Gene regulatory networks to identify new targets for disease – applications to type I diabetes, cell multinucleation and epilepsy

Enrico Petretto, Ph.D.

Associate Professor
Centre for Computational Biology
Program in Cardiovascular & Metabolic Disorders
My aim

Identify new gene targets for complex disease
The traditional genetics approach

Clinical data

Functional gene targets

Pathological pathways

statistical inference

GWAS
WES
WGS

Disease

Genetic variation
Systems-Genetics

Relevant cellular/tissue systems
Systems-Genetics

Relevant cellular/tissue systems

gene regulatory networks

mRNAs
Protein-Protein Interactions
Signaling networks
Systems-Genetics

Relevant cellular/tissue systems

gene regulatory networks

Disease
Systems-Genetics

Disease

Master regulators
Gene networks
Multi-phenotypes

Relevant cellular/tissue systems

gene regulatory networks

GWAS

Disease

genetic variation

causality
Systems-Genetics

- Relevant cellular/tissue systems
- Genetic variation
- Master genetic regulator
- Gene regulatory networks
- Disease
Systems-Genetics

Relevant cellular/tissue systems

gene regulatory networks

regulatory process

master genetic regulator

gene regulatory networks

Disease

genetic variation
Systems-Genetics

Relevant cellular/tissue systems

Gene regulatory networks

Genetic variation

Master genetic regulator

Gene target for disease modification

regulatory process
The systems-genetics workflow

- Disease
- Cellular process
- Regulatory network
- Genetic variation
The systems-genetics workflow

Type 1 diabetes

Inflammatory disease

Epilepsy

Proinflammatory response

Cell multinucleation

Neuroinflammation

EBI2

TREM2

SESN3

Cell Reports 2014

Nature Communications 2015 (in press)
First example of systems-genetics to complex disease

pro-inflammatory gene network

genetic control

rat *Ebi2*

Heinig*, Petretto* et al. *Nature* 2010
Genome-wide expression profiling in 7 tissues
+ Genome-wide SNPs
+ Genome-wide Transcription Factor Binding Sites (TFBS) predictions*

1. TFs expression linked to a genetic variation (i.e., expression QTLs or eQTLs)
2. Differential expression of the TF together with its predicted targets
3. Genome-wide analysis of common genetic control of TFs and eQTLs in seven rat tissues
Transcription-factor-driven regulatory cascade:

genetic variation modulates the expression of *Irf7* and *Irf7* target genes (*irf7*-network)
Transcription-factor-driven regulatory cascade:

genetic variation modulates the expression of *Irf7* and *Irf7* target genes (*irf7*-network)
**Irf7-Driven Inflammatory Network (iDIN)**

(305 genes)

- **Immune response** ($P = 3.6 \times 10^{-19}$)
- **Response to virus** ($P = 2.5 \times 10^{-7}$)
- **Acute inflammatory response** ($P = 2.6 \times 10^{-5}$)
Mapping the genetic control of networks

Genetic regulation?
Mapping of eQTLs and regulatory control points of networks

**New Insights into the Genetic Control of Gene Expression using a Bayesian Multi-tissue Approach**

Enrico Petretto\(^1\), Leonardo Bottolo\(^1\), Sarah R. Langley\(^1\), Matthias Heinig\(^2\), Chris McDermott-Roe\(^1\), Rizwan Sarwar\(^1\), Michal Pravenc\(^4\), Norbert Hübner\(^3\), Timothy J. Aitman\(^1\), Stuart A. Cook\(^1\), Sylvia Richardson\(^2\)

- Fully Bayesian variable selection in the large $p$, small $n$ paradigm
- Multivariate Gaussian distributions for formulating a multiple response model of

$$
Y(n \times q) \quad \text{on} \quad X(n \times p)
$$

$Y \sim \mathcal{N}(I_n, \Sigma)$

$\Sigma$ is a $q \times q$ covariance matrix between the $q$ outcomes

a latent binary vector $\gamma$ is used to induce sparsity and find the subset of predictive SNPs on multiple gene expression levels

**ESS++** explores the $2p$-dimensional model space using an extension of parallel tempering called Evolutionary Monte Carlo that combines Markov chain Monte Carlo (MCMC) and genetic algorithms

(rescaled) Bayes factors for the entire network

$$
m_{ij}^{BF} = f_t \frac{1}{|\mathcal{S}_j|} \sum_{g \in \mathcal{S}_j} BF(\gamma_{tg} = 1; \gamma_{tg} = 0)
$$

average BF for the regulatory ‘hot-spot’

90% range (5th-95th percentiles)

$j \in J$, with $J$ the set 10 putative master regulators

$|\mathcal{S}_j|$, number of associated genes for SNP $j$

$f_t$, rescaling factor such as all the master regulators in the $j$ set control the same number of network genes
Regulatory ‘hot-spots’ for the iDIN

Common genetic control of the iDIN by a single locus on rat chromosome 15q25

Ebi2
From rats … to humans

Is the rat iDIN functionally conserved with humans?

