has revealed clinically actionable alterations and facilitated enrollment on genomically matched clinical trials for many, but certainly not all, individuals. Current efforts to characterize complex genomic features, develop improved clinical decision support tools, and establish additional molecular tests including liquid biopsies promise to further expand the benefits of genome-driven oncology to cancer patients.

References

1. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, Srinivasan P, Gao J, Chakravarty D, Devlin SM, Hellmann MD, Barron DA, Schram AM, Hameed M, Dogan S, Ross DS, Hechtman JF, DeLair DF, Yao J, Mandelker DL, Cheng DT, Chandramohan R, Mohanty AS, Ptashkin RN, Jayakumaran G, Prasad M, Syed MH, Rema AB, Liu ZY, Nafa K, Borsu L, Sadowska J, Casanova J, Bacares R, Kiecka IJ, Razumova A, Son JB,

Stewart L, Baldi T, Mullaney KA, Al-Ahmadie H, Vakiani E, Abeshouse AA, Penson AV, Jonsson P, Camacho N, Chang MT, Won HH, Gross BE, Kundra R, Heins ZJ, Chen HW, Phillips S, Zhang H, Wang J, Ochoa A, Wills J, Eubank M, Thomas SB, Gardos SM, Reales DN, Galle J, Durany R, Cambria R, Abida W, Cercek A, Feldman DR, Gounder MM, Hakimi AA, Harding JJ, Iyer G, Janjigian YY, Jordan EJ, Kelly CM, Lowery MA, Morris LGT, Omuro AM, Raj N, Razavi P, Shoushtari AN, Shukla N, Soumerai TE, Varghese AM, Yaeger R, Coleman J, Bochner B, Riely GJ, Saltz LB, Scher HI, Sabbatini PJ, Robson ME, Klimstra DS, Taylor BS, Baselga J, Schultz N, Hyman DM, Arcila ME, Solit DB, Ladanyi M, Berger MF. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nature Medicine*. 2017; 23(6):703-713.

2. Middha S, Zhang L, Nafa K, Jayakumaran G, Wong D, Kim HR, Sadowska J, Berger MF, Delair DF, Shia J, Stadler Z, Klimstra DS, Ladanyi M, Zehir A, Hechtman, JF. Reliable Pan-Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation Sequencing Data. *JCO Precision Oncology*. 2017 Nov.
3. Mandelker D, Zhang L, Kemel Y, Stadler ZK, Joseph V, Zehir A, Pradhan N, Arnold A, Walsh MF, Li Y, Balakrishnan AR, Syed A, Prasad M, Nafa K, Carlo MI, Cadoo KA, Sheehan M, Fleischut MH, Salo-Mullen E, Trottier M, Lipkin SM, Lincoln A, Mukherjee S, Ravichandran V, Cambria R, Galle J, Abida W, Arcila ME, Benayed R, Shah R, Yu K, Bajorin DF, Coleman JA, Leach SD, Lowery MA, Garcia-Aguilar J, Kantoff PW, Sawyers CL, Dickler MN, Saltz L, Motzer RJ, O'Reilly EM, Scher HI, Baselga J, Klimstra DS, Solit DB, Hyman DM, Berger MF, Ladanyi M, Robson ME, Offit K. Mutation Detection in Patients With Advanced Cancer by Universal Sequencing of Cancer-Related Genes in Tumor and Normal DNA vs Guideline-Based Germline Testing. *JAMA*. 2017; 318(9):825-835.
4. Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, Hyman DM, Solit DB, Robson ME, Baselga J, Arcila ME, Ladanyi M, Tallman MS, Levine RL*, Berger MF*. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell*. 2017; 21(3):374-382.e4.
5. Bielski CM, Zehir A, Penson AV, Donoghue MTA, Chatila W, Armenia J, Chang MT, Schram AM, Jonsson P, Bandlamudi C, Razavi P, Iyer G, Robson ME, Stadler ZK, Schultz N, Baselga J, Solit DB, Hyman DM, Berger MF, Taylor BS. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nature Genetics*. 2018; 50(8):1189-1195.
6. Bielski CM, Donoghue MTA, Gadiya M, Hanrahan AJ, Won HH, Chang MT, Jonsson P, Penson AV, Gorelick A, Harris C, Schram AM, Syed A, Zehir A, Chapman PB, Hyman DM, Solit DB, Shannon K, Chandarlapaty S, Berger MF, Taylor BS. Widespread Selection for Oncogenic Mutant Allele Imbalance in Cancer. *Cancer Cell*. 2018; 34(5):852-862.e4.

7. AACR Project GENIE Consortium. AACR Project GENIE: Powering precision medicine through an international consortium. *Cancer Discovery* 7(8):818-831, 2017.

