MICE, ORGANOIDS AND SINGLE CELLS: COMPUTATIONAL METHODS FOR CANCER TREATMENT

Elizabeth Permina

CTCR (Otago)

GASTRIC CANCER

- It is one of the leading causes of cancer-related deaths in the world

- Although the overall death toll had decreased…

  - Its incidence is >3 fold higher rate in Māori and Pasifika than Pakeha

  - 43% of all cases are metastatic at diagnosis, a stage when the 5 year survival rate is ~4%

CDH1 ROLE

- CDH1 is a tumor suppressor
- The loss of CDH1 has been linked to several cancer types, in particular, Hereditary Diffuse Gastric Cancer (HDGC) and Lobular Breast Cancer (LBC)
- CDH1 germline mutations have a high penetrance effect
- CDH1-/+ is a dominant genotype
- Loss of function in both alleles is embryonic lethal

Māori kindred

Guilford P *et al.* Nature 392:402,1998
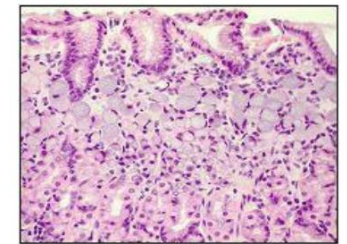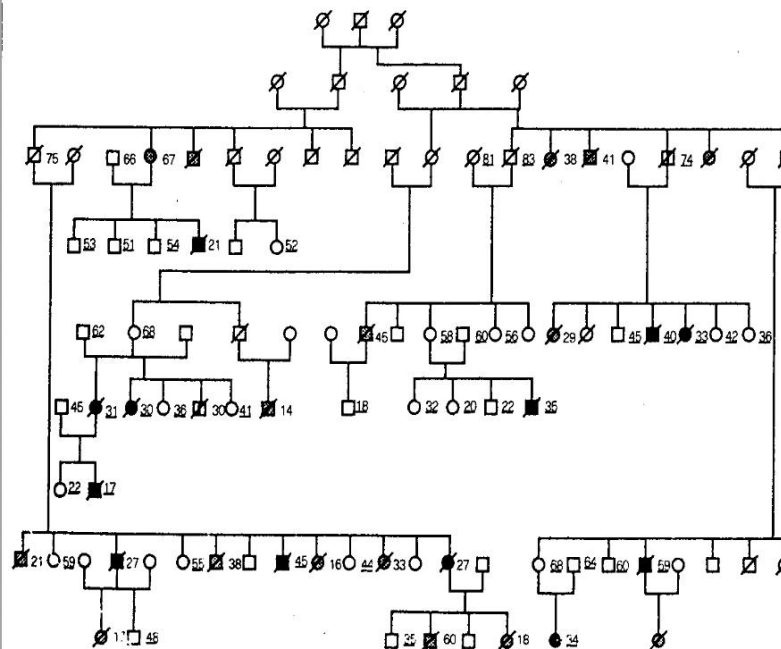
E-cadherin gene (CDH1)*germline mutations

Hereditary Diffuse Gastric Cancer (HDGC)

*Gene map locus: 16q22.1 (MIM ID +192090)

WHEN THINGS GO WRONG

Missense mutations affect cell-cell adhesion, motility and invasion

T340A, A634V, W409R, V832M, E757K

Functional Relevant

Adhesion, Motility, Invasion

A617T, ….

Functional Irrelevant

"neutral variants"

Suriano G *et al.* Hum Mol Genet 12:3007, 2003

WHY HDGC?

- It has a very poor survival rate compared to other types?

- We don't know exactly why – looking for other pathways that could be involved?

- The only reliable option at the moment is genotyping and preventative gastrectomy

- Can we do better?

  - Genomics-informed treatment

  - Drug development

THE MODEL

- HDGC

- Organoids instead of a whole animal or cell lines

- We know the driving gene, so we knock it out

- Study individual organoid cells using genomic technologies (e.g. sc-RNAseq)

ORGANOIDS?