Do the iDIN and its master regulator (*Ebi2*) play a role in human autoimmune disorders (T1D)?
Network replication and cross-species conservation

• Genome-wide expression in human monocytes:
  – Gutenberg Heart Study (n=1490)
  – Cardiogenics (n=758)

• Overlap between iDINs in humans \( (P = 8.3 \times 10^{-23}) \)

• Overlap between rat and human iDIN \( (P = 9.1 \times 10^{-20}) \)
Annotating gene networks with GWAS signal

Gene co-expression network/module

GWAS results for Type 1 Diabetes*

Gene 1: GWAS P-value = 0.00001
Gene 2: GWAS P-value = 0.023
Gene 3: GWAS P-value = 0.8
Gene 4: GWAS P-value = 0.01
Gene 5: GWAS P-value = 0.0002
Gene n: GWAS P-value = 0.056

GWAS enrichment statistic for network

*collaboration with John Todd and the Wellcome Trust Case Control Consortium (WTCCC)
Cross-species integrated iDIN

Association of the iDIN with T1D SNPs
vs randomly selected genes, $P = 2.5 \times 10^{-10}$
vs randomly selected *immune response* genes, $P = 8.8 \times 10^{-6}$
Mapping master regulators for the human IDIN

- human chromosome 13q32
- rat chromosome 15q25
- synteny
- genetic control
Mapping master regulators for the human IDIN

- rs11616269
- rs9557207

FDR < 5%

- human chromosome 13q32
- rat chromosome 15q25

Synteny
Using Systems-Genetics we
- Implicated the **innate viral response** pathway in the aetiology of T1D
- Identified a new **gene target** regulating this process in T1D

---

**pro-inflammatory gene network**

**Type I diabetes**

**rat Ebi2**

**human EBI2**

**genetic control**

**cross-species conservation**

Heinig*, Petretto* et al. *Nature* 2010
The systems-genetics workflow

Disease
Type 1 diabetes
Inflammatory disease
Epilepsy
Cellular process
proinflammatory response
cell multinucleation
neuroinflammation
Regulatory network
EBI2
TREM2
SESN3
Genetic variation
Nature 2010
Cell Reports 2014
Nature Communications 2015 (in press)
Multinucleated cells

The “smiling” multinucleated macrophage
**The WKY rat: a model of kidney disease** (glomerulonephritis\textsuperscript{1,2}) & multinucleated giant cell formation

The WKY rat: a model of kidney disease (glomerulonephritis\textsuperscript{1,2}) & multinucleated giant cell formation

Expression profiling in bone marrow derived macrophages
WKY x LEW back-cross population (n=200 rats)

Mapping the genetic determinants of gene expression
cis vs trans expression QTLs

cis-acting

eQTL gene

trans-acting

eQTL gene

Regulatory gene networks
Boosting detection of trans-eQTLs

Traditional eQTL mapping

-log_{10}(p-value) vs. Rat genome (cM)

FDR = 28%

D15Rat107

P < 0.05

Irf7

Bayesian eQTL mapping

Boosting detection of trans-eQTLs

Traditional eQTL mapping

Bayesian eQTL mapping

Mapping the genetic determinants of macrophage gene networks in WKY x LEW back-cross population (n=200 rats)

190 trans-regulated transcripts posterior probability >80%

Rat genome
Trem2 regulates in trans a network involved in macrophage multinucleation.
Trem2 regulates in \textit{trans} a network involved in macrophage multinucleation.

\textit{trans}-regulated gene network in macrophages

\textit{Trem2} regulates bone homeostasis by controlling the rate of osteoclastogenesis\(^1\)

Network enrichment for osteoclast-expressed genes \((P = 4 \times 10^{-7})\)

Master genetic regulator

Trem2

Kcnn4

Strongest trans-regulated gene in the network (posterior probability 0.999, P-value $\approx 10^{-16}$)

*Kcnn4*, intermediate-conductance calcium-activated potassium channel protein
Using Systems-Genetics we generated *testable hypotheses* on novel gene targets for glomerulonephritis and arthritis.

