International Cancer Genome Consortium
International Facts on Cancer

- In 2007 over 12 million new cases were diagnosed across the planet and approximately 7.6 million cancer deaths occurred.

- In 2050, these numbers will rise to an expected 27 million new cases and 17.5 million cancer deaths if our ability to prevent, diagnose and treat cancer does not improve.

Cancer
A Disease of the Genome

Challenge in Treating Cancer:

- Every tumor is different
- Every cancer patient is different
Goals of
Cancer Genome Research

- Identify changes in the genomes of tumors that drive cancer progression
- Identify new targets for therapy
- Select drugs based on the genomics of the tumor
Large-Scale Studies of Cancer Genomes

- Johns Hopkins
  - > 18,000 genes analyzed for mutations
  - 11 breast and 11 colon tumors

- Wellcome Trust Sanger Institute
  - 518 genes analyzed for mutations
  - 210 tumors of various types

- TCGA (NIH)
  - Multiple technologies
  - brain (glioblastoma multiforme), lung (squamous carcinoma), and ovarian (serous cystadenocarcinoma).
Lessons learned

- Heterogeneity within and across tumor types
- High rate of abnormalities (driver vs passenger)
- Sample quality matters
‘Next Generation’ sequencing instruments are providing new opportunities for comprehensive analyses of cancer genomes

- Capacity greater than one Gigabase per run
- Drastic decrease in costs per genome
- Applications: DNA, RNA, chromatin (i.e. epigenome)
International Cancer Genomics Strategy Meeting
October 1–2, 2007  Toronto (Canada)

22 countries represented
120 participants

34 Genome or Cancer Center Directors
24 Representatives from funding
62 Scientists selected to represent ethics, technologies, statistics, informatics, pathology, clinical oncology and cancer biology
Purpose of the Toronto meeting

- Exchange knowledge and discuss opportunities that could lead to a consortium that would generate a comprehensive atlas of genomic abnormalities in cancer.
Major issues addressed in defining consortium
ICGC Goal

➢ To obtain a comprehensive description of genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes which are of clinical and societal importance across the globe.
Rationale for an international consortium

- The scope is huge, such that no country can do it all
- Independent cancer genome initiatives could lead to relative duplication of effort for common and easy to acquire tumor samples, and incomplete studies for many forms of cancer
- Lack of standardization, and different quality measures across studies could decrease the opportunities to merge datasets, increase power, and detect additional targets
- The spectrum of many cancers is known to vary across the world for many tumor types, because of environmental, genetic and other causes
- An international consortium will accelerate the dissemination of genomic and analytical methods across participating sites, and into the user community
Basic Tenets

- The level of organization is at the specific cancer type or subtype.
- A particular cancer may be investigated by an individual research lab/center or by a collaborative research group, across jurisdictions.
- The key to inclusion of a project in the ICGC is that it should take a comprehensive, genome-wide approach to the analysis of that tumor type (or sub-type).
- The ICGC is open to many organizations willing mount a comprehensive analysis of at least one cancer type or subtype, and that agree to carry out their efforts according to ICGC policies.
Incorporate lessons from pilot projects

- This will be HARD!
- Sample collection can easily be rate limiting
- Much of sample collection will need to be prospective
- Technology landscape is rapidly changing
- Quality assessment is critical
- Truly exciting insights *can* be generated from this kind of comprehensive analysis of cancer
- Creative funding mechanisms will need to be worked out
Interim Organizational Structure

**Executive Committee (EXEC)**
- Initial members (funders): Australia (Observer Status), Canada, China, France, India, Japan (RIKEN; National Cancer Center), Singapore, the UK (The Wellcome Trust; Wellcome Sanger Institute), and the US (NCI and NHGRI), and the European Commission (Observer Status)
- Secretariat: Ontario Institute for Cancer Research

**Scientific Planning Committee (SPC)**

**Working Groups**
- Clinical and Pathology Issues
- Informed Consent and Privacy Protection
- Quality Standards of Samples
- Sample Size/Study Design
- Genome Analyses
- Data Management/Databases & Coordination
International Cancer Genome Consortium

Goals, Structure, Policies & Guidelines

www.icgc.org
ICGC Consent and Privacy Protection Policies

- ICGC membership implies compliance with Core Bioethical Elements for samples used in ICGC Cancer Projects

  ICGC acknowledges that the informed consent process used by ICGC members will necessarily differ according to local, socio-cultural and legal requirements

- To minimize the risk of patient/individual identification, the ICGC has established the policy that datasets be organized into two categories, open and controlled-access.
## Data Releases

