Who am I? **Quaid Morris**

Background:

BSc: CompSci (U of T),

PhD: Machine Learning & CompNeuro (MIT & Gatsby)

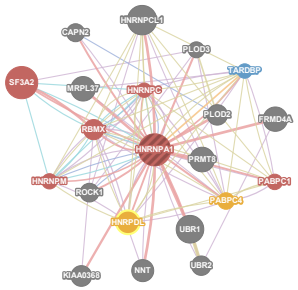
Currently:

Donnelly Centre, Faculty of Medicine, U of T,

Core Expertise:

Machine learning methodology development, and biological and clinical data analysis

Big data

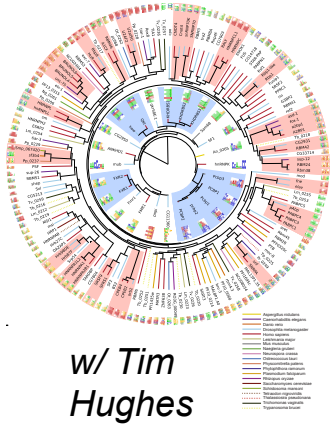


w/ Gary Bader

GeneMANIA: Predicting gene function by big data integration

>100k users
has all of the interaction network data ever generated

Genomics



w/ Tim Hughes

RNA-binding proteins and PTR

Nature 2013

ARTICLE

doi:10.1038/nature12311

A compendium of RNA-binding motifs for decoding gene regulation

Debashish Ray[1*], Hilal Kazan[2*], Kate B. Cook[2*], Matthew T. Weirauch[1*†], Hamed S. Najafabadi[1–4*], Xiao Li[1], Serge Gueroussov[2], Mihai Albu[1], Hong Zheng[1], Ally Yang[1], Hong Na[1], Manuel Irimia[1], Leah H. Matzat[1], Ryan K. Dale[1], Sarah A. Smith[5], Christopher A. Yarosh[5], Seth M. Kelly[1], Behnam Nabet[6], Desirea Mecenas[6], Weimin Li[6], Rakesh S. Laishram[6], Mei Qiao[10], Howard D. Lipshitz[6], Fabio Piano[6], Anita H. Corbett[7], Russ P. Carstens[8], Brendan J. Frey[9], Richard A. Anderson[5], Kristen W. Lynch[6], Luiz O. F. Penalva[10], Elissa P. Lei[5], Andrew G. Fraser[2,3], Benjamin J. Blencowe[1,3], Quaid D. Morris[1,2,3,4] & Timothy R. Hughes[2,3]

Cancer

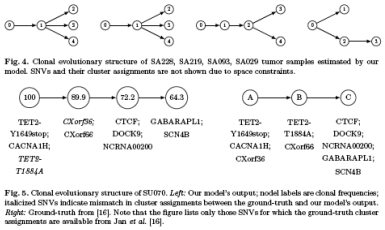


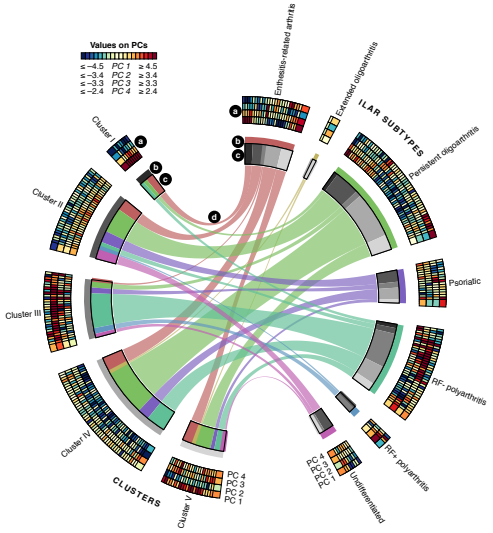
Fig. 4. Clonal evolutionary structure of SA228, SA219, SA093, SA029 tumor samples estimated by our model. SNVs and their cluster assignments are not shown due to space constraints.

Fig. 5. Clonal evolutionary structure of SU070. *Left*: Our model's output; nodal labels are clonal frequencies; italicized SNVs indicate mismatch in cluster assignments between the ground-truth and our model's output. *Right*: Ground-truth from [16]. Note that the figure lists only those SNVs for which the ground-truth cluster assignments are available from Jan et al. [16].

w/ Lincoln Stein, Paul Boutros OICR

Subclonal evolution in cancer

Immunology and genomics



w/ Rae Yeung, SickKids

Re-definition of childhood arthritis using biological and clinical data

Clinical

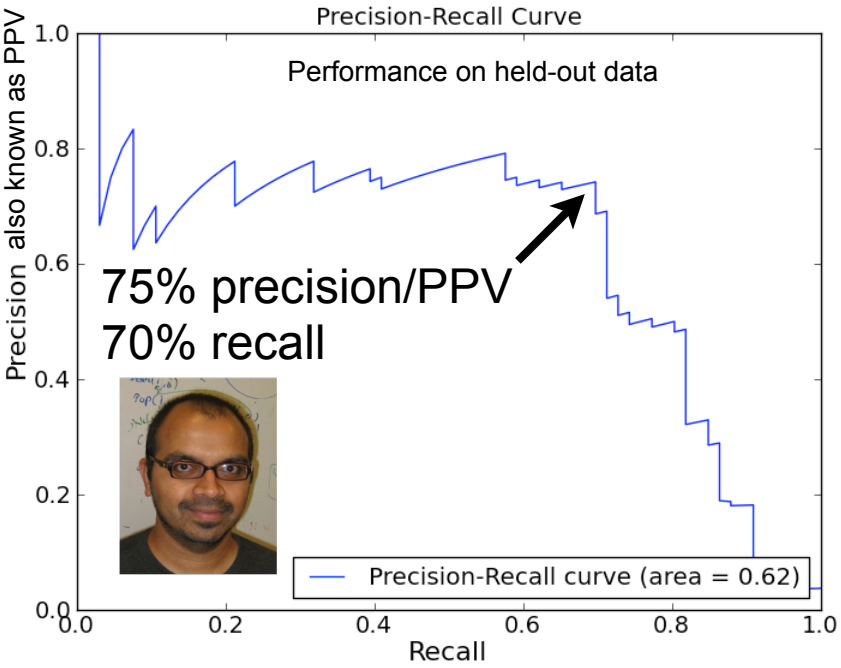
PhD thesis: a clinical diagnostic support system to predict risk of 600 diseases given any of >4000 clinical findings

Using the Primary Care Electronic Medical Records to Identify Complex Patients

w/ **David Frost, MD, FRCP(C)**
Clinical Teaching Unit and Site Director, Toronto Western Hospital, Division of General Internal Medicine

Background: 1% of users consume 50% of health care resources (5% consume 85%)
Goal: identify 'high users' from EMRs before they become high users

Data: ~150 high users identified at Toronto Western (3+ ED visits), ~9k unstructured Western family practice notes from Practice Solutions from: health history, personal traits, active treatments, active problem list. patients > 50yr.
Approach: L1-regularized logistic regression (LASSO) on word frequencies



Words predictive of high use

Disease	Fold enrichment	Social	Fold enrichment	Other	Fold enrichment
thrombocytopenia	35.06	shelter	17.78	admissions	15.86
laminectomy	35.06	odsp (10% of high users)	10.58	multiple (7%)	13.33
tumour	30.3	parents	7.57	cas	13.19
infections	26.67	husband (5%)	6.59	supports	11.8
stent	26.67	estranged	5.31	help	8.88
				abuse (9%)	8.83
				attempt	6.23

Words appear in less than 3% of high users except as noted.