**Disease Models**
- Glomerulonephritis
- Arthritis

**Gene Network**
- TREM2
- KCNN4
- Macrophage multinucleation
- Macrophage & osteoclast multinucleation
Using Systems-Genetics we generated *testable hypotheses* on novel gene targets for glomerulonephritis and arthritis.
The systems-genetics workflow

Type 1 diabetes
- proinflammatory response
- EB12

Inflammatory disease
- cell multinucleation
- TREM2

Epilepsy
- neuroinflammation
- SESN3

Disease

Cellular process

Regulatory network

Genetic variation
Problems with studying disorders of the human brain

Access to ‘fresh’ tissue?

Where is my control tissue?
Finding genetic drivers of networks in pharmacoresistant epilepsy and inform on potential therapeutic targets

- Most of the health cares costs of epilepsy are due to those patients with pharmacoresistant (intractable) seizures

- Therapy: selective removal of the hippocampus, which allows access to ante-mortem ‘fresh’ brain tissue samples

Ante-mortem brain tissues from 129 temporal lobe epilepsy patients
Systems-genetics: application to human epilepsy

N=129 patients with temporal lobe epilepsy (TLE)

Gene expression in hippocampus

Genotype (SNP) data

Clinical data
Systems-genetics: application to human epilepsy

N=129 patients with temporal lobe epilepsy (TLE)

Gene expression in hippocampus

Genotype (SNP) data

Clinical data

Integrate multi-level data using network approaches to identify new gene targets for epilepsy
Gene co-expression network comprising 439 annotated genes identified in the hippocampus of 129 TLE patients

The Temporal Lobe Epilepsy (TLE)-network

Significant association between mRNA expression profiles of genes $(i,j)$ in the human hippocampus

Posterior probability $> 95$

Graphical Gaussian Models\textsuperscript{1,2}


GWAS data implicate the TLE-network in epilepsy susceptibility

The TLE-network is significantly enriched for GWAS signals (31% of genes, $P = 2 \times 10^{-7}$)

GWAS susceptibility to partial epilepsy*

$P_{\text{GWAS}} = 3 \times 10^{-4}$

$P_{\text{GWAS}} = 0.05$

* 1,429 focal epilepsy cases vs 7,358 controls
Functional specialization within the TLE-network

Module-1

Module-2

- Log_{10}(P-value)

0     2     4      6     8    10

-Log_{10}(P-value)

0      1     2      3     4      5

Cytokine-cytokine receptor interaction
NOD-like receptor signaling pathway
Toll-like receptor signaling pathway
Chemokine signaling pathway
MAPK signaling pathway
p53 signaling pathway

Focal adhesion
ECM-receptor interaction
Functional specialization within the TLE-network

Module-1

Module-2

Microarray expression levels (arbitrary units)

Network genes

All other genes on the microarray

**  ***

**  ***

**  ***

-Log_{10}(P-value)

-Log_{10}(P-value)

Pathways:
- Cytokine-cytokine receptor interaction
- NOD-like receptor signaling pathway
- Toll-like receptor signaling pathway
- Chemokine signaling pathway
- MAPK signaling pathway
- p53 signaling pathway
- Focal adhesion
- ECM-receptor interaction
Toll-like receptor (TLR) signaling and epilepsy

Maroso et al., Nat Med 2010
Toll-like receptor (TLR) signaling and epilepsy

Maroso et al., Nat Med 2010
Toll-like receptor (TLR) signaling and epilepsy

Maroso et al., Nat Med 2010
From humans … to mice

Is the human TLE-network functionally conserved with the mouse hippocampus?

Is it associated with epilepsy in the mouse?
RNA-seq analysis of 200 hippocampi from epileptic and control mice

TLE-network is specifically conserved in epileptic mouse hippocampus
RNA-seq analysis of 200 hippocampi from epileptic and control mice

Module-1 genes

**Fold change (epilepsy / control)**

**TLR-signaling and cytokines genes**

(enrichment $P = 9.03 \times 10^{-4}$)

**TLE-network is specifically conserved in epileptic mouse hippocampus**
TLR/IL1 signaling

Human hippocampus epileptic-network

coordinated transcriptional programme

INFLAMMATION

SEIZURES
Primary gene target in the human hippocampus ???

