A Proteomic Survey of the Postnatal Human Brain

Becky Carlyle, Associate Research Scientist
Nairn Lab
NIDA proteomics 06/02/17

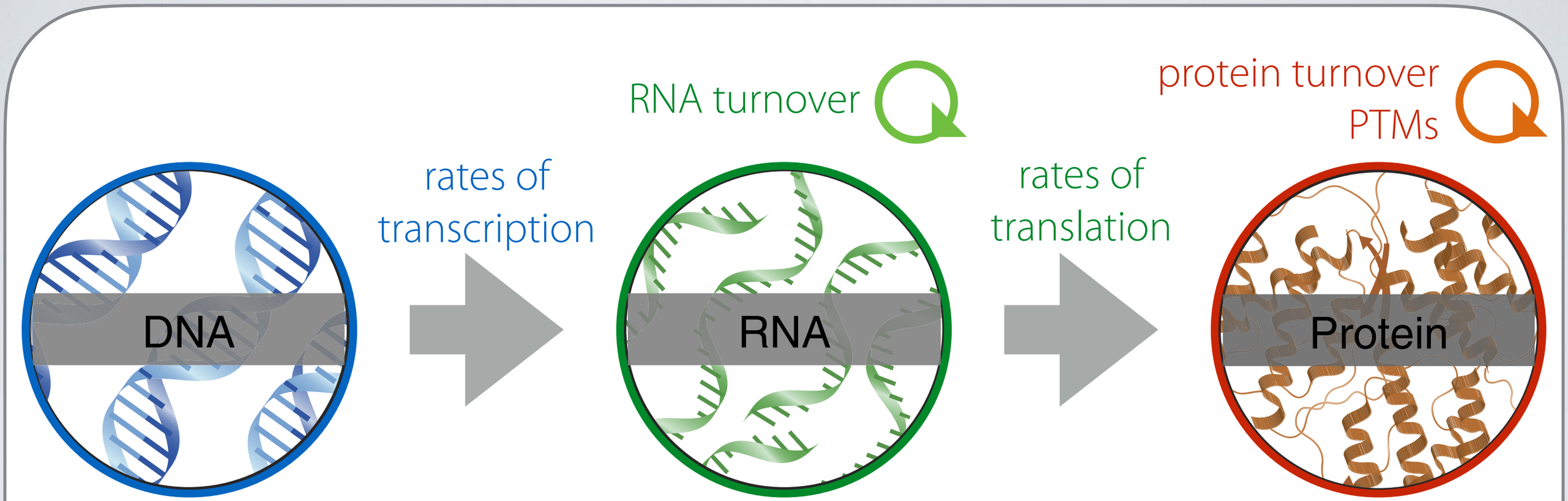
Talk outline

- Why proteomics?
- BrainSpan/psychENCODE project samples
- Proteomic methods
- Results
- Conclusions

Talk outline

- **Why proteomics?**
- BrainSpan/psychENCODE project samples
- Proteomic methods
- Results
- Conclusions