Michael F. Berger, PhD

1997-2001	BA in Physics, Princeton University
2002-2008	PhD in Biophysics, Harvard University
2008-2010	Research Scientist, The Broad Institute
2010-2017	Assistant Member, Memorial Sloan Kettering Cancer Center
2013-present	Associate Director, Marie Josée and Henry Kravis Center for Molecular Oncology
2017-present	Associate Member, Memorial Sloan Kettering Cancer Center

A SYSTEMS LEVEL VIEW OF BREAST CANCER

Carlos Caldas
Cancer Research UK Cambridge Institute
University of Cambridge
Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK
(Carlos.Caldas@cruk.cam.ac.uk)

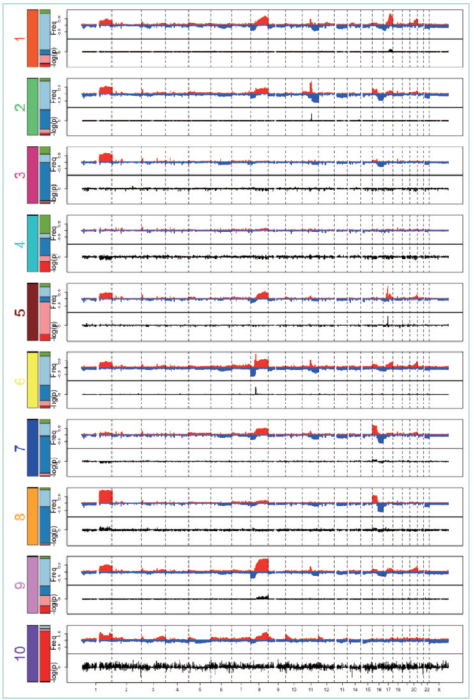
We have redefined the molecular taxonomy of breast cancer showing it is a constellation of 11 distinct genomic subtypes (Integrative Clusters- IntClust). These IntClust have distinct drivers, genomic architectures and TMEs. Explant models representing these IntClust have been generated and have been used for pre-clinical drug testing and for studying drug resistance. Using a combination of genomics, digital pathology, liquid biopsies and explant models allows us to have an integrated view of breast cancers and constitutes a platform for studying its biology and explore novel therapeutics.

My talk covered the following topics:

- Driver-based classification of breast cancer
- Breast cancers as communities of clones and communities of cells
- TME and tumour neighbourhoods
- PBCP: prospective WGS and RNAseq for characterization of 2,250 breast cancers in the clinic
- Cancer cell/TME/therapy 'trialogue'
- Modelling breast cancer- PDTXs/PDTCs
- Integrated breast cancer medicine- a systems level platform

Driver-based classification of breast cancer

Breast cancer: 11 diseases with distinct CNA drivers, prototypical SNV landscapes and characteristic clonal evolution patterns



ARTICLE

The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups

Curtis, 2012

LETTER

The shaping and functional consequences of the microRNA landscape in breast cancer

Dvinge, 2013

Review

A new genome-driven integrated classification of breast cancer and its implications

Dawson, 2013

Genome Biology

RESEARCH Open Access

Genome-driven integrated classification of breast cancer validated in over 7,500 samples

Ali, 2014

nature COMMUNICATIONS

ARTICLE

The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes

Pereira, 2016

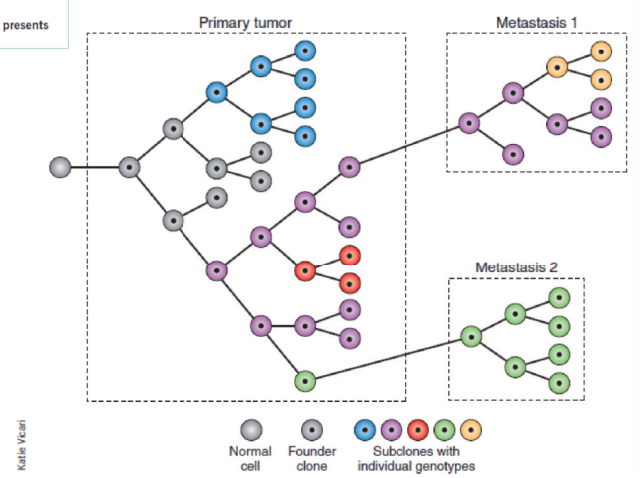
Breast cancers as communities of clones and communities of cells

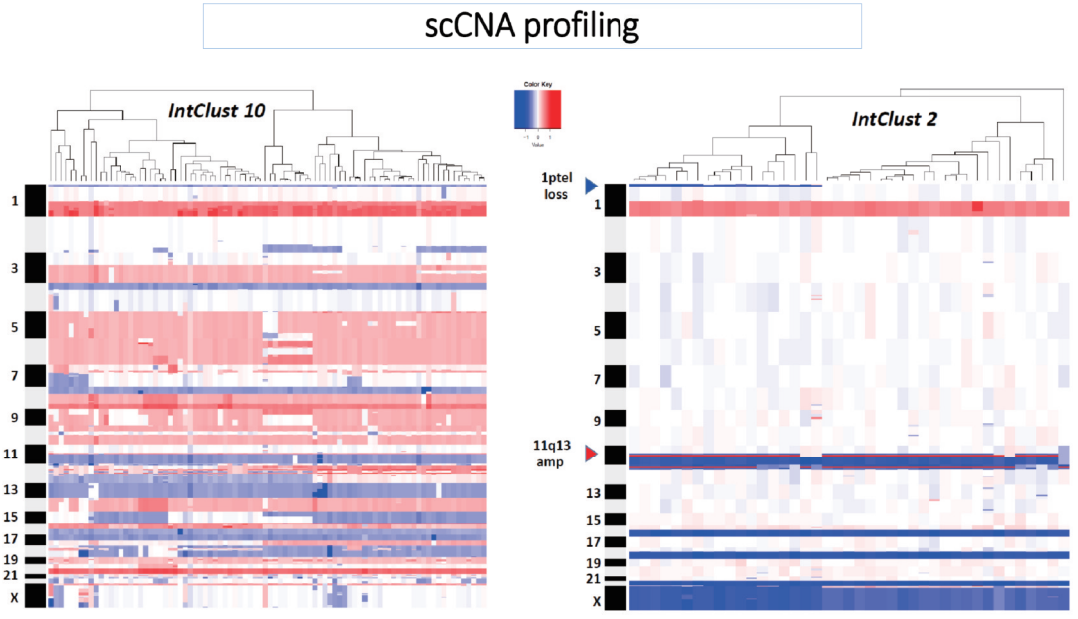
Cancer sequencing unravels clonal evolution