<table>
<thead>
<tr>
<th>ICGC Open Access Datasets</th>
<th>ICGC Controlled Access Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Cancer Pathology</td>
<td>- Detailed Phenotype and Outcome Data</td>
</tr>
<tr>
<td>- Histologic type or subtype</td>
<td>- Patient demography</td>
</tr>
<tr>
<td>- Histologic nuclear grade</td>
<td>- Risk factors</td>
</tr>
<tr>
<td>- Patient/Person</td>
<td>- Examination</td>
</tr>
<tr>
<td>- Gender</td>
<td>- Surgery/Drugs/Radiation</td>
</tr>
<tr>
<td>- Age range</td>
<td>- Sample/Slide</td>
</tr>
<tr>
<td>- Gene Expression (normalized)</td>
<td>- Specific histological features</td>
</tr>
<tr>
<td>- DNA methylation</td>
<td>- Protocol</td>
</tr>
<tr>
<td>- Genotype frequencies</td>
<td>- Analyte/Aliquot</td>
</tr>
<tr>
<td>- Computed Copy Number and Loss of Heterozygosity</td>
<td>- Gene Expression (probe-level data)</td>
</tr>
<tr>
<td>- Newly discovered somatic variants</td>
<td>- Raw genotype calls</td>
</tr>
<tr>
<td></td>
<td>- Gene-sample identifier links</td>
</tr>
<tr>
<td></td>
<td>- Genome sequence files</td>
</tr>
</tbody>
</table>
ICGC Data Release Policies

- The members of the International Cancer Genome Consortium (ICGC) are committed to the principle of rapid data release to the scientific community.

- The individual research groups in the ICGC are free to publish the results of their own efforts in independent publications at any time.

  Investigators outside of the ICGC are free to use data generated by ICGC members, either en masse or specific subsets, but are asked to follow the guidelines developed the “Ft. Lauderdale principles”

  http://tinyurl.com/3klwx4
ICGC Intellectual Property Policy

All ICGC members agree not to make claims to possible IP derived from primary data (including somatic mutations) and to not pursue IP protections that would prevent or block access to or use of any element of ICGC data or conclusions drawn directly from those data.

Note: Users of the data (including Consortium members) may elect to perform further research that would add intellectual and resource capital to ICGC data and elect to exercise their IP rights on these downstream discoveries. However, ICGC participants and other data users are expected to implement licensing policies that do not obstruct further research: (http://tinyurl.com/4rslvy).
Tumor Types and Subtypes

- The ICGC aims to study cancers of all major organ systems
- Studies will cover adult and childhood / adolescent cancers
- Guidelines have been developed for ICGC participants for the selection of Cancer Genome Projects
Policies Regarding Quality Standards of Samples

- A committee of clinical and pathology experts (with representation from different institutions) will be needed to draft and oversee the specific guidelines that will apply for every tumor type or sub-type.
- Tumor types should be defined using the existing international standards of the WHO (including ICD-10 and ICD-O). If novel molecular subtypes are studied, these should be defined with sufficient detail.
- All samples will have to be reviewed by two or more reference pathologists.
- Patient-matched control samples, representative for the germline genome, are mandatory to discern “somatic” from “inherited” mutations.
Policy Regarding Study Design and Statistical Issues

- Every cancer genome project should state a clear rationale for its choice of sample size, in terms of the desired sensitivity to detect mutations. The target number of 500 samples per tumor type/subtype is set as a minimum, pending further information to be provided by ICGC members proposing to tackle specific cancer types/subtypes.
Genome Analyses

- **Mandatory:** Genomic DNA analyses of tumors (and matching control DNA) are core elements of the project.

- **Complementary (Recommended):** Additional studies of DNA methylation and RNA expression are recommended on the same samples that are used to find somatic mutations.

- **Optional:**
  - Proteomic analyses
  - Metabolomic analyses
  - Immunohistochemical analyses
Genome Analyses

- Whole genome shotgun analyses (long-term goal)

- Interim, large-scale, catalogues of somatic mutations
  - Sequencing of all coding exons and other genomic regions of particular biological interest for point mutations.
  - Analysis of low genome coverage of paired-end reads for rearrangements.
  - Genotyping arrays, to detect copy number changes, LOH and breakpoint information.

- Analyses of DNA Methylation

- Expression Analyses: protein coding genes, non-coding RNAs, notably microRNAs.
Similar to other large-scale genome projects, the ICGC will require a Data Coordination Center (DCC) that will:

• provide secure and reliable mechanisms for the sequencing centers, biorepositories, histopathology groups, and other ICGC participants to upload their data;
• track data sets as they are uploaded and processed;
• allow regular audit of the project in order to provide high-level snapshots of the Consortium's status;
• enable the distribution of the data to the long-lived public repositories of genome-scale data;
• provide essential meta-data to each public repository that will allow the data to be understandable;
• facilitate the integration of the data with other public resources, by using widely-accepted ontologies, file formats and data models;
• manage an ICGC data portal that provides researchers with access to the contents of all franchise databases and provides project-wide search and retrieval services.
Genome projects enable research into the complex nature of human disease

- Human Genome Project
- The HapMap Project
- The Cancer Genome Atlas
- ICGC Cancer Genome Projects

The most important contribution to science of these large-scale projects is the generation and transfer of resources, databases and technologies to the scientific community.
The International Cancer Genome Consortium can be the hub of the wheel, but it’s not all of cancer research!
ICGC will be an *enduring* legacy

A comprehensive catalog of somatic changes in the major cancers will be a powerful driver for cancer research and clinical practice for decades.
ICGC will be an *enduring* legacy

A comprehensive catalog of somatic changes in the major cancers will be a powerful driver for cancer research and clinical practice for decades

Early clinical benefits will be stratification of tumors to allow better prediction of prognosis and response to therapy
ICGC will be an *enduring* legacy

A comprehensive catalog of somatic changes in the major cancers will be a powerful driver for cancer research and clinical practice for decades.