- Stem cells make formations that act as miniature organs in early stages of development (but not for long)
- We can still get the genomic profile
- Make a more biologically relevant model than suspended cells
- Potentially we can see multiple cell types
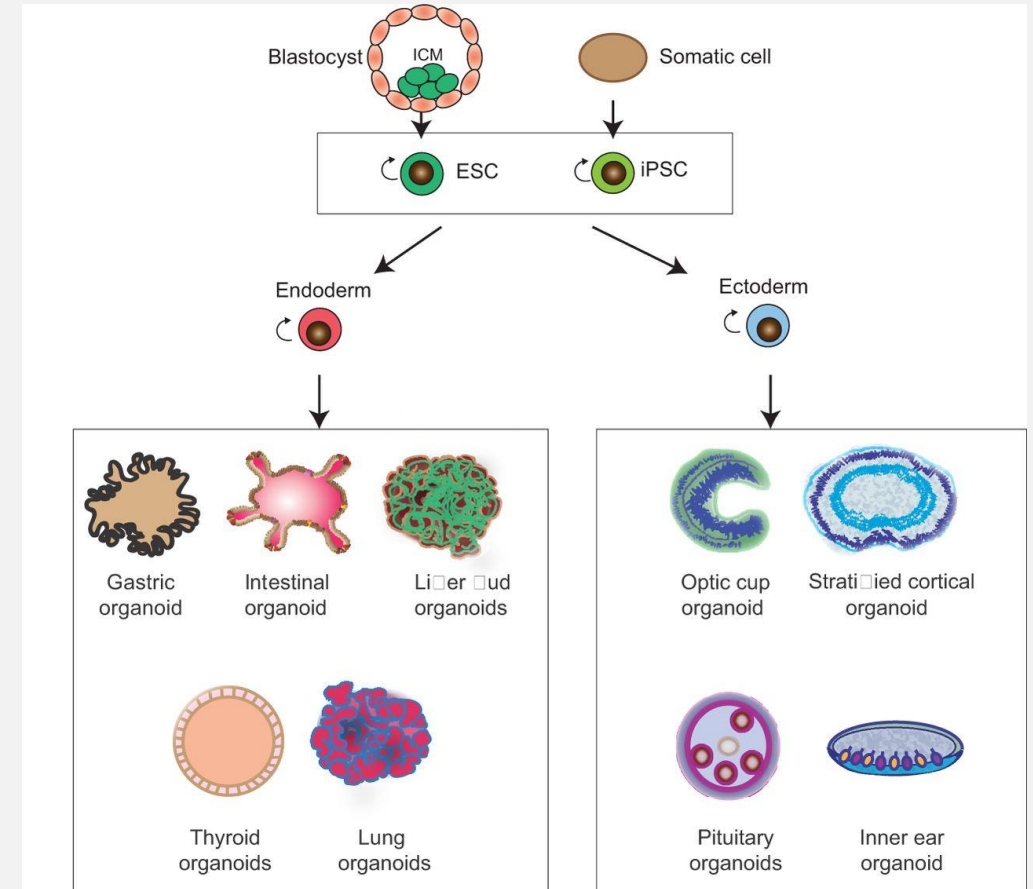
SINGLE-CELL ANALYSIS

- Sequence single cells

- Import data and perform quality control

- Calculate the normalization factors

- Determine the groups to be compared

- Find genes that have significant difference in their expression patterns

- … onwards from there to pathways etc.

STUDY DESIGN - GROUPS

genotype

	Knockout	WT
High (expression)	**High Cdh1 - KO**	**High Cdh1 - WT**
Low (expression)	Low Cdh1 - KO	Low Cdh1 - WT