Protein abundance is the final output of the central dogma

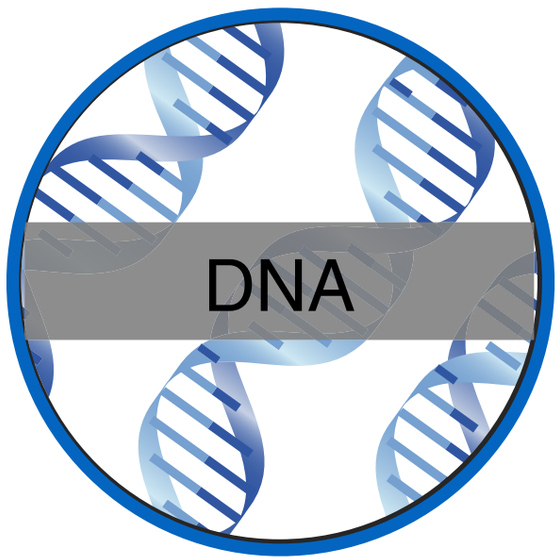


Protein abundance is the final output of the central dogma

RNA turnover

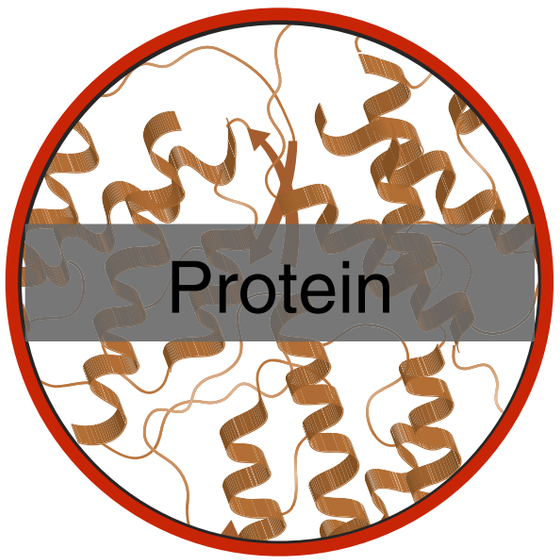
protein turnover

PTMs



rates of transcription

rates of translation