Carlos Caldas

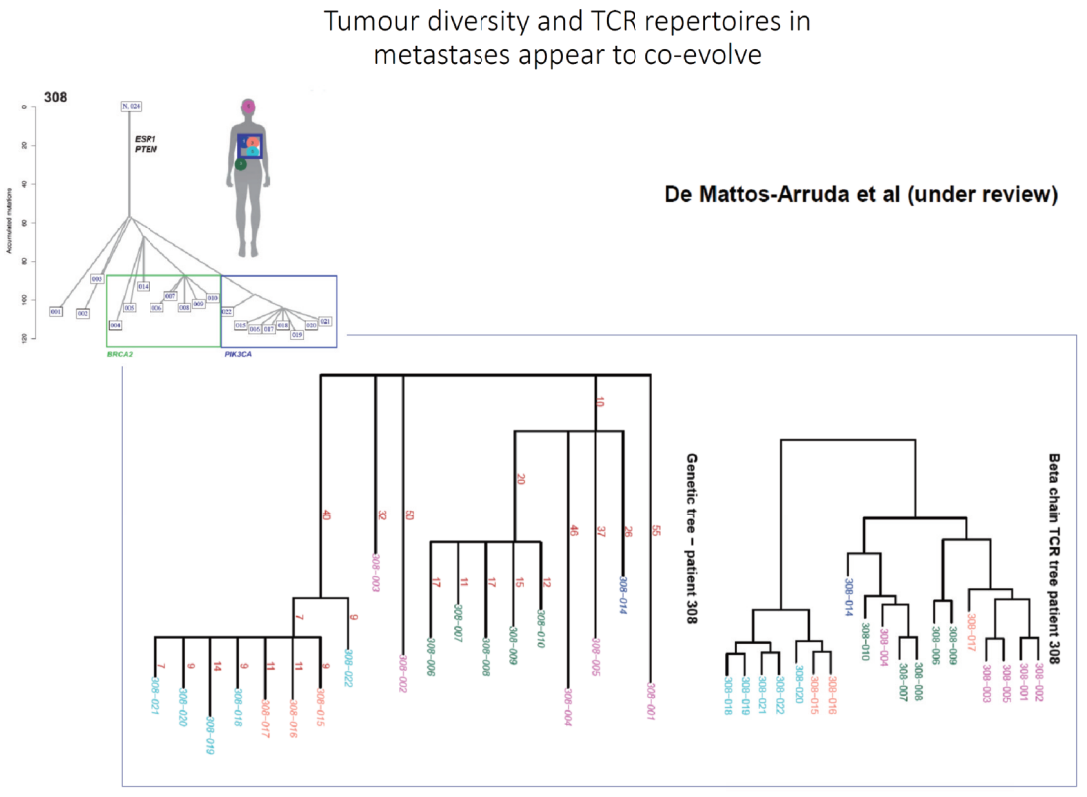
Characterization of tumor heterogeneity at the sequence level presents new challenges and opportunities for targeted therapies.

NEWS AND VIEWS





TME and tumor neighborhoods

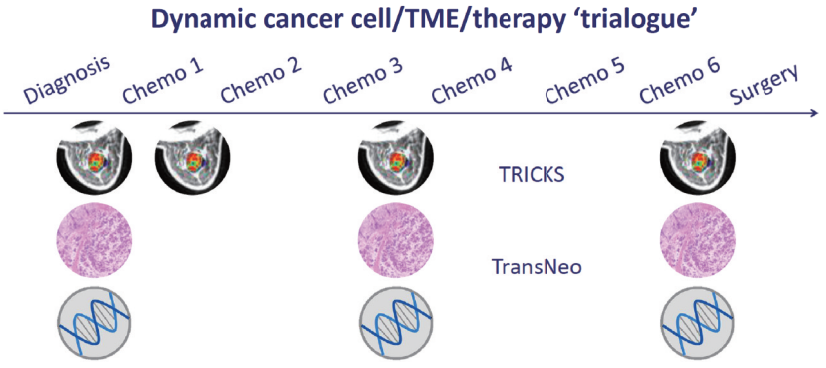


PBCP: prospective WGS and RNAseq for characterization of 2,250 breast cancers in the clinic

PBCP

- 2,250 consecutive pts WGS/RNAseq
- Molecularly stratified cohort
- Molecular tumour monitoring
- Repurposing drugs
- Understanding outstanding responders
- Clinical trials
- Basket/n=1 trials
- Patient participation and empowerment
- Contribution to global cancer genome efforts

Cancer cell/TME/therapy 'trialogue'

