

International Human Epigenome Consortium

Dr Anthony Phillips

3rd International Traumatic Brain Injury Research (InTBIR) Meeting

June 28-29, 2014

At a Glance



Generate reference maps of human epigenomes for key cellular states relevant to health and disease.

→ **Ambitious goal to complete 1000 epigenomes**

Rapid Data Release, Data Coordination, Archiving

Environment

Disease

Aging

Translation into Improved Human Health

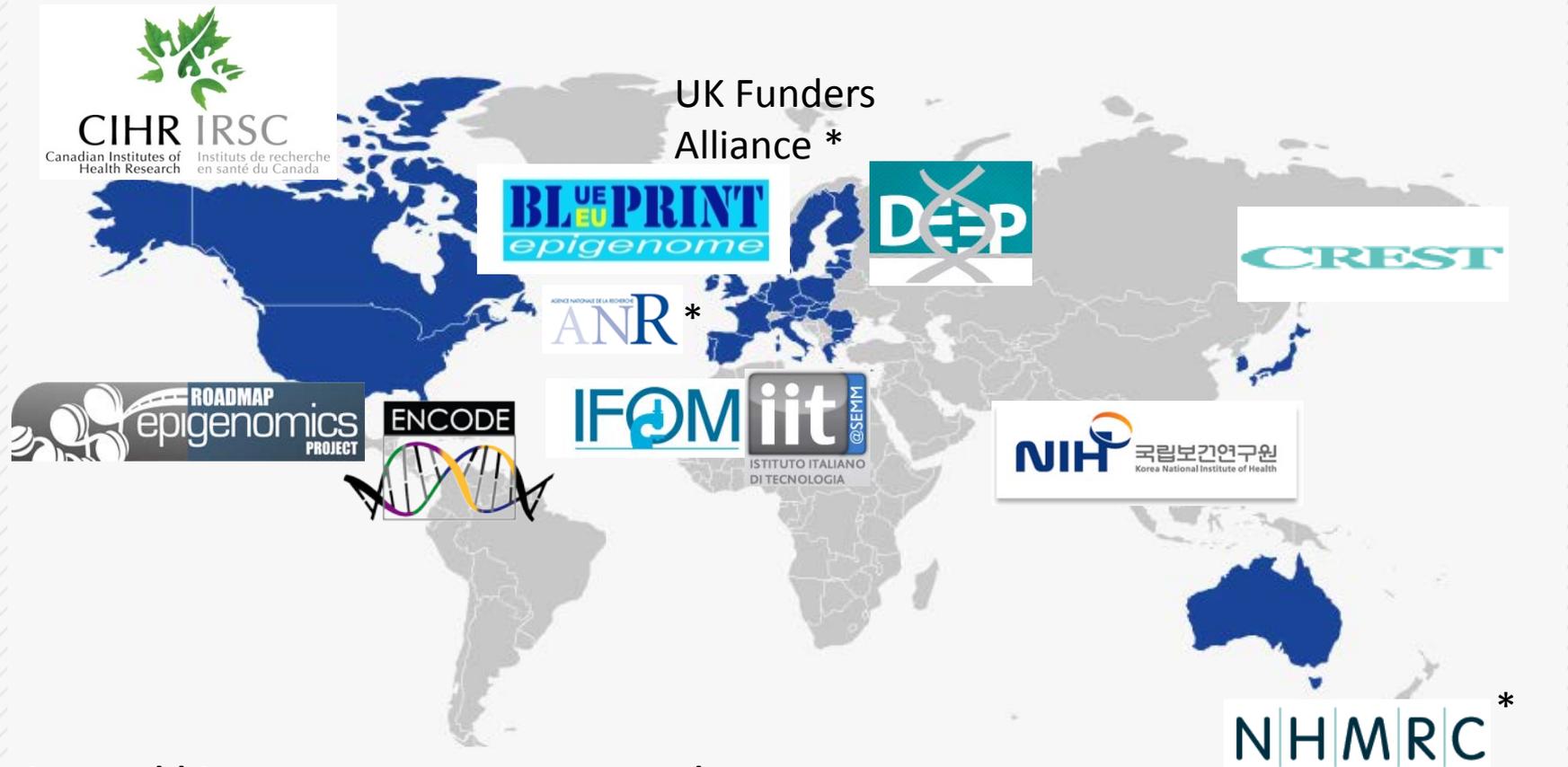
Mechanism

Prevention

Diagnosis

Therapy

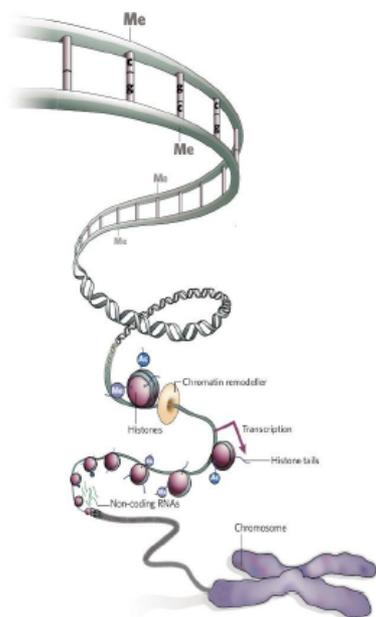
IHEC Member Countries (as of June 2014)



<http://ihec-epigenomes.org/>

GOALS, STRUCTURE, POLICIES & GUIDELINES

January 10, 2013



- Official IHEC policy document first published in December 2011
- Aligned with other international genomics consortia (e.g. ICGC)
- Last revised January 2013
- Currently under review with revised version expected before year-end 2014

Sample Donation Policy (as of January 2013)



Policy Statement: Core Bioethical Elements (to be conveyed to sample donors) include:

- Data derived from the samples collected and data generated by the IHEC members will be made accessible to the IHEC members and other international researchers through either an open or a controlled access database under terms and conditions that will maximize participant confidentiality
 - Those accessing data and samples will be required to affirm that they will not attempt to re-identify samples donors.
 - There is a risk of being identified from data available on the databases
 - Once the data is placed in open databases, it cannot later be withdrawn
 - In controlled access databases, the links to (local) data that can identify an individual will be destroyed upon withdrawal.
 - IHEC members agree not to make claims to possible intellectual properties derived from primary data
-

Data Access and Patient Protection Policy (as of January 2013)



Policy Statement: To minimize the risk of patient/individual identification, IHEC has established the policy that datasets be organized into two categories, open and controlled-access. Table 1 includes a list of data elements and the data access category within which they will be available.

IHEC Open Access Datasets	IHEC Controlled Access Datasets