structural variants / SNPs
promoter usage / splicing
epigenetics / imprinting

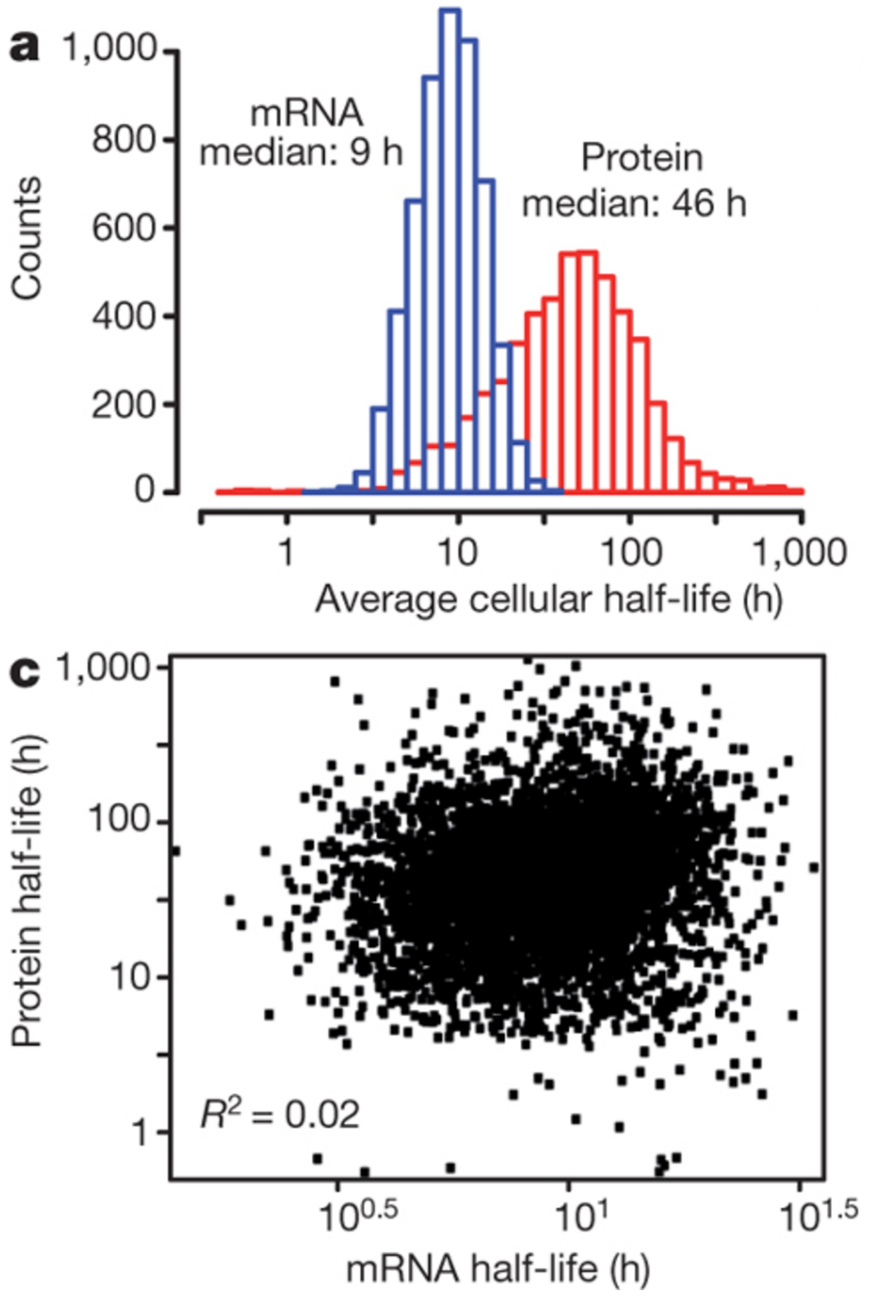
RNA abundance

RNA editing
miRNA repression
NMD

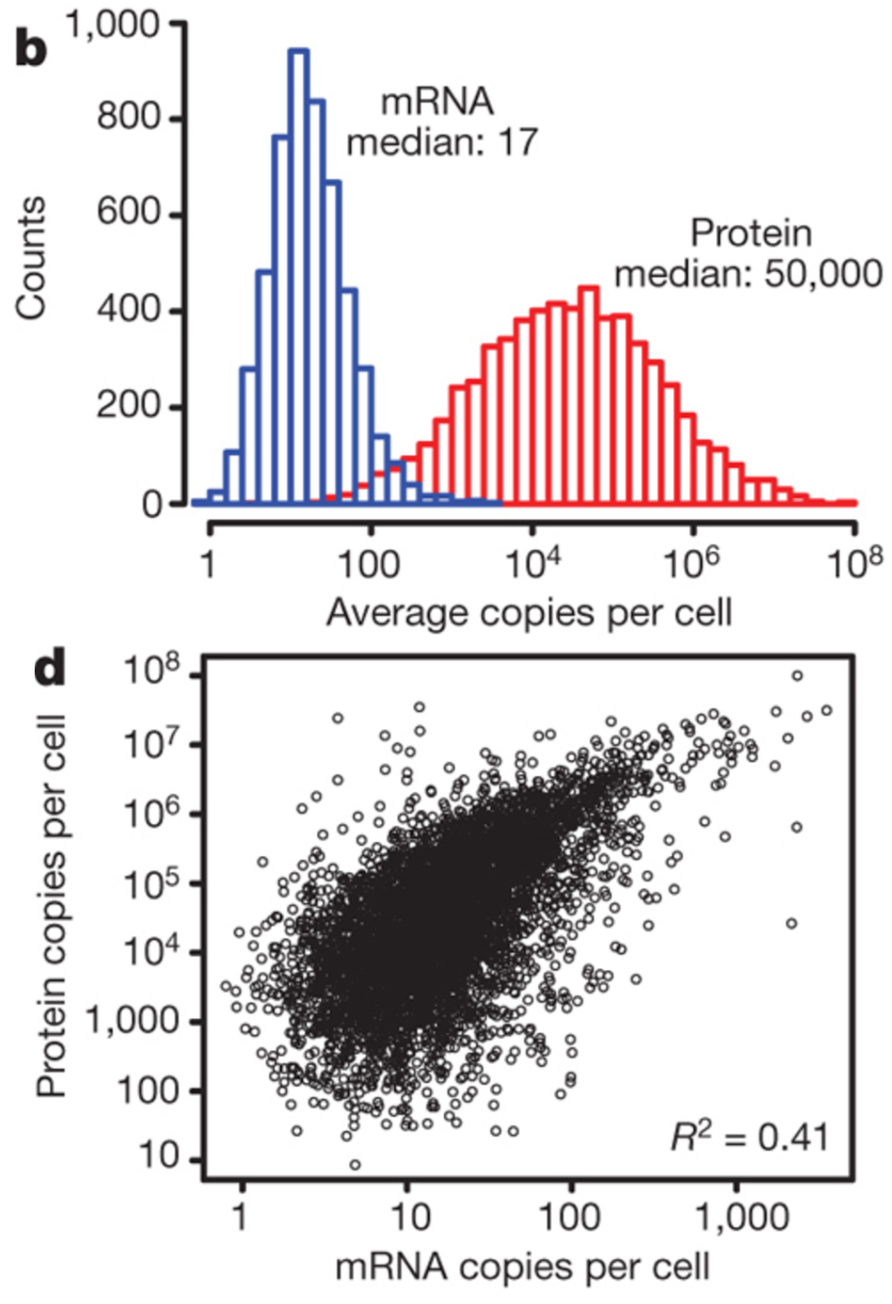
protein abundance

Protein and mRNA may correlate poorly

In both ***turnover*** rates…

