The Human Proteome Project: Chromosome 6, a Canadian Partnership

January 2011
Chromosome 6: overview

**Structure and organization**
- Structure and overview
- Protein coding and regulatory genes

**Role in human diseases**
- Chromosomal abnormalities
- Sequence variations
- Alteration in gene expression

**Teams and technologies**
- Genomics and proteomics teams
- Disease and mechanism-oriented teams
- Implications for health care transition
Chromosome 6: structure & organization

Overview

- The seventh of the 24 human chromosomes to be completed - joining chromosomes 7, 14, 20, 21, 22 and Y. Chromosome 6 is the largest to be fully analysed to date.

- Gene density is 9.2 genes per Mb, similar to Ch 7, 14, 20 and 22. MHC 43 genes per Mb, high density, high polymorphism, high linkage disequilibrium

- 2.2% of sequence occupied by exons; mean 281 bp, maximum length 9,114 bp (ZNF451), maximum number 101 (BPAG1). Largest gene PARK2 (1.4 Mb, 12 exons, 6q24)

- Chromosome 6 includes more than 130 genes that cause, predispose to or protect against certain diseases, with implications for cancer, heart disease, immune and inflammatory disorders and mental health.
Chromosome 6: structure & organization

**Gene content and type**

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>170,899,992 bps</td>
</tr>
<tr>
<td>Protein Coding Genes</td>
<td>1132</td>
</tr>
<tr>
<td>miRNA Genes</td>
<td>17</td>
</tr>
<tr>
<td>miRNA Pseudo Genes</td>
<td>4</td>
</tr>
<tr>
<td>Mouse miRNA Pseudo Genes</td>
<td>1</td>
</tr>
<tr>
<td>snoRNA Genes</td>
<td>25</td>
</tr>
<tr>
<td>snoRNA Pseudo Genes</td>
<td>10</td>
</tr>
<tr>
<td>snRNA Genes</td>
<td>25</td>
</tr>
<tr>
<td>snRNA Pseudo Genes</td>
<td>42</td>
</tr>
<tr>
<td>scRNA Pseudo Genes</td>
<td>35</td>
</tr>
<tr>
<td>rRNA Genes</td>
<td>16</td>
</tr>
<tr>
<td>rRNA Pseudo Genes</td>
<td>29</td>
</tr>
<tr>
<td>tRNA Genes</td>
<td>19</td>
</tr>
<tr>
<td>tRNA Pseudo Genes</td>
<td>22</td>
</tr>
<tr>
<td>miscRNA Genes</td>
<td>56</td>
</tr>
<tr>
<td>miscRNA Pseudo Genes</td>
<td>41</td>
</tr>
<tr>
<td>Pseudo Genes</td>
<td>81</td>
</tr>
<tr>
<td>Pseudo Genes</td>
<td>82</td>
</tr>
</tbody>
</table>

---

[Image of bar chart and legend](#)
Chromosome 6: extracellular proteins

Extracellular proteins involved in immunity and inflammation

- Tumor necrosis factor α
- Lymphotoxin A&B
- 4 Lymphocyte antigens
- 4 complement factors
- 5 Defensins
- 3 Vanins
- Interleukin 17A and 17F
- Serum response factor
- Apolipoproteins A and M
- Vascular endothelial growth factor α
- Connective tissue growth factor
- 3 Serpin proteins
- Endothelin 1
- 6 Collagens
- 2 Laminins
**Chromosome 6: membrane proteins**

**Brain related**
- Cannabinoid receptor
- Opioid receptor mu 1
- 3 GABA receptors
- 2 serotonin receptors (1B & 1E)
- 3 glutamate receptors
- Schizophrenia disorder 3

**Non-brain related**
- 2 interleukin receptors
- Interferon gamma receptor
- 4+ psoriasis susceptibility genes
- Estrogen receptor (breast cancer)
- 9 G-protein coupled receptors
- 15 olfactory receptors
- 9 transmembrane proteins
- 15 solute carrier family members
Transcription factors and other proteins

- 26+ zinc finger proteins
- 7 transcription factors
- 3 PHD finger proteins Fyn and Fyn-related kinase
- Ezrin
- Flotillin
- 4 Gap junction proteins
- 2 Natural cytotoxicity triggering receptors
- 3 BCL-associated proteins
- Hemochromatosis gene
- Parkin 2 (juvenile Parkinson)
- Abelson helper integration site (schizophrenia, autism)
- 37 “similar to” proteins
- 60 open reading frames
Common translocations: genes and proteins re-arranged