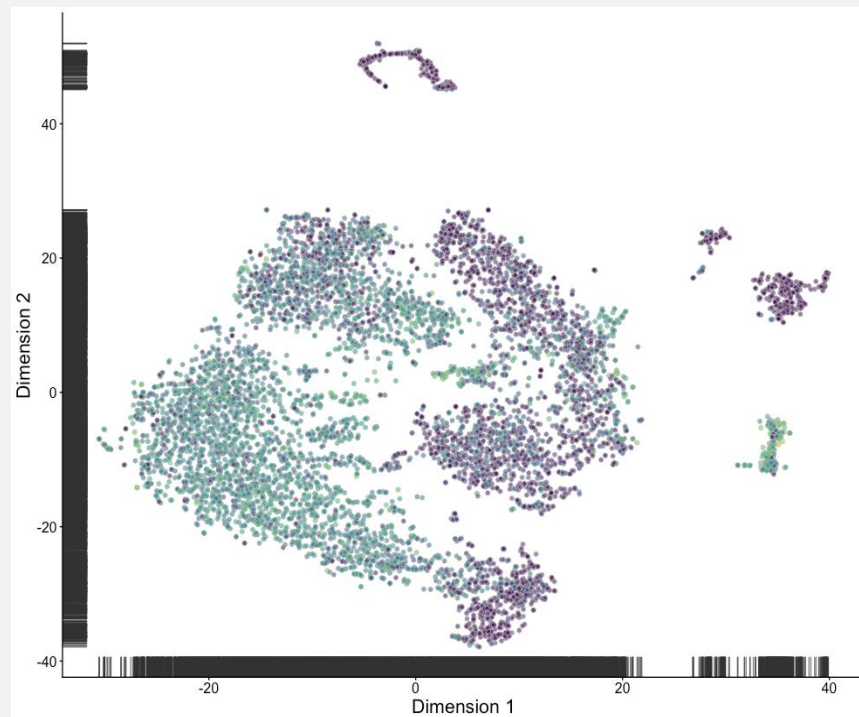# CDH1 KNOCKOUT IS NOT 100% EFFECTIVE...
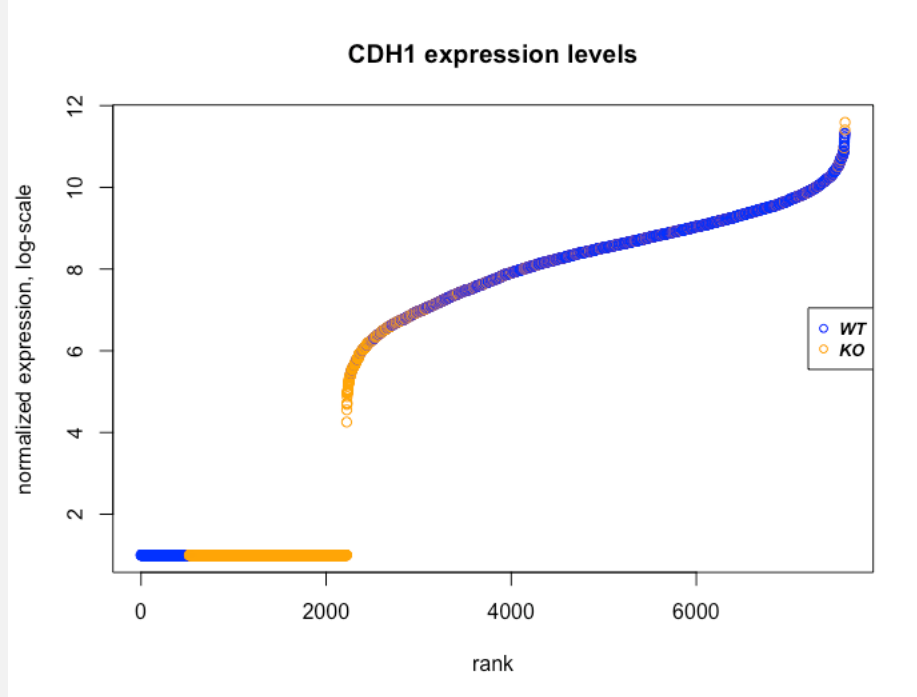
WT / **KO**