<ul style="list-style-type: none">- Disease Pathology<ul style="list-style-type: none">o Histological type or subtype- General Patient/Person Information<ul style="list-style-type: none">o Gendero Age range- Gene expression (normalised)- Epigenetic mapping data<ul style="list-style-type: none">o Methylation markso Histone modification positionso Small RNA levels- Genotype frequencies	<ul style="list-style-type: none">- Detailed Phenotype and Clinical Data- Gene expression (probe-level data)- Raw genotype data- Gene-sample identifier links- Genome sequence files

Data Access and Patient Protection Policy (revisions under discussion)



- Current Open Access datasets have been carefully considered and will continue to be monitored by IHEC
 - Open Access datasets contain data that cannot, *at present, with reasonable efforts*, be aggregated to generate a dataset unique to an individual.
 - The list of epigenome and clinical data in Open and Controlled Access categories is currently under review, and *will be periodically revised by IHEC to reflect the continually evolving fields of epigenetics, bioinformatics, and to comply with ethics and privacy policies and regulations.*
-

IHEC Data Release Policy (as of January 2013)



Policy Statement: The members of the International Human Epigenome Consortium (IHEC) are committed to the principles of rapid data release to the scientific community

- IHEC agrees to follow the recommendations of the Toronto (2009) conference that was sponsored by the NIH, Wellcome Trust, EU Commission and Genome Canada.
 - **Funding agencies** agree to facilitate the specification of data-release policies
 - **Data producers** agree to enable analyses of their data in coordination with central or regional Data Coordination Centers
 - **Data analysts/users** agree to freely analyze released prepublication data and act responsibly in publishing analyses of those data
-

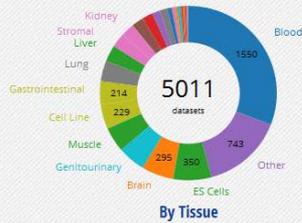
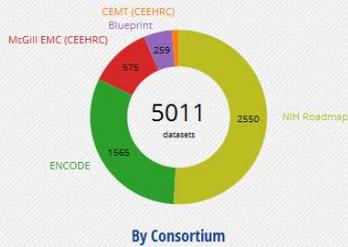
Data Sharing: IHEC Data Portal



Home Data Grid Download Genome Browser IHEC Main Site

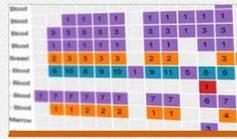
Welcome to the IHEC Data Portal

IHEC makes available comprehensive sets of reference epigenomes relevant to health and disease. You may view, search and download the data already released by the different IHEC-associated projects via the IHEC Data Portal.



