UNIVERSITÀ DEGLI STUDI DI TORINO

Title: Towards understanding and exploiting tumor heterogeneity

Authors (listed alphabetically)

Ash Alizadeh (Department of Medicine, Divisions of Oncology, Hematology, and Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305 USA), Victoria Aranda (Nature Medicine), Alberto Bardelli (University of Torino, Department of Oncology, SP 142, Km 3.95, 10060 Candiolo, Torino, Italy & Candiolo Cancer Institute – FPO, IRCCS, Candiolo, Torino, Italy), Cedric Blanpain (Université Libre de Bruxelles (ULB), Brussels), Christoph Bock, Christine Borowski (Nature Medicine), Carlos Caldas (Department of Oncology, University of Cambridge, Cambridge, UK), Andrea Califano, Michael Doherty, Markus Elsner (Nature Biotechnology), Manel Esteller, Rebecca Fitzgerald (MRC Cancer Unit, Hutchison-MRC Research Centre, University of Cambridge), Bertil Johnsson, Jan Korbel, Peter Lichter, Christopher Mason, Nick Navin, Dana Pe'er, Kornelia Polyak (Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA), Charles Roberts, Lillian Siu (Princess Margaret Cancer Centre, Toronto, Canada), Alexandra Snyder-Charen, Hannah Stower (Nature Medicine). Charles Swanton (University College London Cancer Institute Cancer Research UK Lung Cancer Centre of Excellence and University College London Hospitals NHS Foundation Trust.The Francis Crick Institute, 44 Lincoln's Inn Fields, London, UK), Roland Verhaak (Department of Genomic Medicine, Department of Bioinformatics and Computational Biology, UT MD Anderson Cancer Center, Houston, TX 77030 USA), Jean C. Zenklusen, Johannes Zuber, Jessica Zucman-Rossi

Introduction
In many malignancies, molecular and cellular heterogeneity within a single tumor, between different sites of disease in a single patient, and between tumors in different patients confounds our understanding of tumor evolution and our ability to design and select effective therapies, as well as to curtail treatment resistance.
http://www.nature.com/nrclinonc/journal/vaop/ncurrent/full/nrclinonc.2015.73.html
http://www.ncbi.nlm.nih.gov/pubmed/22513401
http://www.cell.com/cancer-cell/abstract/S1535-6108%2814%2900510-8

We still, however, are at the very beginning of understanding the full extent of tumor heterogeneity (including the contribution of the tumor microenvironment), which types/aspects of tumor heterogeneity are relevant in which tumor types and in which clinical scenarios, and how to counter and/or exploit tumor heterogeneity for therapeutic gain.

To begin to tackle these issues, *Nature Medicine*, *Nature Biotechnology* and the *Volkswagen Foundation* invited a group of 20 scientists from around the globe for a two day brainstorming session in the beautifully restored Herrenhausen Palace in Hanover,

Germany **[callout to Figure depicting Herrenhausen Palace grounds--Herrenhausen and Susan Kim to provide this figure]**.

Reflecting the variety of expertise needed to tackle the issues mentioned above, this group included computational biologists, technology developers, cancer biologists, clinicians, industry representatives and regulators. The aims were to identify the most important questions that need to be answered about tumor heterogeneity, and map paths to answering them. Hopefully the new collaborations and networks forged at the meeting will help make some of these paths a reality.

All in attendance felt that sharing the group's findings--especially the questions identified as most pivotal--with the broader community was key. This Perspective aims to do just that, and is organized in the same manner as the two day meeting. Whereas the first day involved all attendees brainstorming as a single larger group about the most important questions needing answers, the second day was spent in four smaller discussion groups (Cancer Evolution, Beyond the Genome, Clinical/Regulatory and Technology) brainstorming about the answers to four or five selected questions. At the end of the second day, each group crafted a presentation to reveal their conclusions to the larger group. The question and answer period that resulted proved to be a highlight of the meeting.

Cancer Evolution

While many biological aspects of tumor heterogeneity are unknown, the group focused on establishing the basic premises by which we can define and study the parameters of tumor evolution.