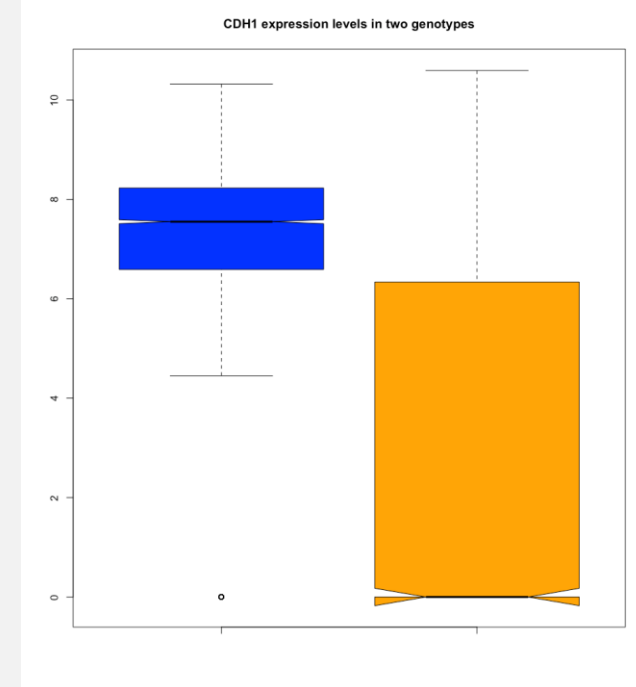
WT / **KO**
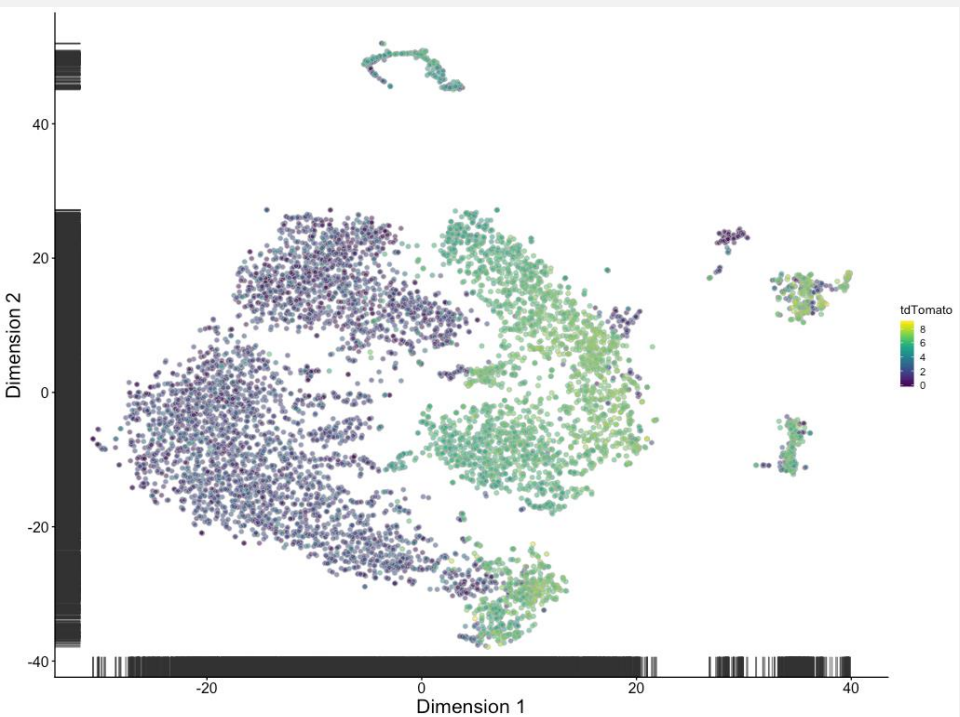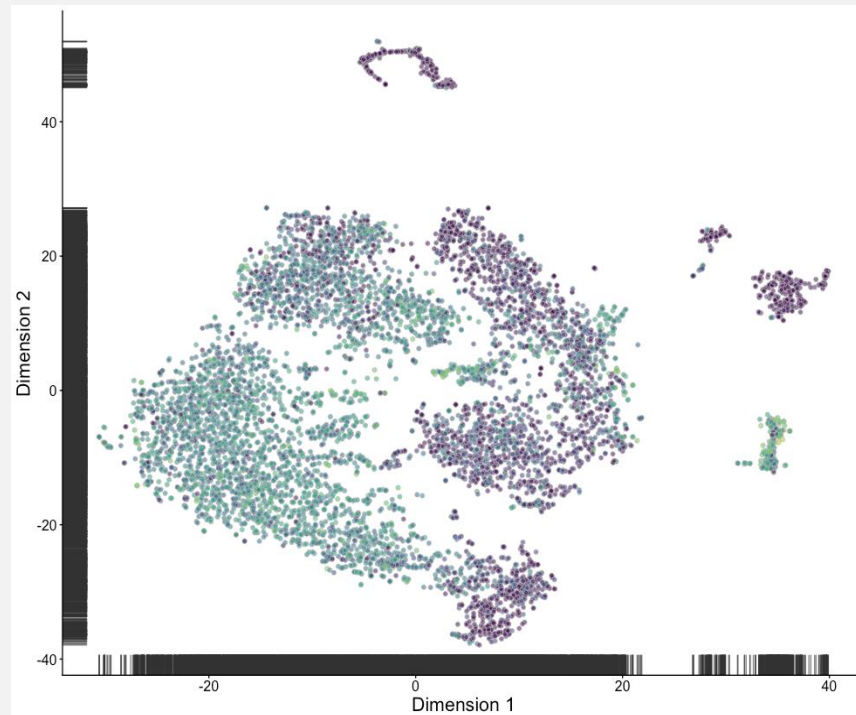
T-SNE GROUPS