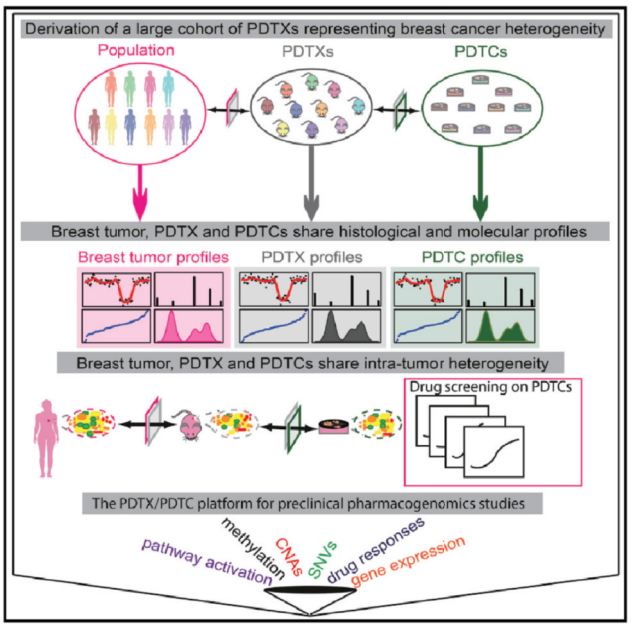


Modelling breast cancer- PDTXs/PDTCs

Cell

A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds

Graphical Abstract



Authors

Alejandra Bruna, Oscar M. Rueda, Wendy Greenwood, ..., Violeta Serra, Mathew J. Garnett, Carlos Caldas

Correspondence

carlos.caldas@cruk.cam.ac.uk

In Brief

Development and analysis of a collection of breast-cancer-patient-derived xenografts indicate that the xenografts and cell cultures derived from them preserve the heterogeneity of the original tumors and can be used for drug screening.

Integrated breast cancer medicine- a systems level platform

- Genomic stratification and monitoring (WGS/RNAseq/ctDNA/TCRseq/BCRseq)
- New pathology: 3D resolved single cell maps of tumours (tumour cells, TME, niches)
- Therapeutic predictive models (PDTXs/PDTCs)
- Anticipatory genomics (Humanized PDTXs/ctDNA) in co-clinical trials

References

1. Curtis C, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. ***Nature***, 486: 346-352, 2012
2. Dawson S-J, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. ***The New England Journal of Medicine***, 368: 1199-1209, 2013
3. Murtaza M, et al. Noninvasive analysis of acquired resistance to cancer therapy by exome sequencing of plasma DNA. ***Nature***, 497: 108-112, 2013 [*co-corresponding/co-senior author]
4. Dvinge H, et al. The shaping and functional consequences of the miRNA landscape in breast cancer. ***Nature***, 497: 378-382, 2013
5. Murtaza M, et al. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. ***Nature Communications***, 6: 8760, 2015
6. Pereira B, et al. The somatic mutation profiles of 2433 primary breast cancers refine their genomic and transcriptomic landscapes. ***Nature Communications***, 7: 11479, 2016
7. Bruna A, et al. A biobank of breast cancer explants with preserved intra-tumour heterogeneity to screen anticancer compounds. ***Cell***, 167: 260-274, 2016
8. Ali HR, et al. Patterns of immune infiltration in breast cancer and their clinical implications: a gene expression-based retrospective study. ***PLOS Medicine***, 13: e1002194, 2016

Carlos Caldas, MD, FMedSci

2002-present	Professor of Cancer Medicine, University of Cambridge
2006-present	Senior Group Leader, Cancer Research UK Cambridge Institute
2016-present	Director, Breast Cancer Programme, Cambridge Cancer Centre, University of Cambridge

THE INTERNATIONAL CANCER GENOME CONSORTIUM – ARGO INITIATIVE: BRINGING A MILLION PATIENT YEARS OF PRECISION ONCOLOGY KNOWLEDGE TO THE WORLD

Andrew V. Biankin
Wolfson Wohl Cancer Research Centre, Institute of Cancer Sciences
University of Glasgow
Garscube Estate, Switchback Road, Bearsden, Glasgow Scotland G61 1BD, UK
(Andrew.Biankin@glasgow.ac.uk)

The most valuable information for someone with the diagnosis of cancer is the ability to predict which treatment will be most effective for them, the likely outcome, and their overall prognosis. Our ability to do this is currently limited to a few examples that as proof of concept demonstrate improved outcomes and efficiency. The most effective way to generate the knowledge required to do this for all people affected by cancer is to amass large datasets of deep molecular and detailed phenotypic data through health systems as part of routine care. However, the pathways and processes are not in place and require development, with the paramount challenge of acquisition and application of rich, reliable clinical data including treatment and response. ICGC-ARGO will build these datasets, initially from clinical trials and from well-annotated cohorts, and together with health systems develop the frameworks that integrate robust clinical data acquired through routine patient care with molecular data. Although multiple regional platforms have been established to generate cancer genomics data, no platform exists that can capture, aggregate, harmonize and appropriately share data globally to enable granular analysis of pooled individual participant data.

ICGC-ARGO is an international network of cancer clinicians, researchers and clinical trials groups that aims to provide **a million patient-years of precision oncology knowledge** in a manner that allows for broad, but ethically responsible, data sharing and research. Clinical and genomic data generated by ICGC-ARGO members will be exclusively available to its membership for a short period of time before being released to the broader research community.

ICGC-ARGO will acquire detailed treatment and response information from patients along with genomic, transcriptomic and proteomic data from their tumours and healthy tissues. This data will allow for precise correlation of molecular aberrations with clinical features, allowing for the discovery of molecular markers of response, resistance and toxicity of treatment. This knowledge will translate into new approaches to improve outcomes for people affected by cancer.