- t(6;9)(p23;q34): DEK and CAN; AML (1%), myelodysplastic syndrome (rare)
- t(6;9)(q21-25;p13-24): adenoid cystic carcinoma
- t(6;11)(p21;q12-13); TFEB and Alpha; renal neoplasm of children and young adults
- t(6;12)(q23;q15): ? and HMGA2/HMGIC; hyaline vascular Castleman’s disease
- t(6;14)(p21.1;q32.3); cyclin D3 and IgH; gastrointestinal stromal tumors, multiple myeloma (4%), diffuse large B cell lymphoma
- t(6;14)(p25;q32): MUM/IRF4 and IgH; multiple myeloma (20%)
- 6p23: DEK; see t(6;9)(p23;q34)
- 6p21: pim-1
- 6p21: TFEB; see t(6;11)(p21;q12)
- 6p21.1: cyclin D3; seet(6;14)(p21.1;q32.3)
- 6p21.2: p21 WAF1/CIP1
- 6q22: ros
- 6q22-24: myb
- 6q24-27: mas
- 6p25: MUM/IRF4; see t(6;14)(p25;q32)
Chromosome 6: disease associations

Over 100 major disease associations recognized so far.

- 3-M syndrome
- 21-hydroxylase deficiency
- Acute promyelocytic leukemia
- **Alzheimer’s disease**
- age-related macular degeneration
- **ankylosing spondylitis**
- anterior segment dysgenesis
- arginase deficiency
- arrhythmogenic right ventricular cardiomyopathy
- autosomal recessive cerebellar ataxia type 1
- **autism**
- Axenfeld-Rieger syndrome
- **Behcet’s disease**
- bipolar disorder
- branchiooculofacial syndrome
- breast cancer
- bullus pemphigoid
- cbIf combined homocystinuria and methylmalonic aciduria
- Charcot-Marie-Tooth disease
- Char syndrome
- chordoma
- cleidocranial dysplasia
- **celiac disease**
- coenzyme Q10 deficiency
- collagenopathy (types II and XI)
- colon cancer
- **complement deficiency**
- cone dystrophy
- congenital muscular dystrophy
- Creutzfeld-Jacob disease
- 3/14/11
- 9
Chromosome 6: disease associations

Over 100 major disease associations recognized so far.

- craniometaphyseal dysplasia (AR type)
- **Crohn’s disease**
- D-2-hydroxy glutaric aciduria
- Dandy-Walker malformation
- **diabetes mellitus type 1**
- diabetic nephropathy
- dyslexia
- Ehlers-Danlos syndrome
- epilepsy
- factor XIII deficiency
- Fancony anemia
- gastric cancer
- Hashimoto thyroiditis
- head and neck cancer
- Hemochromatosis
- Huntington disease-like syndrome
- hypermethioninemia
- IgA nephropathy
- Ischemic heart disease
- Joubert syndrome
- Lafora progressive myoclonus epilepsy
- Leber congenital amaurosis type V
- Leigh syndrome
- ligneous conjunctivitis
- lung cancer
- **maple syrup urine disease**
- melanoma
- methylmalonic academia
- multiple epiphyseal dysplasia
- **multiple sclerosis**
- narcolepsy
- neural tube defects
Chromosome 6: disease associations

Over 100 major disease associations recognized so far...

- nephritis
- neuroblastoma
- non-ketotic hyperglycinemia
- nonsyndromic deafness
- oculodentodigital dysplasia
- otospondylomegaepiphyseal dysplasia
- palmoplantar keratoderma
- Parkinson disease
- Pemphigus vulgaris
- pituitary adenoma
- plasminogen deficiency
- PLO/sclerosing leukoencephalopathy
- polycystic kidney disease
- polymyalgia rheumatica

- polypoidal choroidal vasculopathy
- POEMS syndrome
- pontocerebellar hypoplasia
- porphyria
- porphyria cutanea tarda
- premature ovarian failure
- primary ciliary dykinesia
- prostate cancer
- proxisome biogenesis disorders
- psoriasis
- Refsum disease
- restless legs syndrome
- retinal neovascularization
- retinitis pigmentosa
Chromosome 6: disease associations

Over 100 major disease associations recognized so far..