CDHI expression
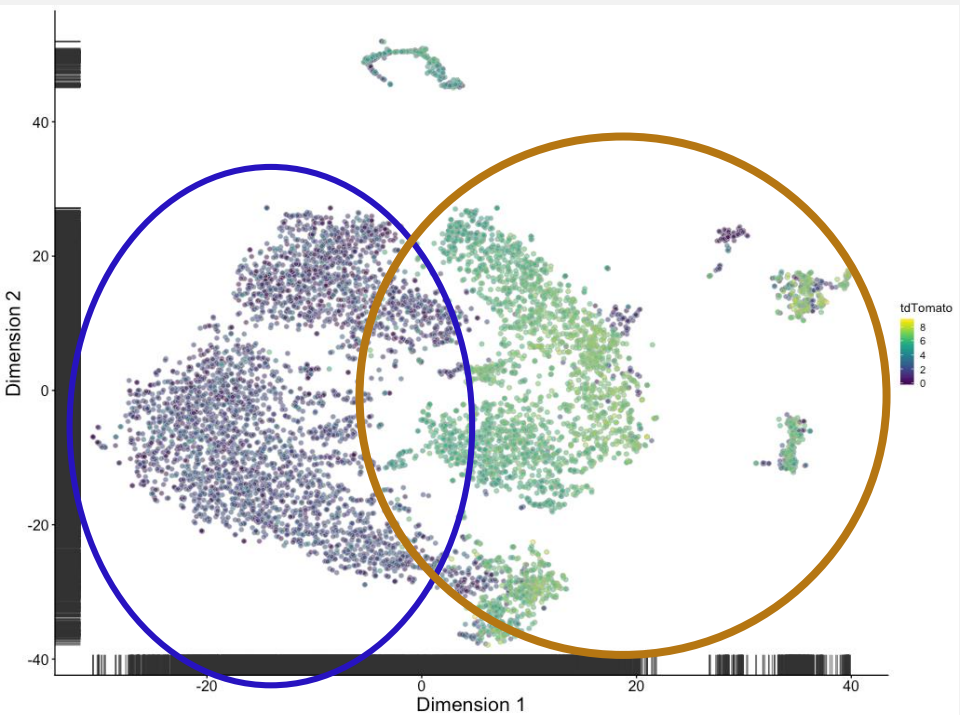
Expression of genotype marker
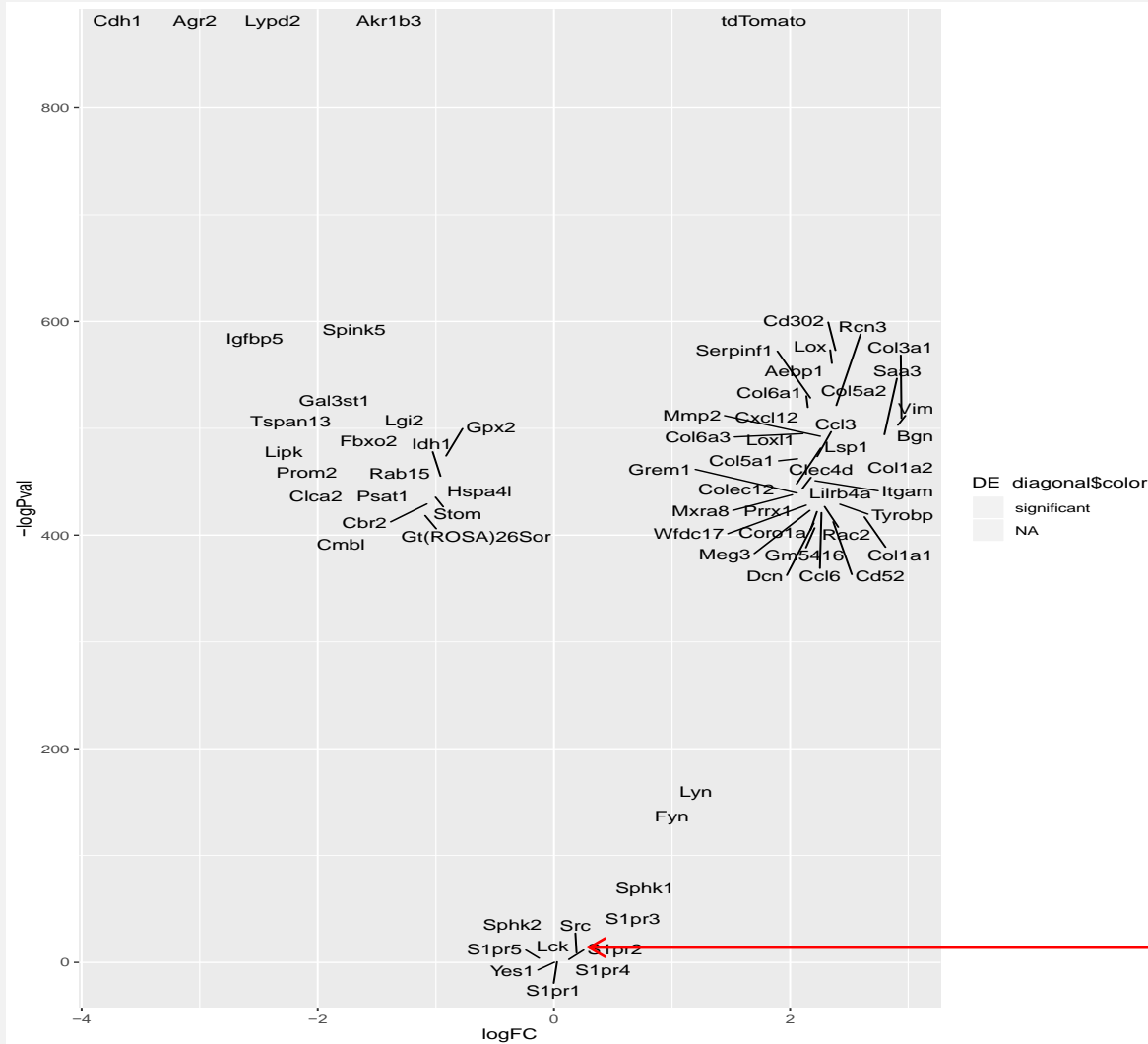
T-SNE GROUPS