TLR/IL1 signaling

coordinated transcriptional programme

Human hippocampus epileptic-network

INFLAMMATION

SEIZURES
Bayesian genetic mapping of “master regulators” for networks

Fully-multivariate and model based (>>> probability of eQTL detection)

\[
\Omega = \begin{bmatrix}
\omega_{11} & \cdots & \omega_{1j} & \cdots & \omega_{1p} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\omega_{k1} & \cdots & \omega_{kj} & \cdots & \omega_{kp} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\omega_{q1} & \cdots & \omega_{qj} & \cdots & \omega_{qp}
\end{bmatrix}
\]

- \( p = 1 \ldots \text{number of predictors (SNPs)} \)
- \( q = 1 \ldots \text{number of expression traits} \)
Bayesian genetic mapping of “master regulators” for networks

Fully-multivariate and model based (>>> probability of eQTL detection)

\[
\Omega = \begin{bmatrix}
\omega_{11} & \cdots & \omega_{1j} & \cdots & \omega_{1p} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\omega_{k1} & \cdots & \omega_{kj} & \cdots & \omega_{kp} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\omega_{q1} & \cdots & \omega_{qj} & \cdots & \omega_{qp}
\end{bmatrix}
\]

\( p = 1 \ldots \text{number of predictors (SNPs)} \)

\( q = 1 \ldots \text{number of expression traits} \)

\( \omega_{kj} = \omega_k \times \rho_j \)  \hspace{1cm} “Propensity” of the \( j \)th predictor be a master regulator of the entire network (or multiple trans-eQTLs)

“Base level probability” of association for the \( k \)th trait with predictors

Finding the master regulatory loci of transcriptional Modules

Module-1

Module-2

rs10501829

FDR <5%

FDR <5%
Finding the master regulatory loci of transcriptional Modules

- Log₁₀(Bayes Factor) vs. chromosome
- Proportion of Module-1 genes associated with each SNP
- Chromosome 11q21 locus
- Significance of correlation - Log₁₀(P-value)
- rs10501829, rs7107661, rs6483435
- Module-1 genes: SESN3, KDM4DL, SRSF8, ENDO01, FAM76B, CEP57, MTMR2, MAML2

Module-1

Average correlation with Module-1 genes expression

Pearson correlations

-0.3, 0, 0.4

Proportion of Module-1 genes associated with each SNP

Mbp

383 kb
Finding the master regulatory loci of transcriptional Modules

Proportion of Module-1 genes associated with each SNP

Significance of correlation

Average correlation with Module-1 genes expression

Module-1
*SESN3* is associated with epilepsy and regulates neuro-inflammatory gene expression (Module-1) in mouse epileptic hippocampus.

Pearson correlation between Module-1 genes and Sesn3 expression

\[ R^2 = 0.35 \ (P = 9.4 \times 10^{-6}) \]
**SESN3 regulates neuro-inflammatory gene expression (in vitro)**

**LPS-stimulated microglial cells**

**Unstimulated microglial cells**

**Module-1 genes**
SESN3 regulates neuro-inflammatory gene expression (in vitro)
Zebrafish model of epilepsy


Locomotor activity assay

Control

PTZ induction

Induction of neuronal activity regulated c-fos

in collaboration with Vincent Cunliffe (University of Sheffield)
SESN3 regulates PTZ-induced seizures \textit{in vivo}

Reduced PTZ-induced locomotor activity in SESN3 morpholinos

\begin{itemize}
\item Control morpholino
\item Control morpholino + PTZ
\item SESN3 morpholinos
\item SESN3 morpholinos + PTZ
\end{itemize}

\begin{itemize}
\item Cumulative distance swam (mm)
\item Control morpholino
\item Control morpholino + PTZ
\item SESN3 morpholinos
\item SESN3 morpholinos + PTZ
\end{itemize}

\begin{itemize}
\item P = 2.7 \times 10^{-4}
\end{itemize}

in collaboration with Vincent Cunliffe (University of Sheffield)
Rescue of zebrafish sesn3 morphant phenotype by co-injection of synthetic sesn3 mRNA

in collaboration with Vincent Cunliffe (University of Sheffield)
**SESN3 regulates Module-1 gene expression (in vivo)**