The ARGO Project (Accelerating Research in Genomic Oncology)

ICGC-ARGO is the second phase of the International Cancer Genome Consortium. The first initiative comprehensively mapped the structural aberrations of cancer genomes and advanced our understanding of the molecular basis of cancer. ICGC focused on primary cancers prior to therapy, and to date has collected and distributed knowledge on nearly 21,000 primary cancers that have been characterised at the genomic, transcriptomic and epigenomic levels. Data generated through the ICGC has transformed research strategies in academia and industry alike, with hundreds of seminal works arising from ICGC data, and landmark articles appearing in the world's elite scientific journals. No therapeutic is developed today without, in some way, applying the knowledge that ICGC has provided the world. Although ICGC has achieved much, pivotal outstanding challenges remain to be addressed; unanswered questions that are vital in our quest to defeat cancer.

The key questions addressed by ARGO are:

1. *How do we use current treatments better?*
2. *How does a cancer change with time and treatment?*
3. *How do we practically implement these approaches in healthcare and therapeutic development?*
4. *How do we advance early detection and ultimately prevent cancer?*

The ICGC and its membership have established global networks, genomic sequencing capacity, and computational expertise for the collection, integration, analysis and distribution of large cancer genome data sets. Tens of thousands of scientists, clinicians and technologists have built the methodologies and mechanisms to advance genomic health. The ICGC forms a strong foundation on which to build and advance cancer research through integrating the expertise and capacity of the ICGC with clinical research and cancer care.

Over the next ten years ICGC-ARGO will generate, acquire and integrate genomic and clinical data on more than 200,000 patients (1 million patient-years). Clinical data will include information concerning lifestyle, co-morbidity, diagnostics, toxicity, response to therapy and survival. Using this large-scale integrated data, researchers, scientists,

policymakers and clinicians will be able to work with patients, health care providers, industry, and others to advance therapeutic development with interventions based on matching the patient's disease molecular subtype with the most effective treatment; develop preventative strategies; markers for early detection of disease; and more specific criteria and methods for diagnosis and prognostication.

Joining ICGC-ARGO

Any researcher or research group from academia or industry may become an ICGC-ARGO member. Members must commit to the acquisition of molecular and clinical data from a minimum number of cancer patient donors; there are several levels of membership that correspond to the number of donors committed. The sources of cohorts of patients that would constitute ICGC-ARGO projects may include:

- Biospecimens from participants enrolled in active clinical trials;
- Analyses of banked samples from past clinical trials;
- Analyses of samples from clinically well-annotated cohorts that satisfy ICGC-ARGO clinical data requirements;
- Longitudinal cohort studies;
- Autopsy studies with detailed clinical data
- Population-based studies with detailed clinical and lifestyle data
- Real World Data acquired through health systems.

To apply, prospective members are asked to fill in an expression of interest that describes their project and level of commitment.

Benefits of Membership

ICGC-ARGO members will be able to submit patient genomic, transcriptomic and clinical data to one of a series of ARGO regional data processing centres, where it will be subjected to state-of-the-art QC, alignment, variant calling, annotation and clinical harmonization. The harmonized data will be returned to members in a form that allows it to be compared to data collected by all other ARGO participants, and a copy of the data will be retained by the processing centre for merging into a central compute cloud-accessible database of all ARGO results. Members will have exclusive access to the data they generate for a period of 12 months, after which it will become available to all members of the consortium. After 24 months, the data will be available to external parties.

Industry Partners

Industry partners are welcome to be part of ICGC-ARGO. If organisations wish to contribute data, then they may also become ICGC Members as detailed above. For

prospective partners that will not contribute data, in-kind contributions of either $5 million or $10 million US gives them Gold or Platinum Industry partner status, which includes membership of Working Groups and select committees.

Andrew V. Biankin, BMedSc, MBBS(HONS), FRACS, FFS(RCPA), FRCS(Glasg), FRCS(Edin), PhD, FRSE, FMedSci

2011	Specialist Consultant in Precision Oncology and Therapeutic Development
2013	Regius Chair of Surgery, University of Glasgow
2013	Director, Wolfson Wohl Cancer Research Centre, University of Glasgow
2016	Chairman, Precision-Panc Therapeutic Development Platform
2017	Executive Director, International Cancer Genome Consortium

HIGH-THROUGHPUT ANNOTATION OF VARIANTS OF UNKNOWN SIGNIFICANCE

Hiroyuki Mano
Division of Cellular Signaling, National Cancer Center Research Institute
5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(hmano@ncc.go.jp)

Numerous variants of unknown significance (VUS) have been identified through large-scale cancer genome projects, although their functional relevance remains uninvestigated. The L858R substitution or exon 19 deletions are, for instance, known to confer activation of epidermal growth factor receptor (EGFR) kinase, but there are more than one thousand of nonsynonymous mutations within *EGFR* reported in the COSMIC database. It is still obscure whether such mutations have some relevance to EGFR kinase activity or are merely passenger mutations (Figure 1).