CDH1 expression

Expression of genotype marker

DIFFERENTIALLY EXPRESSED GENES?

- Genes are marked red for statistical significance
- Problem – too much red

Corresponds to FDR = 0.01, standard measure for stat significance for multiple comparison

Standard RNA-seq analysis aims at complexity reduction as a way to reduce the noise…

- Single cells are treated as biological replicates – and there is a lot of cells

- "Too many replicates" results in "too much" statistical power

- This workflow is adjusted to zero-inflated data* – but the other problem remains

  *A. Lun et al., 2018; Pal, Smyth et al., 2017

- Too many significant genes – hard to pick differentially expressed pathways (most pathways appear to be enriched)

- If we had to cut off higher on the volcano plot (choose even more conservative statistical significance) – how to pick a cut-off?

- Possible approaches:

  - Repeat experiment with more true biological replicates

  - ~~Randomly select samples from each group – and compare ranking (?)~~

  - Cluster at a lowest level (to get to the cell subtype) – and treat small clusters as technical replicates

PATHWAY ANALYSIS RESULTS

PATHWAY ANALYSIS RESULTS
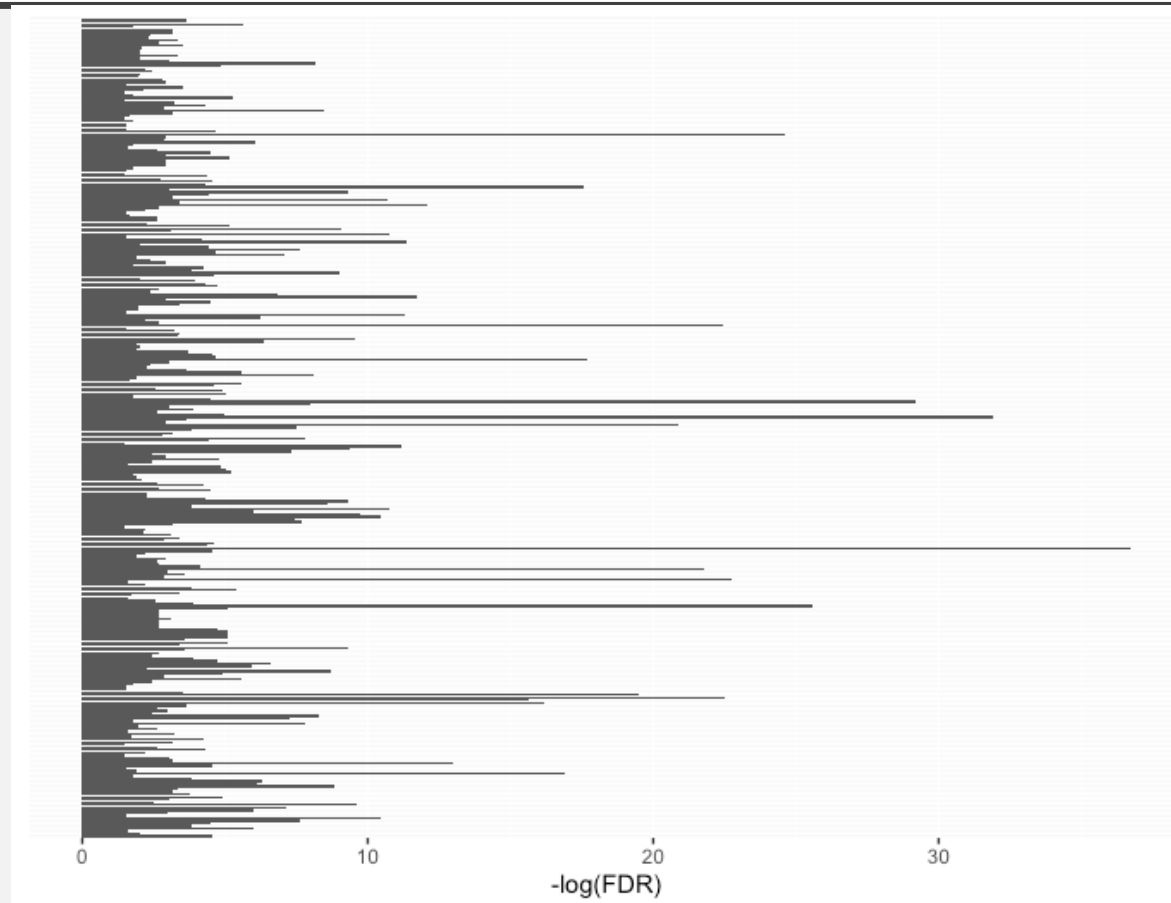
NEW WORKFLOW

- Clusterization step added

- Clusters are treated as technical replicates and merged on the low level

- New results: look better

FURTHER WORK

- Unsupervised clustering and merging replicates makes pathway analysis look better – will we get the same results when we get more data?

- We have a list of potential drug targets - the disruption of membrane-associated cytoskeleton leads to abnormal cell survival signaling

- Candidate drugs are tested on organoids

ACKNOWLEDGEMENTS

CENTRE FOR
Translational Cancer Research
Te Aho Matatū
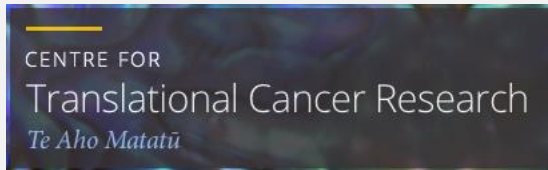