<http://epigenomesportal.ca/ihec/>

Data Grid



The Data Grid offers a way to easily visualize available datasets, and to generate dynamic track hubs.

Go

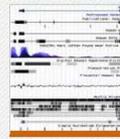
Downloads



The Download section provides a way to navigate through all datasets for direct download.

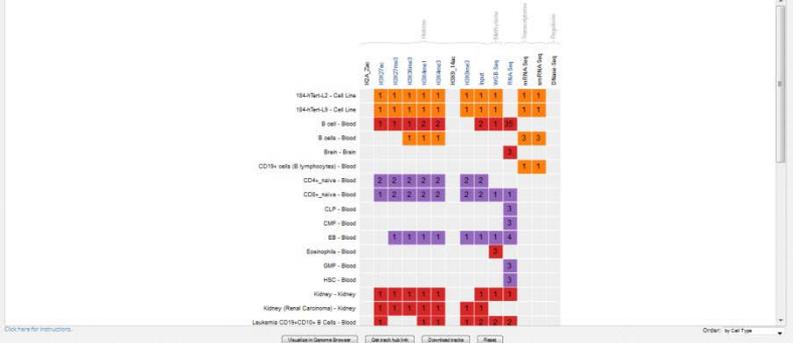
Go

Genome Bro



A direct access to our UCSC preloaded track hub of all av

Go



Datasets

Track Hubs <input type="checkbox"/> Consortium <input checked="" type="checkbox"/> McGill EMC (CEEHRC) 575 <input checked="" type="checkbox"/> CEMT (CEEHRC) 62 <input checked="" type="checkbox"/> Blueprint 259 <input checked="" type="checkbox"/> ENCODE 1565	Tissues <input checked="" type="checkbox"/> Adrenal <input checked="" type="checkbox"/> Blood <input checked="" type="checkbox"/> Bone <input checked="" type="checkbox"/> Bone Marrow	Assay Categories <input checked="" type="checkbox"/> Histone <input checked="" type="checkbox"/> Methylome <input checked="" type="checkbox"/> Transcriptome <input checked="" type="checkbox"/> Regulome	Overview <input checked="" type="checkbox"/> Consortium <input checked="" type="checkbox"/> Tissue <input checked="" type="checkbox"/> Assay Category
---	---	--	---



Contact Us

Data Sharing: IHEC Data Portal



- IHEC Data Portal can be used to *view, search and download* all Open Access data released by the IHEC-associated projects
- Portal developed and maintained by one of the CIHR-funded CEEHRC Epigenomic Platform Data Coordinations Centres (McGill)
- Takes advantage of Compute Canada high-performance computing cluster to manage large volume of data associated with all IHEC reference epigenomes
- Preliminary data release June 2, 2014 (5011 datasets), next release planned for September 2014

<http://epigenomesportal.ca/ihec/>



Canada



Canadian contribution to IHEC

Canadian Epigenetics, Environment and Health Research Consortium (CEEHRC)

Canadian Epigenetics, Environment and Health Research Consortium (CEEHRC)



Goal: To position Canada for the rapid translation of epigenetic discoveries into diagnostic procedures and the eventual development of new therapeutics that can improve human health

Funding Opportunities - \$50M initiative

Platform Centres \$15M

Catalyst
\$1.4M

Fellowships (multiple launches)

Team Grants \$10.1M

Canada-Japan Teams \$13.2M

Consortium \$2M

2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020



Communication & Outreach

- Café Scientifiques
- Journalist Workshop
- Best Brains Exchange

CEEHRC Platform Centres: 200 Reference Epigenomes Planned



CEEHRC PLATFORM

Canadian Epigenetics, Environment and Health Research Consortium Platform

Centre for Epigenomic Mapping Technologies (CEMT) - Martin Hirst, Marco Marra and Steven Jones

- Established Cell Groups, Biospecimen Core, Library Core, Sequencing Core and Informatics Core – with clear process flows
- Focusing on a number of cancer and stem cell types, including peripheral and lymphoid cells, hematopoietic cells, iPSCs, as well as thyroid cells and mammary, endometrial, fallopian and colonic epithelium.
- CEMT Data Access: <http://www.epigenomes.ca/>



- **Mark Lathrop, Tomi Pastinen, Guillaume Bourque, and Alan Evans**
- Focus on mapping population variation in human epigenomes
- Emphasis on longitudinal studies, population sampling in selected tissues, disease focus - especially autoimmune, inflammatory, metabolic, and neuropsychiatric
- McGill Dynamic Interactive Grid View (DIG-V): <http://epigenomesportal.ca/>