To drastically increase the speed of functional annotation, here we have developed a mixed-all-nominated-mutants-in-one (MANO) method to evaluate the transforming potential and drug sensitivity of oncogene VUS in a high-throughput manner. Each mutant cDNA is stably integrated into the genome of assay cells (such as mouse 3T3 fibroblasts and an IL-3–dependent, murine pro-B cell line, Ba/F3) along with 6 bp-barcode sequences

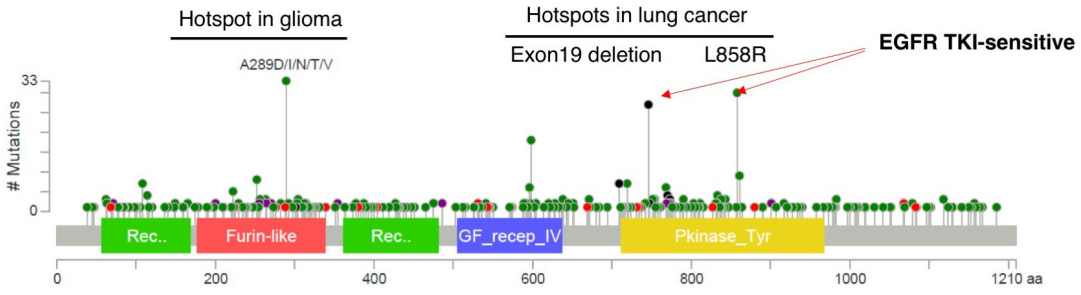


Figure 1 More than 1,000 nonsynonymous mutations are reported for *EGFR*

(Figure 2). Separately transduced assay cells are then pooled together, and cultured in a competitive manner to evaluate their transforming potential or drug sensitivity *in vitro* or *in vivo*. At the end of the expansion period, genomic DNA is extracted from the assay cells to PCR-amplify the barcode sequences that are then subjected to deep sequencing with the Illumina MiSeq platform in order to quantitate the relative abundance.

To validate its usefulness, we treated with various tyrosine kinase inhibitors (TKIs) a pool of 16 Ba/F3 cells expressing active EGFR mutants (n = 11) or other oncoproteins (n = 5). Whereas treatment with a cytotoxic compound, puromycin, induced uniform cell death across the cell clones, that with EGFR TKIs (gefitinib, erlotinib, afatinib, osimertinib and rociletinib) resulted in the dose-dependent death of cells for five TKIs-sensitive EGFR mutant (L858R, E746_A750 del, G719S, E861Q, and S768I) in the pool (Figure 3). As expected, Ba/F3 expressing EGFR(T790M) was resistant against the first- and second-generation EGFR TKIs (gefitinib, erlotinib and afatinib), but sensitive to the third-generation EGFR TKIs (osimeritinib and rociletinib). On the contrary, cells expressing EGFR(T790M/C797S) were shown resistant to such third-generation TKIs. Similarly, crizotinib, a TKI for ALK and ROS1, suppressed the growth of the cells expressing EML4-ALK or CD74-ROS1, and another inhibitor for ALK and RET, alectinib, inhibited growth of the cells expressing EML4-ALK or KIF5B-RET.

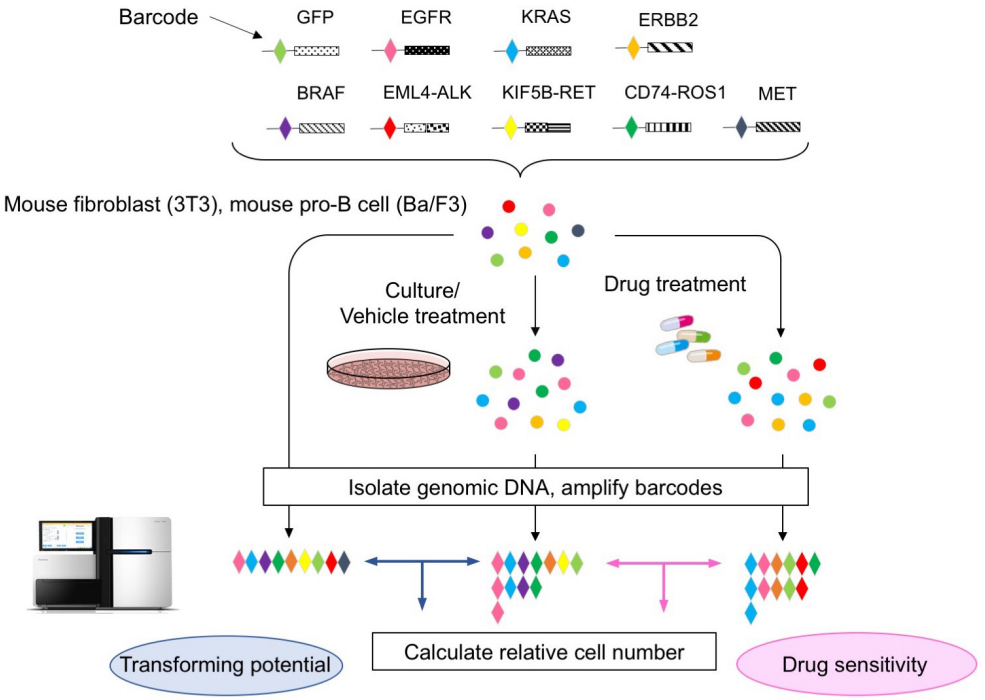


Figure 2 The MANO method

We further applied this method to 101 nonsynonymous *EGFR* mutants, and discovered a number of mutations conferring resistance to EGFR TKI, including gefitinib- and erlotinib-insensitive missense mutations within exon 19 and other gefitinib-resistant mutations such as L833V, A839T, V851I, A871T, and G873E. Importantly, 12.8% of *EGFR(L858R)*-positive tumors harbored compound mutations in the *cis* allele, which decrease the gefitinib sensitivity of these tumors. The MANO method further revealed that some EGFR mutants that are highly resistant to all types of TKIs are sensitive to cetuximab [1].