- Parry Guilford
- Tanis Godwin
- Tom Brew
- Mik Black

hrc Health Research
Council of
New Zealand

BC CAN CER MOLECULAR ONCOLOGY RESEARCH
Provincial Health Services Authority

David Huntsman

@ Huntsman lab
University of British
Columbia, Vancouver,
CA

CLUSTERS

- Number of clusters is selected in agreement with most indices results (max number taken)

- **Samples that fall in one cluster are treated as technical replicates**

CDH1 MUTATION MAP

E-cadherin mutations in diffuse gastric cancer

Validation of *CDH1* germline missense mutations

Table 2 E-cadherin germline missense mutations used for the statistical analysis

Variant	<1%	Co-segregation	Recurrence	SIFT[a]	Functional effect
Neutral (N)[b]	–	–	–	–	–
T118R	+	–	–	–	+
L214P	+	–	–	+	+
G239R	+	–	–	+	+
A298T	+	–	–	–	+
T340A	+	–	+	–	+
W409R	+	–	–	+	+
P429S	+	–	–	+	+
A592T	–	–	+	–	+
A617T	+	–	–	–	–
A634V	+	–	+	–	+
R732Q	+	–	–	+	+
P799R	+	–	–	+	+
V832M	+	+	–	+	+

Suriano G et al. J Mol Med 84:1023, 2006

CDH1 IN CANCER

E-cadherin & gastric cancer
(sporadic and hereditary)

2010/.../14

2009
- Second hit of *CDH1* inactivation

2008
- NMD mRNA surveillance downregulates aberrant *CDH1* transcripts

2007
- E-cadherin repressors in gastric ancer
- Novel germline *CDH1* mutations
- Experimental model in Drosophila

2006
- Cleft lip/palate and *CDH1* mutations in families with HDGC

2005
- First Portuguese family with HDGC

2004
- Model of development of HDGC

2003
- Functional analyses of E-cadherin (*CDH1*) germline missense mutations

2001
- Prophylactic gastrectomies in asymptomatic carriers of germ-line E-cadherin mutations
- Second hit E-cadherin gene (*CDH1*) inactivation (promoter methylation) in sporadic diffuse gastric carcinoma

1998/9
- Familial gastric cancer: overview and guidelines for management. (IGCLC)