What is a clone?

This term is used widely in the field but the discussion in this group revealed that perhaps surprisingly there is no consensus as to what it indicates; in fact this question sparked some of the most animated discussion at the meeting. In principle, under the assumption that tumors arise from a single cell, each tumor can be considered a clone. In this scheme, trunk mutations have thus a cancer cell fraction (CCF) of 1. All cells within a tumor with a CCF<1 could be considered subclones, at least with regard to their relative population frequency within a given lesion. However, the group recognized that even this definition is misleading, due to an illusion of clonality within a single biopsy, where a particular mutation can appear clonal within one biopsy with a CCF of 1, but on subsequent tumour sampling the mutation may be subclonal or absent altogether.

What is a driver?

The term 'driver' is typically used to denote a genetic event associated in some way with tumor progression. Although it might traditionally be viewed as a tumor cell-autonomous alteration that promotes tumor proliferation, after discussion we felt it may be useful to extend this definition to encompass a broader slice of the complex biology of pro-tumorigenic events. In other words, a broader biological definition of cancer driver would be a cell autonomous *or* non-autonomous alteration that contributes to tumor evolution at any stage, including initiation, progression, metastasis and resistance to therapy, by promoting a variety of functions including proliferation,

survival, invasion, or immune evasion. Accordingly, drivers can be identified based on statistical analyses of genetic or epigenetic alterations, or by functional screens, and should ideally be confirmed by experimental evidence including preclinical *in vitro* and *in vivo* data, as well as clinical data. Complicating the matter further, as the role of a driver is constrained by spatial and temporal contexts, genetic events can act as drivers at one stage of tumorigenesis, but as passengers at another stage and vice versa.

What is the source of heterogeneity in cancer, and what is the contribution of heterogeneity to cancer evolution?

Regarding the source(s) of heterogeneity, while evolution is driven by selection of phenotypes according to their relative fitness, not all somatic genetic alterations have a recognizable phenotypic consequence and even fewer provide a fitness advantage. Selection for phenotypic alterations can favor the outgrowth of cells with genetic alterations associated with that phenotype. Therefore, when studying cancer evolution, it is likely that multidimensional phenotyping--measuring signaling, epigenetic, transcriptional, metabolic, and other alterations in addition to genetic alterations-- together with functional screening will be most informative in revealing the source(s) of the phenotype(s) that is driving tumorigenesis. Generating and interpreting these data is not trivial and the unanswered questions related to these issues are covered in the Technology section, below. Regarding the contribution of heterogeneity, while heterogeneity can be broadly considered as a trait that allows tumors to overcome evolutionary pressures, it can also reflect vulnerabilities that could be exploited therapeutically. This makes it even more important to develop tools to quantify and model tumor heterogeneity.

How can we model tumor heterogeneity in preclinical experiments?

One challenge to assessing the dynamic contribution of heterogeneity as a trait of tumor progression is the fact that current preclinical tumor models do not recapitulate the condition under which heterogeneous tumors arise and evolve in humans [http://www.nature.com/nm/journal/v21/n5/abs/nm.3853.html]. For example, although genetically engineered mouse models (GEMMs), have been instrumental in revealing crucial aspects of tumor biology, their tumors are relatively small and homogeneous, driven by a small number of genetic alterations, and can be polyclonal in nature. Tumor burden, metastatic potential and tumor longevity are also not recapitulated adequately in mouse models.. We need to apply new technologies to these problems; for example CRISPR/Cas-mediated genome editing can help recapitulate the genetic variability accumulated during human tumor evolution. Patient-derived xenograft (PDX) models do capture, at least at the start, some of the heterogeneity of patient samples. However, some subclones can be selected for increased fitness for growth in the mouse host, which lacks the proper microenvironmental and immune components that may otherwise influence subclonal selection. Ongoing efforts to humanize mouse models may help incorporate relevant features that shape tumor evolution in humans. Beyond animal models, *in vitro* approaches such as tumor slice cultures can be exploited to recapitulate a snapshot of

the tumor in its native environment, and organoids can be used to model tumorigenesis in human cells. *In silico* models that use multiscale parameters can also create interesting hypotheses that are experimentally testable. However, because no model is perfect, many in the group felt that there was no substitute for studying tumour evolution within the patient.