**c-fos (in situ hybridization)**

Control morpholino  Sesn3 i3e4 + e4i4 morpholinos

+ PTZ

- PTZ

---

**Reduction of PTZ-induced gene expression in the Sesn3-morphants**

<table>
<thead>
<tr>
<th>Gene</th>
<th>mRNA expression reduction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-fos</td>
<td>35-64%</td>
</tr>
<tr>
<td>Atf3</td>
<td></td>
</tr>
<tr>
<td>Fosb</td>
<td></td>
</tr>
<tr>
<td>Egr2b</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 6.** qPCR analysis of seizure-induced gene expression in PTZ-treated control and sesn3 morphant larvae. For each sample, total RNA was extracted from 8-20 stage-matched, sibling zebrafish larval heads. Relative expression levels, normalized to beta actin gene expression, were determined by using the $2^{-\Delta\Delta Ct}$ method. At least 4 biological replicates were used for the qPCR analysis. Bottom right panel: both c-fos and Module-1 genes (Atf3, Egr2b, Fosb) showed consistent reduction (35-64%) in mRNA expression in sesn3 morphant larvae, as compared with control morphant larvae, after PTZ treatment. mRNA expression data are reported as mean ± sem. NS, not significant.

in collaboration with Vincent Cunliffe (University of Sheffield)
**SESN3 regulates Module-1 gene expression (in vivo)**

**c-fos (in situ hybridization)**

- **Control morpholino**
- **Sesn3 i3e4 + e4i4 morpholinos**

**Baseline induction of pro-inflammatory Module-1 genes by Sesn3**

- **Control (uninjected embryos)**
- **Embryos injected with sesn3 mRNA**

Reduction of PTZ-induced gene expression in the Sesn3-morphants

- **Control**
- **Embryos injected with sesn3 mRNA**

---

**Module-1 genes**

- **c-fos**
- **Atf3**
- **Fosb**
- **Egr2b**
- **Gapdh**

**Baseline induction of pro-inflammatory Module-1 genes by Sesn3**

- **P = 0.0001**
- **P = 0.007**
- **P = 0.02**
- **P = 0.05**
- **ns**


in collaboration with Vincent Cunliffe (University of Sheffield)
**Gene co-expression network analysis**

**Genetic susceptibility analysis**

**Gene co-expression network analysis**

**Genome-wide Bayesian mapping of the genetic control of co-expression networks**

**New gene target for disease modification in human epilepsy**

**GWAS data for susceptibility to human focal epilepsy**

**Functionally enriched co-expression modules**

**surgical hippocampal samples from patients with temporal lobe epilepsy**

129 patients

**Functional conservation**

**Systems genetics identifies Sestrin 3 as a regulator of a proconvulsant gene network in human epileptic hippocampus**

*Article Details*

** nature COMMUNICATIONS**

**DOI: 10.1038/s41467-019-09331**


**New gene target for disease modification in human epilepsy**
The systems-genetics workflow

Type 1 diabetes

proinflammatory response

Inflammatory disease

cell multinucleation

Epilepsy

neuroinflammation

EBI2

TREM2

SESN3

Nature 2010

Cell Reports 2014

Nature Communications 2015
The systems-genetics workflow

Heart disease
(dilated cardiomyopathy)

Heart disease
(cardiac fibrosis)

Memory performance

long term potentiation (LTP)
Ca signaling and axon guidance

rat/human heart

human brain

...
Acknowledgments

**Duke-NUS**
Sarah Langley
Nathan Harmston
Owen Rackham

**Imperial College London**
Aida Moreno-Moral
Xiaolin Xiao
Tiziana Rossetti
Maxime Rotival
Prashant Srivastava
Antonio Marvao
Kirill Shkura
Andree Delahaye-Duriez

**Collaborators**

**Heart studies**
- Stuart Cook (Duke-NUS)
- Paul Barton (IC London)
- Norbert Hubner (MDC Berlin)

**Epilepsy, memory and cognition**
- Michael Johnson (IC London)
- Albert Becker (Uni Bonn)
- Rafal Kaminsky (UCB)
- Vincent Cunliffe (Uni Sheffield)

**Inflammation and kidney disease**
- Jacques Behmoaras (IC London)
- Terry Cook (IC London)

**Bayesian modeling**
- Leonardo Bottolo (IC London)

**Funding and industrial partnerships**

- Wellcome Trust
- EURATRANS
- British Heart Foundation
- UCB
- University of Bonn