In addition, the MANO method was used to assess comprehensively the transforming activities as well as drug sensitivities of *ERBB2* mutations. We evaluated 55 non-synonymous *ERBB2* mutations which are reported recurrently in COSMIC database, and discovered several novel activating mutations which probably drive tumorigenesis. Furthermore, *ERBB2* mutations showed varying drug sensitivities to ERBB2-targeted inhibitors. Thus, the MANO method may be a novel approach for assessing VUS, and our findings in this study will be beneficial to deliver personalized medicine to cancer patients harboring *ERBB2* mutations [2]. Thus, these data support the importance of examining the clinical relevance of uncommon mutations within oncogenes and of evaluating the functions of such mutations in combination.

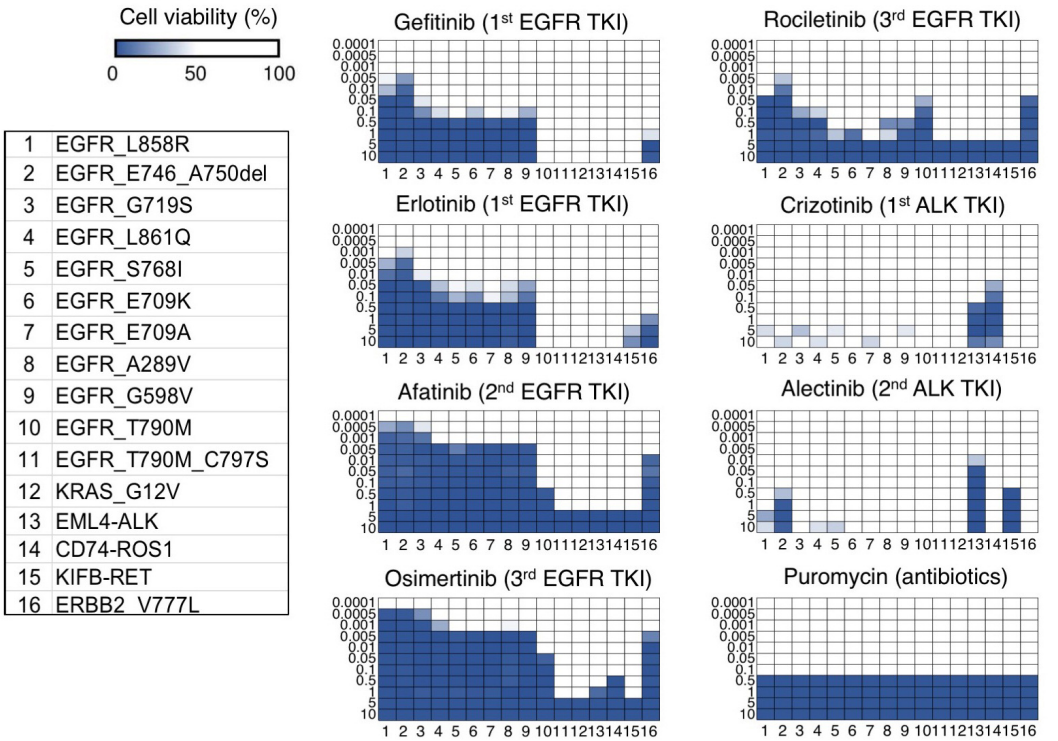


Figure 3 Proof-of-concept of the MANO method

References

1. Kohsaka S, Nagano M, Ueno T, Suehara Y, Hayashi T, Shimada N, Takahashi K, Suzuki K, Takamochi K, Takahashi F, Mano H. A method of high-throughput functional evaluation of EGFR gene variants of unknown significance in cancer. *Sci Transl Med* 9, eaan6556, 2017.
2. Nagano M, Kohsaka S, Ueno T, Kojima S, Saka K, Iwase H, Kawazu M, Mano H. High-throughput functional evaluation of variants of unknown significance in ERBB2. *Clin Cancer Res*, doi: 10.1158/1078-0432.CCR-18-0991, 2018.

Hiroyuki Mano, MD, PhD

1984	Graduated from School of Medicine, Faculty of Medicine, The University of Tokyo
1986-1989	Clinical Fellow, The Third Department of Internal Medicine, Faculty of Medicine, The University of Tokyo
1989-1991	Postdoctoral Researcher, Department of Biochemistry, St. Jude Children's Research Hospital
1991-1993	Assistant Professor, The Third Department of Internal Medicine, Faculty of Medicine, The University of Tokyo
1993-2001	Associate Professor, Department of Molecular Biology, Jichi Medical University
2001-2013	Professor, Division of Functional Genomics, Jichi Medical University
2013-2018	Professor, Department of Cellular Signaling, Graduate School of Medicine, The University of Tokyo
2016-present	Chief, Division of Cellular Signaling, Director, National Cancer Center Research Institute
2018-present	Director, Center for Cancer Genomics and Advanced Therapeutics, National Cancer Center

CONCLUDING REMARKS

Tatsuhiro Shibata
Laboratory of Molecular Medicine
The Institute of Medical Science, The University of Tokyo (IMSUT)
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
(tshibata@ims.u-tokyo.ac.jp)

Division of Cancer Genomics
National Cancer Center Research Institute
5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(tashibat@ncc.go.jp)