Beyond the Genome

What is the contribution of the epigenome to tumor phenotype and clinical outcome?

Cell states are defined by the interplay of the genome, epigenome, transcriptome and proteome in each tumor cell. Because cell states tend to be self-stabilizing, there are typically many fewer distinct cell states in a tumor than the amount of genetic, epigenetic, and transcriptional heterogeneity would suggest. Thus, even genetically distinct cells may be in a similar cell "state" and hence may be susceptible to treatment with the same drugs. On the other hand, even genetically identical cells can express a large number of substantially different "cell states" due to influence of the micro-environment and hence the epigenome. However we must stop thinking about the genetic and epigenetic contributions to cell state separately because both of their contributions to cell state may be intertwined. Furthermore, it is known that epigenetic defects, such as promoter CpG island hypermethylation-associated silencing of DNA repair genes, cause genetic changes; and translocations and mutations can also target epigenetic disruption thus resulting in further interaction between epigenetic and genetic traits.

We must strive to identify different cell states by integrating different datasets and once these are identified, we must work toward therapeutic strategies based on inferred cell states. Although epigenetic data forms only a part of such integrative analysis, because epigenetic modifications are dynamic and responsive to environmental pressures, they may exert a particularly strong role in the definition of the cell state and behavior at any given moment in time in response to therapy. In addition, although epigenetic marks are dynamic, they represent the history of the cancer, as once a cell has passed through a particular cell state, some of these epigenetic marks remain. Moreover, epigenetic marks, specifically regions of open chromatin, can also reflect the potential of the tumor to respond to an environmental or therapeutic pressure. Epigenetic marks are therefore unique in their ability to provide information about the previous, present and potential future states of a cell.

Because epigenetics provides a different and complementary paradigm to the analysis of genetic mutations, it may be possible in the future, once we have defined these states to use two or three important epigenetic markers to infer cell states. Furthermore, as the epigenetic state of cancers is more plastic than that of normal development, such contributions may be critical to understanding phenotypic changes of cancers such as the epithelial-mesenchymal transition, the capacity to disseminate beyond the primary site, and drug resistance.

What methods and samples do we need to describe and understand the heterogeneity and influence of the tumor microenvironment?

To understand the influence of the microenvironment on cell state, we need to coordinately characterize DNA sequence, epigenome, transcriptome, protein, metabolites and infiltrating immune cells in both the tumor and the stroma. Evaluation of data from single cells will provide additional insights into heterogeneity. Only through integration of such data, can we begin to develop a more robust understanding of cancer states. As a consequence, there will continue to be increasing need for computational biologists. Such technologies are discussed further below.

How do we increase the immunogenicity of tumors?

Immunogenicity depends on mutations that generate epitopes that are not recognised as "self" by tumor-infiltrating T lymphocytes. Therefore, chemotherapy and other genotoxic drugs may improve the outcome of immunotherapy interventions including adoptive T cell transfer and immune checkpoint blockade by generating mutations or modifying the immune microenvironment. (http://clincancerres.aacrjournals.org/content/20/21/5384.long). However, it is unclear whether subclonal changes in immunogenicity will be enough to cause the whole tumor to be eradicated by the immune system. It is possible that applying radiation therapy prior to checkpoint blockade will result in increased efficacy. Isolated cases have suggested an abscopal effect of such treatment, but this has yet to be confirmed in a randomized clinical trial (http://www.nejm.org/doi/full/10.1056/NEJMoa1112824). Recent data suggest that sustained benefit of radiation combined with CTLA4 blockade may also require PD-L1 blockade to reverse T cell exhaustion, and that radiation increases the diversity of the T cell receptor repertoire on intratumoral T cells (http://dx.doi.org/10.1038/nature14292). Oncolytivc viruses may also be used to increase immunogenicity, via the induction of an inflammatory response upon local injection of virus, leading to control of distant tumors by an increase in tumor infiltrating cytotoxic populations (http://stm.sciencemag.org/content/6/226/226ra32.long). For example, Talimogene laherparepvec (T-VEC) has shown promising data in Phase III clinical trials (http://jco.ascopubs.org/content/early/2015/06/16/JCO.2014.58.3377.long).