- rheumatoid arthritis
- rhizomelic chondrodysplasia punctata
- schizophrenia
- Schmid type metaphyseal chondrodysplasia
- sialic acid storage disease
- sialidosis
- specific language impairment
- Spinocerebellar ataxia
- spondylometaphyseal dysplasia
- spinocerebellar ataxia
- spontaneous and recurrent abortion
- Stargardt macular degeneration
- Stickler syndrome
- succinic semialdehyde dehydrogenase deficiency
- sudden infant death syndrome
- systemic lupus erythematosus
- trichothiodystrophy
- Tourette syndrome
- tumor angiogenesis
- viral resistance and responsiveness
- vitelliform macular dystrophy
- Wegener granulomatosis
- Weissenbacher-Zweymuller syndrome
- xeroderma pigmentosum
- X-linked sideroblastic anemia
- Zellweger syndrome
Chromosome 6: disease associations

Over 100 major disease associations recognized so far.

Malignancies
- Lymph node metastasis in gastric cancer
- Gall bladder cancer
- Breast cancer in Russia
- Skin cancer
- Melanoma
- Thyroid cancer
- Germ-cell testis tumors
- Burkitt's lymphoma
- Cervical cancer in HPV-16-positive patients
- Protective effect from renal cell carcinoma

Leukemias
- Adult acute myeloblastic leukemia
- Childhood acute lymphoblastic leukemia
- Childhood acute lymphoblastic leukemia
- Chronic myeloid leukemia
- Chronic lymphoid leukemia
- Large granular lymphocyte leukemia with arthritis
Chromosome 6: structure & organization

MHC region: 6p21.3

Fig 3.20 © 2001 Garland Science

Fig 3.21 © 2001 Garland Science
Chromosome 6: alloimmunity

Transplantation
Chromosome 6: autoimmunity

Rheumatoid disease
Chromosome 6: uncertain etiology

Schizophrenia and MS

Early and Late Gray Matter Deficits in Schizophrenia

EARLIEST DEFICIT

5 YEARS LATER (SAME SUBJECTS)

STG, DLPFC

Thompson et al., 2001

Basic Science

- genomics, proteomics, biormarkars, metabolomics, neuroimaging, international cooperation

Advanced Analysis

- metabolic pathways, brain banks, databases, endophenotypes, candidate proteins, candidate genes

Schizophrenia Biomarkers

- high sensitivity and specificity, validated, standardized, reliable and reproducible, noninvasive, simple to perform, inexpensive

Clinical Application

- understanding pathogenesis, diagnosis, classifications, prognosis, clinical management, drug discovery and development
Chromosome 6: disease associations

Schizophrenia: gene associations

- NQO2: NAD(P)H dehydrogenase, quinone 2
- JARID2: jumonji, AT rich interactive domain 2
- ATXN1: ataxin 1
- SIRT5: sirtuin (silent mating type information regulation 2 homolog) 5
- DTNBPI: dystrobin binding protein 1
- MOG: myelin oligodendrocyte glycoprotein
- GABBR1: gamma-aminobutyric acid (GABA) B receptor, 1
- BAK1: BCL2-antagonist/killer 1
- C4A: complement component 4A (Rodgers blood group)
- C4B: complement component 4B (Childo blood group)
- DDR1: discoidin domain receptor tyrosine kinase 1
- ATF6B: activating transcription factor 6 beta
- GRM4: glutamate receptor, metabotropic 4
- LTA: lymphotoxin alpha (TNF superfamily, member 1)
- MICB: MHC class I polypeptide-related sequence B
- NOTCH4: Notch homolog 4 (Drosophila)
- TNF: tumor necrosis factor (TNF superfamily, member 2)
- TNXB: tenasin XB
- DAAM2: dishevelled associated activator of morphogenesis 2
- SLC25A27: solute carrier family 25, member 27
- CNR1: cannabinoid receptor 1 (brain)
- GRIK2: glutamate receptor, ionotrophic, kainate 2
- FABP7: fatty acid binding protein 7, brain
- TAAR6: trace amine associated receptor 6
- AHII: Abelson helper integration site 1
- SOD2: superoxide dismutase 2, mitochondrial
- TCP1: t-complex 1
- TBP: TATA box binding protein

Chromosome 6: disease associations

- 6p25.2
- 6p24.3
- 6p22.3
- 6p21.3
- 6p12.1
- 6q12
- 6q14.3
- 6q16.3
- 6q22.3
- 6q22.33
- 6q23.2
- 6q24.1
- 6q24.3
- 6q25.2
- 6q26