The 49th International Symposium of the Princess Takamatsu Cancer Research Fund, entitled "Deciphering, Simulating and Editing of the Cancer Genome", focused on the current status of cancer genomics, data informatics, and the potential application of gene editing technology in the cancer research area. Since cancer is a disease of *the genome*, exploration of somatic and germline genetic and epigenetic alterations in the cancer genome is fundamental and essential for the diagnosis, treatment and prevention of cancer. Here, 30 speakers, who are experts in cancer genomics, tumor heterogeneity, or gene editing, and 196 participants have joined together and discussed the current status and future directions in this research area. In particular, this meeting aimed to accomplish the following three purposes: Understanding and sharing information on the current cutting-edge developments in the field of cancer genomics, Promoting communication, interactions and future collaborations among participants, especially young Japanese scientists and speakers from abroad, and Discussion about the future of cancer genomics.

In this meeting, a wide range of topics related to cancer genome research have been intensively discussed, including coding and non-coding cancer driver gene landscapes, genetic changes in precancerous lesions, germline contribution, mutational signatures in carcinogenesis processes, international and trans-ethnic-scale cancer genome projects, analytical informatics, immunogenomics, deciphering tumor genome heterogeneity at the single cell level, and application of gene editing for cancer research. Hereafter, several topics are highlighted.

Large-scale collections of cancer genomes by the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) has now enabled inference of the saturating list of cancer driver genes, especially in the coding genome. Most cancer types have long-tailed lists of driver genes with a minority of frequently altered ones and many rare ones. Functional annotations of these cancer drivers enable identification of essential and therapeutically targetable molecules, and these efforts have led to the application of cancer genome sequencing in the clinic. To further promote this, an international consortium, called ICGC-ARGO, is planning to collect and share more such detailed clinical information together with cancer genome information. On the other hand, driver events in the non-coding genome, such as promoter mutations and, more frequently, structural alterations that affect regulatory regions, such as enhancers, remain to be well explored, and hence, are likely to be some of the hottest fields in future research.

Recent multi-region sequencings of cancer tissues have uncovered clonal evolution processes during carcinogenesis. In addition to the natural selection of the fittest clones, neutral evolution or branched evolution in cancers have also been reported, which lead to genetic heterogeneity in the cancer cell population. Recent advances in single-cell genomics have enabled exploration of the detailed processes and status of intra-tumor heterogeneity. Accumulation of clonal and subclonal mutations during tumor evolution naturally generates non-self-antigens, called *neo-antigens*, which are targets of host immune surveillance. Consequently, cancer cells would evade immune cell attack through multiple ways, including by activating immune checkpoint molecules, such as PD-L1, or inactivating the antigen presenting system. These are new types of cancer driver genes and have been intensively discussed together with co-existing cancer-immune microenvironments.

Cancer is also an accumulative process of genetic alterations. Large-scale cancer genome data has elucidated that there exist characteristic patterns of these processes. In the case of somatic substitutions, more than 40 mutational patterns, called mutational signatures, have been currently identified in human cancer genomes, and, importantly, these signatures mirror unique carcinogenesis processes, such as exposure to carcinogens (smoking, Aflatoxins, etc.) or DNA repair deficiencies. Understanding and elucidating mutational processes would lead to efficient cancer prevention. Mutational processes also operate in precancerous lesions, and deep sequencing of precancerous lesions in the liver have uncovered unique mutational processes, patchy clonal structures, and non-random virus genome integrations. In the near future, more detailed genomic views of precancerous lesions will be uncovered, which will contribute to advancing our understanding of the processes involved in human carcinogenesis.

Since key scientists in the field of cancer genomics have participated in this meeting, we

circulated a questionnaire on the future directions of this research field, which was generated following a discussion with Ms. Potenski, a senior editor of the *Nature Genetic*s journal. Many valuable suggestions and opinions were voiced. As an example, responses to the question 'What topics/challenges in the cancer genomics field will be important in the next 2-5 years?' included: deciphering the complete structure of cancer genomes and the impact of non-coding elements (Aneuploidy, structural alterations, regulatory elements), determining how cancer genomes are different from those of benign neoplasms, genetic changes and clonal structure of precancerous tissue, understanding the degree of plasticity of cancer genomes and their impact on cancer evolution, understanding the genomic characteristics of the immune response to cancer, early detection and interception of cancer, interplay between germline variants, somatic alterations and lineage, cancer epigenome, and artificial intelligence for supporting clinical decision making, some of which have been further discussed in this meeting.

At this meeting, we also introduced the system of asking participants to submit a document pledging their non-disclosure of confidential data that was discussed during this meeting, which would promote the presentation of unpublished data by speakers. This is the first trial of this kind of a disclosure, but we believe that this should be continued in subsequent meetings.

Finally, I would like to express my sincere appreciation to the co-organizers of this meeting, Drs. Matthew Meyerson, Peter Campbell and Hiroyuki Aburatani, for their conception of the meeting sessions and selection of speakers, and all the speakers and participants for the active and fruitful discussions, all of which have greatly contributed to the success of the meeting. I also thank the general secretary of the meeting, Dr. Shinichi Yachida, staffs from the National Cancer Center, Osaka University, The University of Tokyo, and the secretarial office of the Princess Takamatsu Cancer Research Fund for their efficient management. I hope that this meeting will promote future international collaborations and help young cancer scientists to join this very exciting field.