Clinical/Regulatory

What aspects of tumor heterogeneity actually matter in the clinic, and how can they be transformed into diagnostic strategies and treatment guidelines including biomarkers of response?

We need more information on the degree to which heterogeneity affects the clinical management of patients. We need more work to document the phylogeny and generate atlases or road maps for each tumor subtype. This will allow us to identify more confidently the trunk mutations for each subtype and begin to understand the branching properties. Although one could easily assume that between The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), and other consortia we have at our disposal the tumor genomic data needed to generate road

maps for each tumor subtype, this group felt that none of our existing tumor genome repositories are actually sufficient for this sort of analysis. This is because these programs were not designed to address the heterogeneity component of cancer, and because all platforms used characterized the tumors in 'bulk', giving results that average across all tumor clones. Although some bioinformatics tools have been developed to tease out the clonal data in these datasets, their inherent limitations still exist. What we need for each subtype is a minimum number of primary tumors; the minimum number will likely vary according to the tumor subtype and its inter-patient heterogeneity. Ideally we will obtain multiple regions from each tumor to capture spatial heterogeneity. To differentiate trunk mutations from subclones, we need to sequence each region deeply. Ideally epigenetic and other analyses will also be performed. Patients who donated their tumors must then be followed longitudinally and tissue--where practical--and blood collected at regular time points and subjected to more deep sequencing to follow the molecular changes. Clinical annotation of samples and phenotypic correlation is essential at each step. This sort of analysis should reveal a finite number of trunk (clonal) and tree (subclonal) mutations, which can inform about the signaling pathways involved, for each tumor type. Excitingly, some new studies such as TRACERx incorporate several of these design elements, albeit in a single tumor type (https://www.clinicaltrials.gov/ct2/show/NCT01888601?term=TRACERx&rank=1).

How can we maximize the extraction of molecular and clinical data that is sharable and likely to feed back to benefit patients?

We need new consortia composed of academic medical centers *and* industry partners *and* regulatory agencies. Prior to clinical sample collection or data generation, all stakeholders need to agree on a minimal set of metadata that need to be collected for each tumor in a format that enables sharing; characterization must be systematic and agnostic to the tumor subtype. Genomic, clinical and any other data--once collected--must be added within a time frame agreed upon at the outset to a repository identified by all as suitable. To maximize the extent of effective data sharing and minimize sharing limitations caused by differences in consent practices across institutes, municipalities and nations, new harmonized consent practices consisting either of universal consent forms, or a universal option for patients to waive all restrictions on global sharing of data--even data such as germline sequences which have the potential to reveal identity of patients and their relatives--are needed up front. Patients, for example through formation of new patient advocacy organizations, should be empowered to drive data sharing. Encouragingly, several of these considerations are being incorporated into new consortia such as Cancer Core Europe (http://www.ncbi.nlm.nih.gov/pubmed/?term=Eggermont+AM%2C+Caldas+C).

How can we 'drug' tumor heterogeneity?

Although combination drug studies are challenging, adaptive trial designs to test combinations of targeted therapies with chemotherapies and/or immunotherapies based on molecular information extracted from individual tumors will be needed. Whether these combinations are given simultaneously at the start of treatment, or