- 4835
- 3720
- 6310
- 23408
- 84062
- 4340
- 2550
- 578
- 720
- 721
- 780
- 1388
- 2914
- 4049
- 4277
- 4855
- 7124
- 7148
- 23500
- 9481
- 1268
- 2898
- 2173
- 319100
- 54806
- 6648
- 6950
- 6908
- 6p24-p23
- 6p21.3
- 6p21.2
- 6p12.3
- 6q16.1
- 6q22.1
- 6q22.33
- 6q23.2
- 6q24.1
- 6q24.3
- 6q25.2
- 6q26

Pr. Dermot Kelleher, 2010
**Recent disease associations**

**Dementia Revealed: Novel Chromosome 6 Locus for Late-Onset Alzheimer Disease Provides Genetic Evidence for Folate-Pathway Abnormalities**

**Genetics of Type 1A Diabetes**

**Somatic mutations of the Parkinson's disease–associated gene PARK2 in glioblastoma and other human malignancies.**
Et al. Nature Genetics, Volume: 42, 77–82, 2010
Examples of key diseases in Canada

**Rheumatoid disease:** Prevalence almost one quarter million. Relapsing progressing inflammatory joint disease with deformity and incapacitation. Management improving, but restoration of specific tolerance is the ultimate goal. Economic burden very high.

**Juvenile Diabetes:** Prevalence 1% of population. Progressive blindness, renal failure, vascular damage and amputations. Life reduced by at least 15 years. Management improving but inadequate, no cure. Economic burden estimated almost $2 billion per year.


**Alzheimer’s Disease:** Prevalence over half a million and increasing. Most important form of dementia (65%). No effective management, no cure. Profound societal and economic burden estimated at $15 billion per year.

**Multiple sclerosis:** Prevalence 75,000. Canada is a high-risk region of the world. Progressive demyelinating disease leading to paralysis and immobility. No effective management or cure. Profound societal and economic burden estimated at $1 billion.
Chromosome 6: project plan

Long-term, HPP-linked Goals:

• Confirm expression of all protein-coding genes (~1132), including tissue distribution

• Identify archetypical peptides of isoforms and for modifications

• Contribute spectra to spectral libraries

• Identify all splice variants, modifications: phospho and glyco, N and C-termini

• Characterize binding partners: interactome (proteins, metabolites, nucleic acids, etc).

• Make tools & data available

Short-term Goals: 3-year plan

• Start with short arm (6p)

• confirm expression

• confirm splice variants, modifications & binding partners

• focus on gene families/groups of particular interest
Chromosome 6: project plan

**Histone Cluster**
- Obvious biological relevance:
  - Modifications of histone tails
  - Tie in with current epigenetic efforts
- Driving questions:
  - When & where are the marks laid down?
  - What is the connectivity between marks?
- Economic opportunities:
  - Novel diagnostics
  - MRM methodology

**MHC Cluster**
- Obvious biological relevance:
  - Antigen presentation, regulation
  - Immunity, infection, transplantation
- Driving questions:
  - What are the peptide repertoires?
  - How do the relate to disease?
- Economic opportunities:
  - T / B cell tolerance
  - Vaccines, therapeutics

**Other Cluster**
- Obvious biological relevance:
  - What are the other genes / proteins
  - Are they expressed
- Driving questions:
  - What are their interactions?
  - What are their functions?
- Economic opportunities:
  - To be determined...

**Start** Year 1 Year 2 Year 3 Year 4 Year 5 Year 6
### Chromosome 6: project plan

#### Informatics
- Before: Define what we know, and don’t know
- During: establish a repository to collect data from project participants.
- Do not re-invent the wheel
- work with UniProt (or HPRD) to make information available
- tissue distribution (or a stronger statement that a protein may not exist)
- SRM transitions (also into Skyline & SRMAAtlas)
- splice variants & PTMs already handled

#### Confirming expression
- Avoid brute force approach
- Make use of plentiful microarray data to target high probability tissues
  - Clinician partnerships
  - Some deep sequencing
- SRMs: faster, cheaper, more sensitive, more applicable

#### Splice variants
- What evidence exists?
  - EST/cDNA libraries
  - PeptideAtlas, GPM, etc.
- Add in data from expression confirmation studies
  - no new deep sequencing
- Design SRMs to cover all important predicted variants of families of interest
Defining HPP standards

- As a leader of the Human Proteome Project, Canada will define & shape the quality and standards for the international community.

- The size and medical and societal importance of chromosome 6 is commensurate with Canada’s leading role in proteomics.