Int J Surg Pathol 6: 135, 1998
Histopathology 35: 477, 1999
Lab Invest 79: 459, 1999
J Med Genet 36: 873, 1999
N Engl J Med 344:1904, 2001
Oncogene 20: 1525, 2001
Gastroenterol Clin Biol 25: 931, 2001
Hum Mutat 19:510, 2002
Hum Mol Genet 12: 575, 2003
Hum Mol Genet 12: 3007, 2003
Oncogene 22:5716, 2003
J Pathol 203: 681, 2004
Virchows Arch 446: 18, 2005
Clin Cancer Res 11:5401, 2005
J Med Genet 43:138, 2006
J Mol Med 84:1023, 2006
Hum Mol Genet15:1704, 2006
Hum Mutat 28:203, 2006
J Pathol 211:507, 2007
Oncogene 27: 4255, 2008
J Clin Pathol 61:25, 2008
J Pathol 216:295, 2008
Gastroenterology 136:2137, 2009
J Med Genet 47: 436, 2010
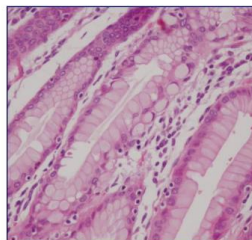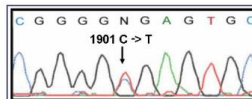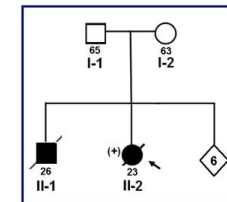PLoS One 6:e23188 ,2011
Cell Mol Life Sci, 2011

Validation of *CDH1* germline missense mutations

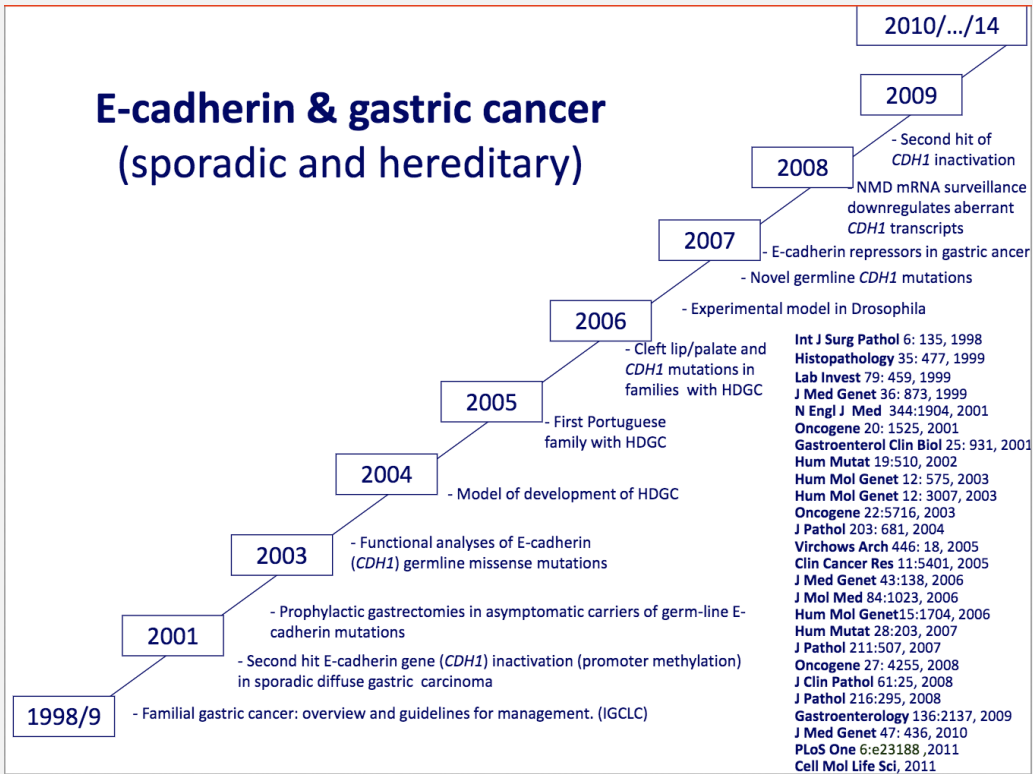**Table 2** E-cadherin germline missense mutations used for the statistical analysis

Variant	<1%	Co-segregation	Recurrence	SIFT[a]	Functional effect
Neutral (N)[b]	–	–	–	–	–
T118R	+	–	–	–	+
L214P	+	–	–	+	+
G239R	+	–	–	+	+
A298T	+	–	–	–	+
T340A	+	–	+	–	+
W409R	+	–	–	+	+
P429S	+	–	–	+	+
A592T	–	–	+	–	+
A617T	–	–	+	–	–
A634V	+	–	+	–	+
R732Q	+	–	–	+	+
P799R	+	–	–	+	+
V832M	+	+	–	+	+

Suriano G *et al.* J Mol Med 84:1023, 2006

1901 C -> T