sequentially as new resistance or other subclonal mutations appear during longitudinal analysis of patient samples obtained through non-invasive methods, may vary depending on the appearance of the road map of each tumor subtype generated by the consortia mentioned above, and on the therapeutic window of each drug alone and in combination. Ideally we could always block druggable trunk mutations and then add drugs to block emerging subclones. To simplify development of combination therapies, drugs showing a high degree of tumor vs. normal tissue selectivity (e.g. those targeting a mutant but not wild type version of a tyrosine kinase) may be prioritized. To minimize legal and financial hurdles that prevent testing of combinations of different drugs from different companies, 'honest broker' approaches that negotiate these issues with companies (along the lines of the Cancer Research Institute Clinical Accelerator, Cancer Core Europe, and the NCI Cancer Therapy Evaluation Program (CTEP)) should be proactively incorporated into the consortia mentioned above. It is likely that tumour heterogeneity in the form of increased somatic mutational diversity in some cases represents an Achilles' Heel for tumours due to the increased likelihood of tumour neo-antigens being recognised as non-self by T cells.

Why do clinical trials fail and what is the clinical trial of the future?

The ideal clinical trial will incorporate patients whose tumors have been selected as likely to respond based on molecular markers of response that have been well validated in preclinical studies. However, trials in which a single agent is tested in a cohort with a matched biomarker do not provide information about the impact of heterogeneity, nor the longitudinal evolution of clonal or subclonal cells. The reason for lack of response, either in the cohort or at the individual level requires understanding of the spatial and longitudinal heterogeneity of the tumor. The ideal clinical trial will be dynamic--in real time--to molecular changes revealed by frequent characterization of tumor evolution in response to therapy. This characterization will require material from the primary tumor or metastases (not always accessible) or could be achieved by studying nucleic acids and/or cells in the blood, as emerging data suggests liquid biopsies are feasible (http://www.nature.com/nm/journal/vaop/ncurrent/full/nm.3870.html, http://www.nature.com/nm/journal/v21/n6/full/nm.3854.html, http://www.nature.com/nm/journal/v20/n5/full/nm.3519.html http://www.nejm.org/doi/full/10.1056/NEJMoa1213261). Imaging approaches may not have sufficient resolution, information content, or speed to reveal molecular changes indicative of emerging resistance to therapy, although new imaging modalities such as ^{13}C-based magnetic resonance spectroscopy might provide metabolic readouts of response (http://www.nature.com/nm/journal/v20/n1/full/nm.3416.html). Changes in clinical practice and regulatory procedures may be needed. For example, are we ready to conduct trials in which treatment is adapted based on changes in circulating free tumor DNA as a surrogate of progression? Similarly, in the scenario where a resistance mutation is detected in the blood of a patient by analysis of circulating free tumor DNA, would a clinician be comfortable discontinuing that targeted therapy and switching to a different targeted therapy, even if the patient's tumor remains stable or continues to shrink as revealed by imaging analysis?

Technology

What sources of heterogeneity can we measure and which are difficult to assess?

DNA Assessing the mutational landscape of tumors by high throughput DNA sequencing of bulk samples is the most mature of all the technologies used for the molecular characterization of tumor heterogeneity. Single nucleotide and structural variations with a high allele frequency can be robustly detected with the sequencing depth that can be routinely achieved in experimental as well as clinical settings. For comprehensive cataloging of mutations that occur with a frequency of less than 1-2%, the required sequencing depth for robust variant calling (400-500x) is still prohibitive for larger scale studies, but it can be expected that with the continuing development in sequencing technology this issue will be solved in the near future. A major advantage of DNA sequencing is that it is relatively robust towards sample treatment and high quality data can be obtained from most specimens, although accurate enumeration of subclonal tumor heterogeneity in archival samples that are generally formalin-fixed and paraffin embedded (FFPE) can be more challenging

Optimizing analysis pipelines for variant calling has been an intensive focus of research in recent years. A low false positive or false negative mutation detection rate has little effect in TCGA/ICGC-like cohort studies but may lead to artefactual differences between related mutation profiles, and cause critical misinterpretations of study results. What is currently missing is an independent systematic evaluation of the many pipelines currently used for mutation calling in cancer samples. In this context, a valuable community resource would be the availability of benchmarking reference specimens with defined clonal composition as assessed by a gold-standard. We expect the results from comparative evaluation such as ICGC-TCGA DREAM Genomic Mutation Calling Challenge will provide a good estimate of the relative performance of the different methods for the processing of whole genome datasets, but further investigations are likely to be needed to benchmark tools for calling of subclonal mutations and the estimation of allele frequencies. Conservative approaches, such as >60x coverage thresholds and mutation filtering using multiple germlines, are recommended when determining the amount of heterogeneity between tumor samples. A relatively uncharted area is the development of metrics that quantify similarity and differences between samples from the same clonal origin, which includes multiple biopsies from the same tumor, pre- and post-treatment samples from the same patient, or comparison of tumor samples and (xenotransplanted) model systems. .