- Goal: have all countries measure expression (by SRM) in same tissue samples – will need a bank spanning tissues & populations.

- Goal: minimum data quality requirements
  - e.g., S/N & specificity for SRM detection
  - e.g., FDR or % coverage for MuDPIT
  - e.g., 99% confidence for specific PTM sites.
# Chromosome 6: collaborators

<table>
<thead>
<tr>
<th>National</th>
<th>International</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. Overall, UBC</td>
<td>G. Nepom, USA</td>
</tr>
<tr>
<td>L. Foster, UBC</td>
<td>F. Claas, Netherlands</td>
</tr>
<tr>
<td>D. Figeys, UO</td>
<td>G. Opelz, Germany</td>
</tr>
<tr>
<td>K. Siminovich, UT</td>
<td>C. Seusal, Germany</td>
</tr>
<tr>
<td>K. Tinckam, UT</td>
<td>P. Dyer, UK</td>
</tr>
<tr>
<td>P. Campbell, UA</td>
<td>F. Christiansen, Australia</td>
</tr>
<tr>
<td>L. West, UA</td>
<td>R. Duquesnoy, USA</td>
</tr>
<tr>
<td>N. Berka, UC</td>
<td>H. Gebel, USA</td>
</tr>
<tr>
<td>P. Nickerson, UM</td>
<td>M. Leffel, USA</td>
</tr>
<tr>
<td>M. Barnett, UBC</td>
<td>E. Reid, USA</td>
</tr>
<tr>
<td>T. Nevill, UBC</td>
<td>A. Zachary, USA</td>
</tr>
<tr>
<td>M. Cantarovich, MU</td>
<td>M. Zeevi, USA</td>
</tr>
<tr>
<td>S. Chung, UBC</td>
<td>M. Gerbase-deLima, Brasil</td>
</tr>
<tr>
<td>E. Cole, UT</td>
<td></td>
</tr>
<tr>
<td>J. Tchervenkov, MU</td>
<td></td>
</tr>
<tr>
<td>G. Warnock, UBC</td>
<td></td>
</tr>
<tr>
<td>N. Reiner, UBC</td>
<td></td>
</tr>
<tr>
<td>E. Keystone, UT</td>
<td></td>
</tr>
<tr>
<td>A. Jevnikar, UWO</td>
<td></td>
</tr>
<tr>
<td>S. Keshavjee, UT</td>
<td></td>
</tr>
<tr>
<td>G. Levy, UT</td>
<td></td>
</tr>
<tr>
<td>Q. Madrenas, UWO</td>
<td></td>
</tr>
<tr>
<td>D. White, UWO</td>
<td></td>
</tr>
<tr>
<td>D. Thompson, UBC</td>
<td></td>
</tr>
<tr>
<td>B. Verchere, UBC</td>
<td></td>
</tr>
<tr>
<td>J. Shapiro, UA</td>
<td></td>
</tr>
<tr>
<td>B. McManus, UBC</td>
<td></td>
</tr>
<tr>
<td>R. McMaster, UBC</td>
<td></td>
</tr>
<tr>
<td>T. Feasby, UA</td>
<td></td>
</tr>
<tr>
<td>A. Traboulsee, UBC</td>
<td></td>
</tr>
<tr>
<td>J. Stoessl, UBC</td>
<td></td>
</tr>
<tr>
<td>L. Tibbles, UC</td>
<td></td>
</tr>
<tr>
<td>D. Grant, UT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organizations</strong></td>
<td><strong>Canadian Society of Transplantation</strong></td>
</tr>
<tr>
<td></td>
<td><strong>The Transplantation Society</strong></td>
</tr>
</tbody>
</table>
Chromosome 6: partners

Preliminary list of clinical, academic and industrial partners

In affiliation with:
- University of British Columbia
- University of Victoria
- McGill University
- University of Calgary
- Faculty of Medicine & Dentistry, University of Alberta
- Simon Fraser University
- Ottawa Institute of Systems Biology
- University of Victoria-Genome BC
- Proteomics Centre
- USC/CHLA
- Microarray Core
- iCAPTURE Centre for Cardiovascular and Pulmonary Research
- PROOF Centre of Excellence
- Genome British Columbia
- Genome Canada
- Genome Québec
- Vancouver Coastal Health Research Institute
- Providence Health Care
- Novartis
- Astellas
- Roche
- AstraZeneca
- Janssen
- Affymetrix
- IDT

Networks of Centres of Excellence of Canada