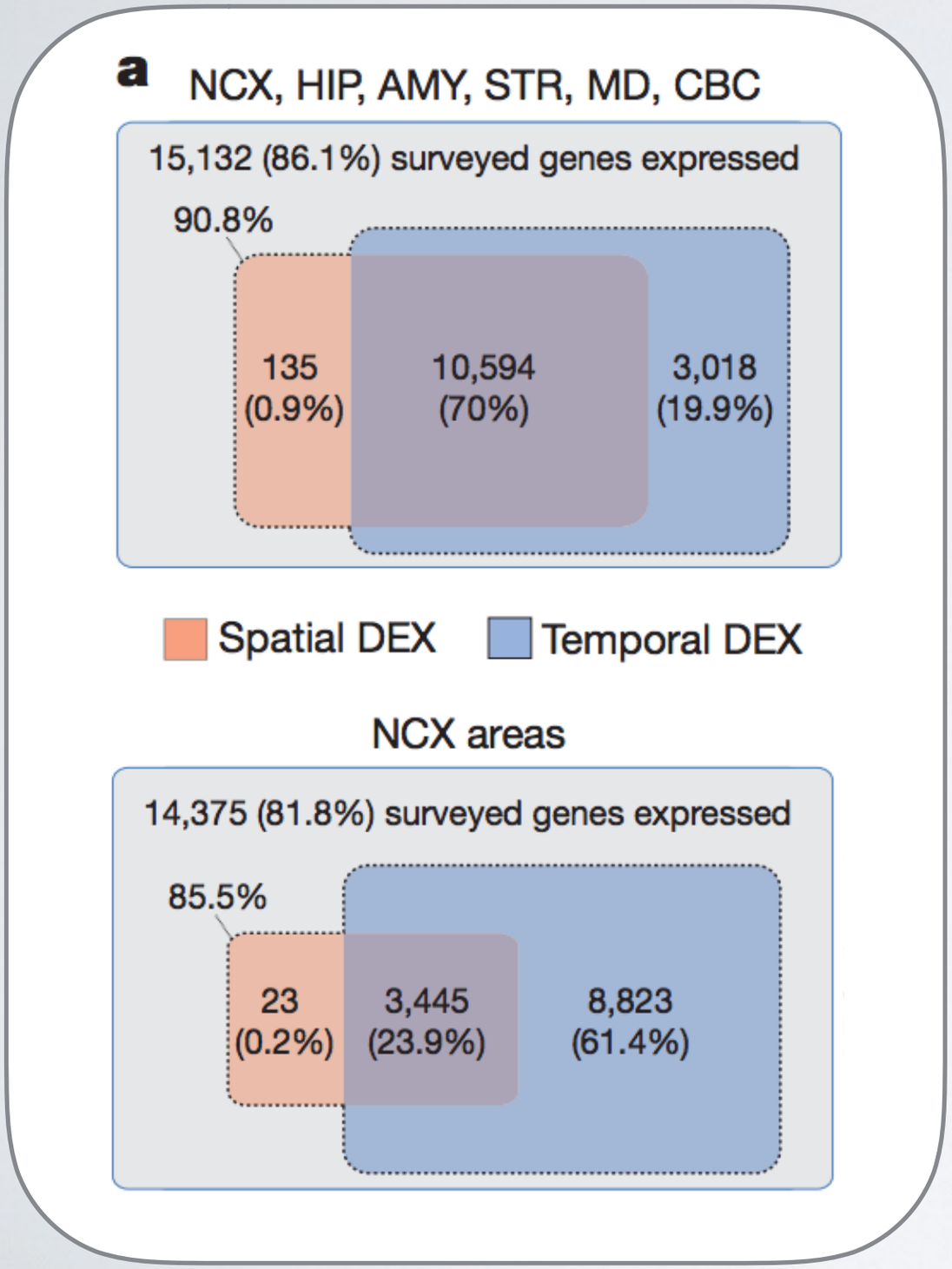


… and ***abundance***



Schwanhausser et al., Nature, 2011

Human brain - mRNA abundance



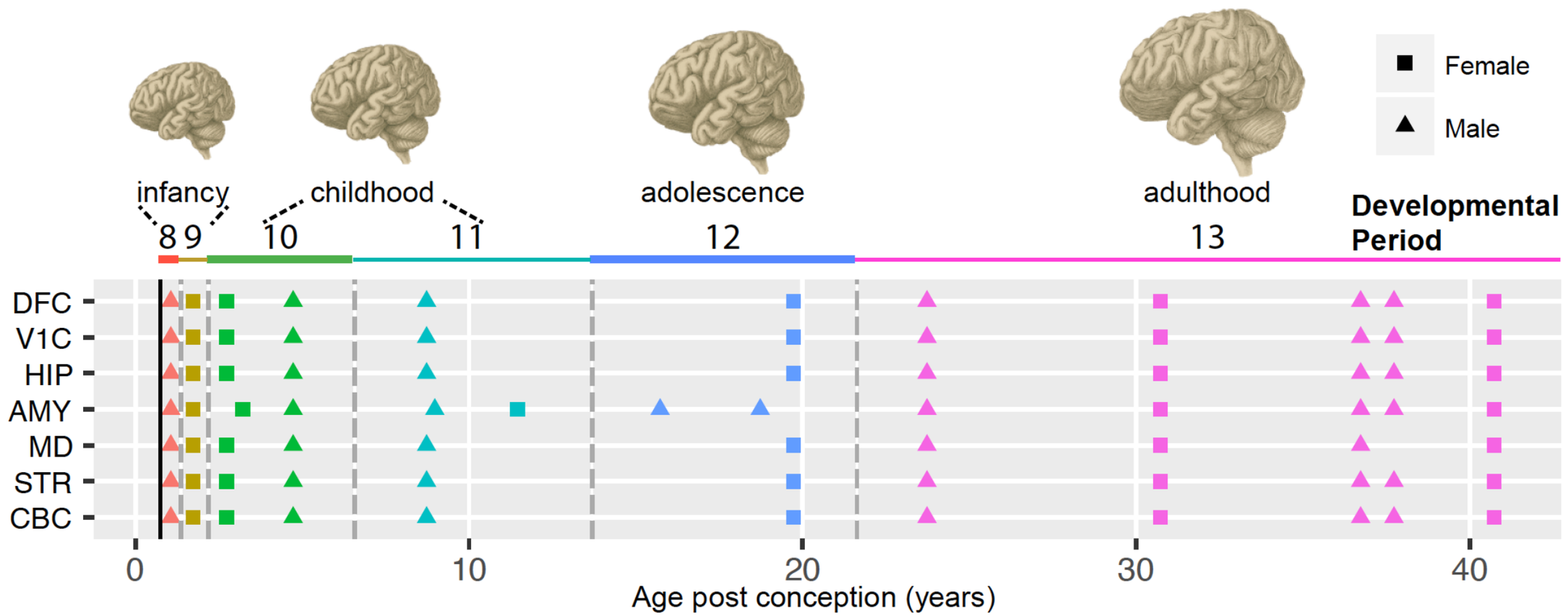
- The vast majority of mRNA variation between regions comes during development
- Fewer genes differ in abundance between cortical regions in the adult
- Given the wider range in protein half lives and abundance, there may be differences in protein between these regions that are not reflected in mRNA

Kang et al., Nature, 2011

Talk outline

- Why proteomics?
- **BrainSpan/psychENCODE project samples**
- Proteomic methods
- Results
- Conclusions