In the context of heterogeneity the more recent development of single cell genome sequencing is very exciting as it allows not just for an estimation of the frequency of individual mutated alleles in a cancer sample, but also for the determination of co-occurring or mutually exclusive alterations. Currently, the main limitations for single cell genome sequencing are the relatively low throughput, and the partial genome coverage and high error rate when using amplification methods such as multiple displacement amplification (MDA). Also, algorithms for calling SNV and structural variations have not been optimized for single cell data yet.

RNA As with DNA, sequencing is now the method of choice for investigating the RNA composition of tumors. In contrast to DNA data, it is difficult to learn much about the heterogeneity of bulk samples from RNA-seq data, beyond what can be done by sequencing samples from different parts of the tumor. The tumor microenvironment may represent as much as 90% of some tumor samples and contributes proportionally to the RNA pool, which affects measures of heterogeneity and the overall resulting transcriptional profile. Computational deconvolution of different expression components within a single sample are able to distinguish between cells from different lineages, but have limited applicability in samples with low transcriptional diversity.

Single cell RNA-seq is a robust technology that, with the recently developed Drop-seq methods (http://www.ncbi.nlm.nih.gov/pubmed/?term=26000488, http://www.ncbi.nlm.nih.gov/pubmed/?term=26000487), can analyze tens of thousands of cells simultaneously in a cost-effective and efficient manner. That said, sensitivity for lowly expressed genes still needs to be increased for all RNA-seq protocols and we need better methods for controlling amplification biases and technical noise. Optimized analysis tools for single cell RNA-seq methods are also being developed, but a thorough comparative benchmarking of these tools has also been lacking. For single cell methods, obtaining full-length RNA molecule sequences or information about RNA modifications still remains challenging.

As the transcriptome is highly dynamic, sample handling is a critical hurdle in the acquisition of quality transcriptomes. Issues to consider include how quickly the sample is processed or frozen following its extraction from the patient and even more importantly the protocol through which the cells are disassociated from solid tumors. The availability of fresh or rapidly frozen samples is essential, as FPPE samples, although they can be processed for RNA-seq, are unlikely to provide a reasonable picture of cancer cell states.

Proteins
Compared to nucleic acid techniques the investigation of proteins lags behind, especially in terms of sensitivity and comprehensiveness. Although it is possible to get a complete picture of the protein content of a sample using mass spectrometry-based proteomics the comparatively large amount of material needed makes proteome-wide experiments on cancer samples unfeasible in most instances at the moment. Antibody-body based techniques are the method of choice when sample material is limited, but they are limited by the availability of high quality antibodies and throughput. On the single cell level, technologies such as FACS or CyTOF allow the investigation of currently up to about 17 (FACS) or 45 (CyTOF) proteins per cells with very high throughput. Future development of CyTOF technology might increase the number of proteins that can be monitored, but no technology that can provide a truly comprehensive protein atlas for single cells is on the horizon.

The protein content is less dynamic than the transcriptome or the epigenome, which reduces the requirement for sample freshness. However, the phosphor-proteome critical for the understanding of cancer signaling is even more sensitive and rapidly changing than the transcriptome. With regard to FFPE samples, they can be processed for proteomics experiments, but not for CyTOF or FACS.