Early clinical benefits will be stratification of tumors to allow better prediction of prognosis and response to therapy.

Longer term benefits will be development of new and more effective targeted therapies.
Acknowledgements

**INTERIM ICGC EXECUTIVE (EXEC)**

Warwick Anderson, National Health and Medical Research Council, Australia (Observer Status)
Cindy Bell and Karen Kennedy, Genome Canada, Canada (Observer Status)
Tom Hudson, Ontario Institute for Cancer Research, Canada
Henry Yang, Chinese Cancer Genome Consortium, China
Jacques Remacle, Patrik Kolar and Iiro Eerola, European Commission (Observer Status)
M.K. Bhan and T.S. Rao, Ministry of Science & Technology, Department of Biotechnology, India
Edison Liu, Genome Institute of Singapore, Singapore
Alan Schafer and Michael Stratton, The Wellcome Trust: Wellcome Trust Sanger Institute, United Kingdom
Anna Barker and Daniela Gerhard, National Cancer Institute, United States
Francis Collins, Jane Peterson, Mark Guyer and Brad Czenberger, National Human Genome Research Institute, United States
# Acknowledgements

<table>
<thead>
<tr>
<th>ICGC Scientific Planning Committee (SPC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warwick Anderson</td>
</tr>
<tr>
<td>Anna Barker</td>
</tr>
<tr>
<td>Cindy Bell</td>
</tr>
<tr>
<td>Martin Bobrow</td>
</tr>
<tr>
<td>Wylie Burke</td>
</tr>
<tr>
<td>Francis Collins</td>
</tr>
<tr>
<td>Carolyn Compton</td>
</tr>
</tbody>
</table>
## Acknowledgements

### International Cancer Genome Consortium (ICGC) Working Groups

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lynda Chin</td>
<td>Peter Lichter</td>
<td>Mike Stratton</td>
<td>Baitha Knoppers</td>
<td>Eric Lander</td>
<td>Lincoln Stein</td>
<td>Mark Guyer</td>
</tr>
<tr>
<td>Jean-Yves Blay</td>
<td>Carolyn, Compton, Andy Fuhrer</td>
<td>Olli Kalionieni</td>
<td>Martin Bobrow</td>
<td>Fon DePinho</td>
<td>Cameron Brennan</td>
<td>Daniela Gerhard, Karen Kennedy, Brad Ozenberger</td>
</tr>
<tr>
<td>William Dalton</td>
<td>Youyoung Lu, Miguel Angel Pris</td>
<td>Ed Liu</td>
<td>Wylie Burke</td>
<td>Doug Easton</td>
<td>Arul Chinnaiyan</td>
<td>signage, John Golden, J Gourishankar</td>
</tr>
<tr>
<td>Tony Green</td>
<td></td>
<td>Marco Marra</td>
<td>Kazuto Kimura</td>
<td>Gaddy Getz</td>
<td>Peter Good</td>
<td>David Haussler, David Housman, Tim Hubbard, Subha Madhavan</td>
</tr>
<tr>
<td>Stan Hamilton</td>
<td></td>
<td>John McPherson</td>
<td>Karen Kennedy</td>
<td>Fartha P. Majumder</td>
<td>Joe Gray</td>
<td>Paul Spellman</td>
</tr>
<tr>
<td>Timothy Ley</td>
<td></td>
<td>Brad Ozenberger</td>
<td>Brad Ozenberger</td>
<td></td>
<td>Jai DiGiovine</td>
<td></td>
</tr>
<tr>
<td>Ed Liu</td>
<td></td>
<td>Henk Stunnenberg</td>
<td>Daniel Tan</td>
<td></td>
<td>J Gourishankar</td>
<td></td>
</tr>
<tr>
<td>Paul Mischel</td>
<td></td>
<td>Daniel Tan</td>
<td>Susan Wallace</td>
<td></td>
<td>Jai DiGiovine</td>
<td></td>
</tr>
<tr>
<td>Kenneth Penta</td>
<td></td>
<td>Brandon Wainwright</td>
<td></td>
<td></td>
<td>Jai DiGiovine</td>
<td></td>
</tr>
<tr>
<td>Rajiv Sarin</td>
<td></td>
<td>Rick Wilson</td>
<td>Henry Yang</td>
<td></td>
<td>Jai DiGiovine</td>
<td></td>
</tr>
<tr>
<td>Daniel Tan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jai DiGiovine</td>
<td></td>
</tr>
</tbody>
</table>

ICGC