BrainSpan/psychENCODE project samples



Same subjects used in BrainSpan for RNA-seq

~6 subjects spanning postnatal development

5 adult subjects, 7 brain regions

Talk outline

- Why proteomics?
- BrainSpan/psychENCODE project samples
- **Proteomic methods**
- Results
- Conclusions

Simple proteomic workflow

Human brain dissection

Homogenize by sonication

Proteins solubilised in 8M Urea

+ DTT, IAN

Proteins reduced & denatured

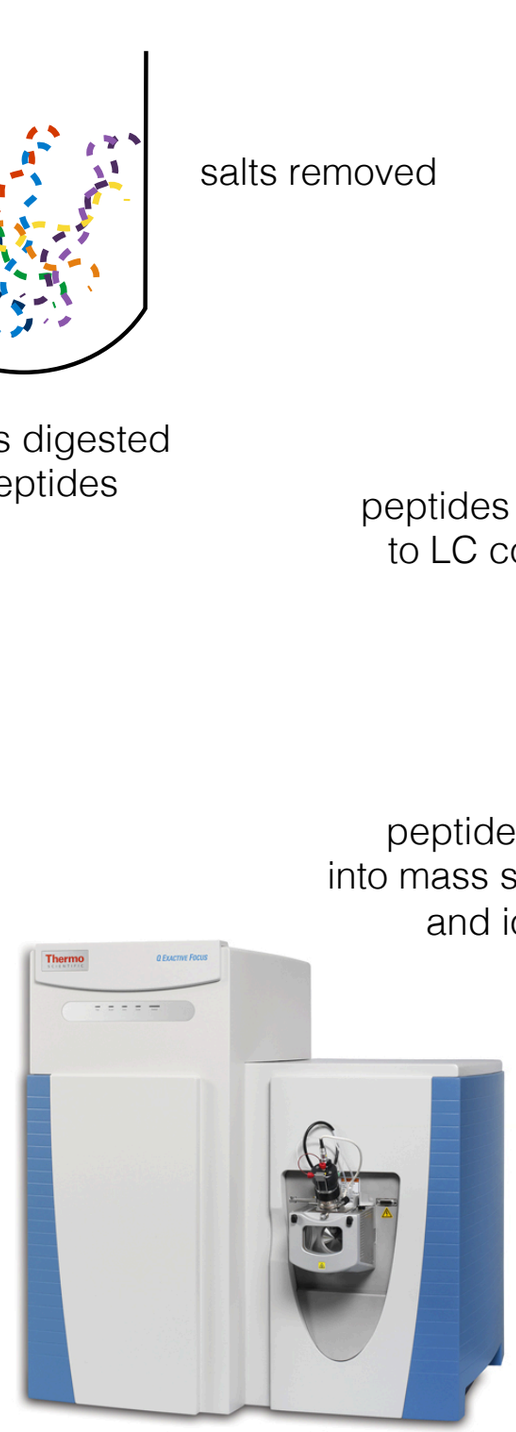
+ trypsin

Proteins digested to peptides

salts removed

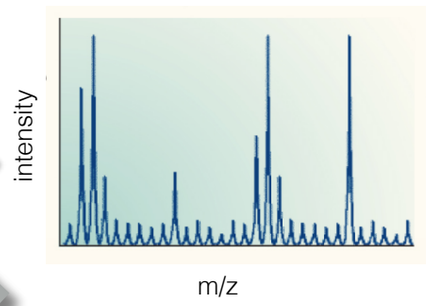
peptides loaded to LC column

peptides **eluted** into mass spectrometer and ionised



MS1

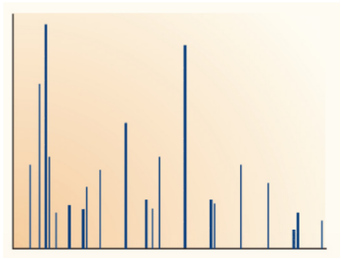
Precursor intensity
QUANTIFICATION



MS2

MS1 selected for further fragmentation

Fragmentation spectra