Epigenetic marks

As discussed in the 'Beyond the Genome' section, the epigenetic features of chromatin includuing histone modification, DNA methylation and DNA accessibility provide information about both the cell state and the evolutionary history of a tumor. Robust technologies have been developed to provide genome-wide maps of most marks. For histone marks, Chip-seq methods have been developed that can reliably applied to very small samples (1,000 cells or less). Several techniques are currently routinely used to assess methylation levels. The most comprehensive picture can be obtained from whole genome bisulfite sequencing (WGBS), but precipitation techniques (methylated DNA immunoprecipitation sequencing and methylated DNA binding domain sequencing) or reduced representation bisulfite sequencing (RRBS) are also still commonly used. WGBS can be applied also be applied to small samples, but the DNA damaging effects of the bisulfite treatment limit genome coverage. The development of alternative chemistries that are less harsh will help reduce experimental artifacts. The Illumina Infinium 450k BeadChip platform provides an array-like alternative that has been found to provide acceptable DNA methylation profiles, even with FFPE samples. Assays for the various oxidized forms of 5-Methylcytosine have been developed, but not thoroughly validated in terms of reproducibility and sensitivity. DNA accessibility and nucleosome positioning can also be readily measured, most commonly using DNAase I-based assays and more recently ATAC-seq for bulk samples. As there are a large number of epigenetic modifications that one would like to measure in a given sample, a big need in the field is the development of multiplexing strategies that allow measuring many marks at the same time in the same sample.

Although some estimation of cellular heterogeneity can be obtained from bulk experiments in the case of DNA methylation, single cell assays would provide advantages in terms of capturing the amount of heterogeneity, although they would be limited by throughput. Single cell assays for histone marks have yet to be developed, but given the rapid development in the field, single cell ChIP-seq assays can be expected in the near future.

With regard to sample preparation, the epigenome, like the transcriptome, is highly dynamic and sensitive to changes in the environment. As such fresh or rapidly frozen samples is essential.

Multiplexing

A complete picture of a cell state will often require measuring differ parameters in the same single cell. Although it is usually possible to perform multiple assays on a bulk

sample, this is in many cases not possible with single cell measurements (with the exception of RNA and DNA-seq

How can we assess spatial organization of tumors?
Traditionally, when a spatial resolution higher than what can achieved by multiple biopsies is desired, assessment of spatial heterogeneity in tissue samples has been limited to microscopy-based methods. For example, immunofluorescence and fluorescence in situ hybridization can localize proteins, RNAs and DNA mutations in tissue slices with high sensitivity, potentially down to the single molecule level. In practice both methods suffer from difficulties in quantifying expression levels and in comparing results within and between different samples due to variable background and target accessibility. These techniques are also very low throughput and the number of mRNAs or proteins that can be imaged simultaneously are currently limited to a handful using standard technology. Imaging site-specific epigenetic modifications is currently not routinely done, although at least one method has been developed to visualize histone modifications in fixed tissues.

Excitingly, new technologies have emerged in recent years that promise to revolutionize our ability to assess spatial heterogeneity of protein and RNA expression. For proteins, CyTOF has been developed into an imaging tool that can image the localization of currently up to 32 proteins (and potentially up to 100) with subcellular resolution. For RNA, in situ sequencing methods can provide information about the RNA content of individual cells in the context of a fixed tissue. The practical applications of these technologies are still in their infancy and a thorough benchmarking of reproducibility and sensitivity is yet to be done. Also throughput of these new technologies still seems to be severely limited at the moment.

With all these new technologies on the rise, the critical bottleneck has become the development of computational methods to analyze each technology individually, integrate information from the different technologies together and connect these data to prognostic and actionable clinical value.

What non- or minimally invasive technologies can be used to obtain information about tumor heterogeneity?

Currently the best sources of information about the molecular makeup of a cancer that can be obtained without biopsying the tumor itself are found in the blood. Cell-free DNA and circulating tumor cells are especially rich sources of information. Currently we still lack a sufficient number of high quality studies to assess how well data obtained from these blood-borne biomarkers reflects the tumor itself, although a number of recent studies have already highlighted the power of this approach for the noninvasive characterization of tumor heterogeneity in carcinomas of the colon, breast, and lung (http://www.nature.com/nm/journal/vaop/ncurrent/full/nm.3870.html, http://www.nature.com/nm/journal/v21/n6/full/nm.3854.html,

http://www.nature.com/nm/journal/v20/n5/full/nm.3519.html
http://www.nejm.org/doi/full/10.1056/NEJMoa1213261). It is also unclear whether primary tumors or metastases contribute more to the pool of circulating cancer material. It seems clear however that even if it is determined that circulating material faithfully reflects the tumor itself, more efficient ways of isolating the cells and nucleic acids from the blood, and data analysis tools that can more faithfully reconstruct the parent tumor need to be developed.

Although *in vivo* imaging technologies are currently not able to provide many insights into intratumor heterogeneity in patients due to resolution and labeling issues, some work indicates that heterogeneity in PET and MRI imaging is clinically predictive of response (http://www.ncbi.nlm.nih.gov/pubmed/25421725). Image-guided biopsies may also make an important contribution to the analysis of genome based intratumoral heterogeneity by providing the spatial context to relate different regions.

How do we benchmark and validate the methods used for assessing heterogeneity?
Validation of the accuracy and robustness of the various assays discussed above will require the development of gold standard samples that are readily available and can be recreated in reproducible manner by individual labs. For some data types, such as DNA sequence mutations, simple mixtures of cell lines will be sufficient, while others such as the epigenome or RNA expression are too sensitive to environmental changes and will require test samples with more intrinsic control of biological variation, such as 'spike in' standards. For each assay a set of quality control metrics will have to be agreed upon that can be used to assess the performance of improved methods and that each investigator can apply to his or her own experiments.
One should also keep in mind that the levels of accuracy needed for understanding biology and for informing clinical decision making may differ, and these should be investigated separately.
What computational methods do we need to integrate the different types of data?

Although we have a wealth of methods that can be applied to existing datasets, and methods to simulate tumor dynamics, we should apply our current knowledge to envision a more integrated experimental pipeline that can extract the most information from patients. Looking forward, several considerations should be taken into account when designing both individual lab experiments as well as cooperative projects for high-throughput data acquisition and analysis. At the outset, a critical point to bring these emerging technologies to the clinic is improving sample handling from collection, through processing and into proper allocation towards different assays.

Other questions that need to be carefully considered in an experimental design that aims to study tumor heterogeneity and evolution include: 1) Which tumor type should be chosen? Tumors need to be relatively large to provide enough material for the various assays, need to be readily resectable and need to progress sufficiently quickly to make a reasonable timeline possible. 2) When should a tumor be sampled and should it

be sampled repeatedly? 3) Which parts of the tumor should be analyzed? Should we include information about metastatic sites?

Ideally, a coordinated effort to produce these kind of samples would generate gold standard datasets from a large number of patients with well annotated clinical history and comprehensive tumor imaging. Each patient's tumor could then be analyzed with a wide array of experimental techniques that provide information about different levels of heterogeneity. The core group of methods should include multi-focal bulk and single cell DNA sequencing, single cell RNA-seq, multi-focal bulk and single cell mapping of epigenetic marks and single cell CyTOF-based analysis of candidate marker proteins. Both tumor and micro-environment, including tumor infiltrating leukocytes would ideally be assayed. These data could be complemented with other assays that can measure spatial heterogeneity or investigate cell free DNA or circulating tumor cells. These assays would provide detailed protein, genome and RNA maps, but in order to be able to reconstruct patient-specific regulatory networks, algorithms will need to be substantially improved.

This wealth of data should then be made available to the research community to develop new methods that analyze and integrate the information provided by different assays, in order to predict disease outcome and therapeutic success, so we can continue to gain insight into the importance and impact of cellular heterogeneity.

Closing remarks

One take-home message from this meeting was that the phenomenon of tumor heterogeneity is likely to influence--for some time to come--all aspects of cancer research, from how we perceive tumor biology, to how we develop techniques to study tumors, to how we treat patients. This conference was unique in its goal of identifying questions rather than answers, and we hope that this description of the 'known unknowns' identified by this small group of experts sparks research and collaboration in the community at large.