IDENTIFICATION
through database matching

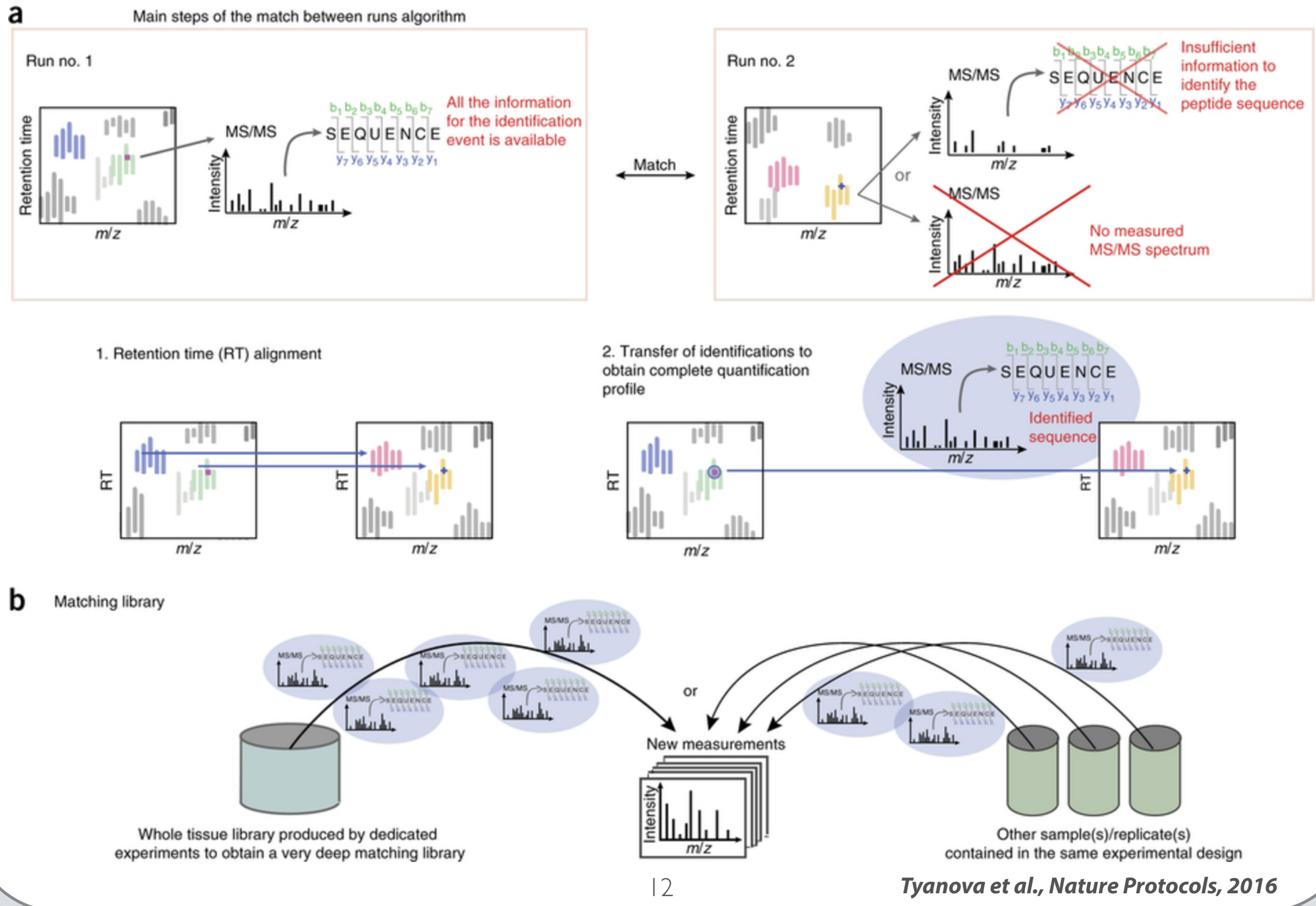


Downstream analysis

MBR

MS1 not selected - may be feature matched

Match between runs feature

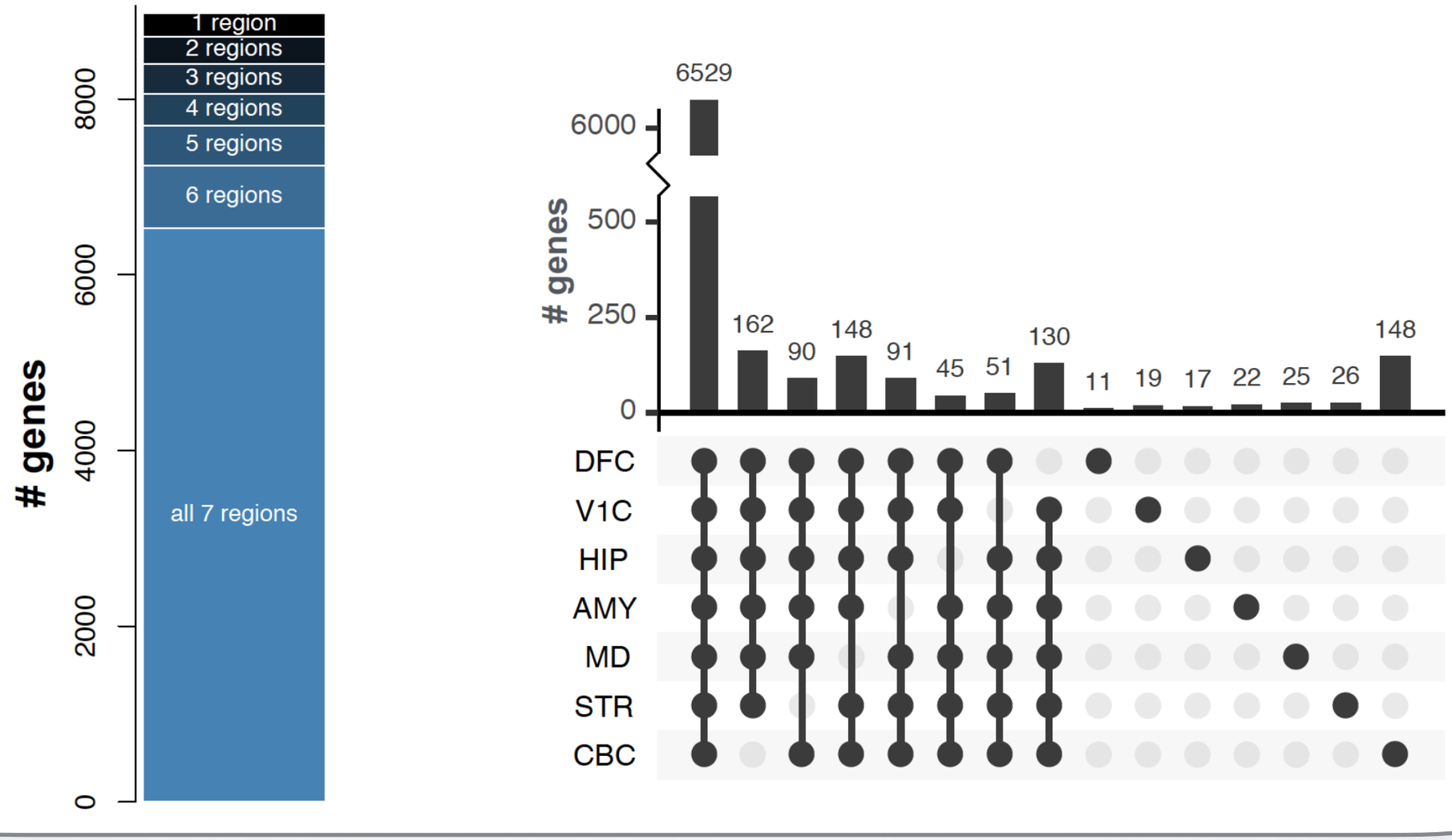


Tyanova et al., Nature Protocols, 2016

Talk outline

- Why proteomics?
- BrainSpan/psychENCODE project samples
- Proteomic methods
- **Results**
- Conclusions

Fractionated regions

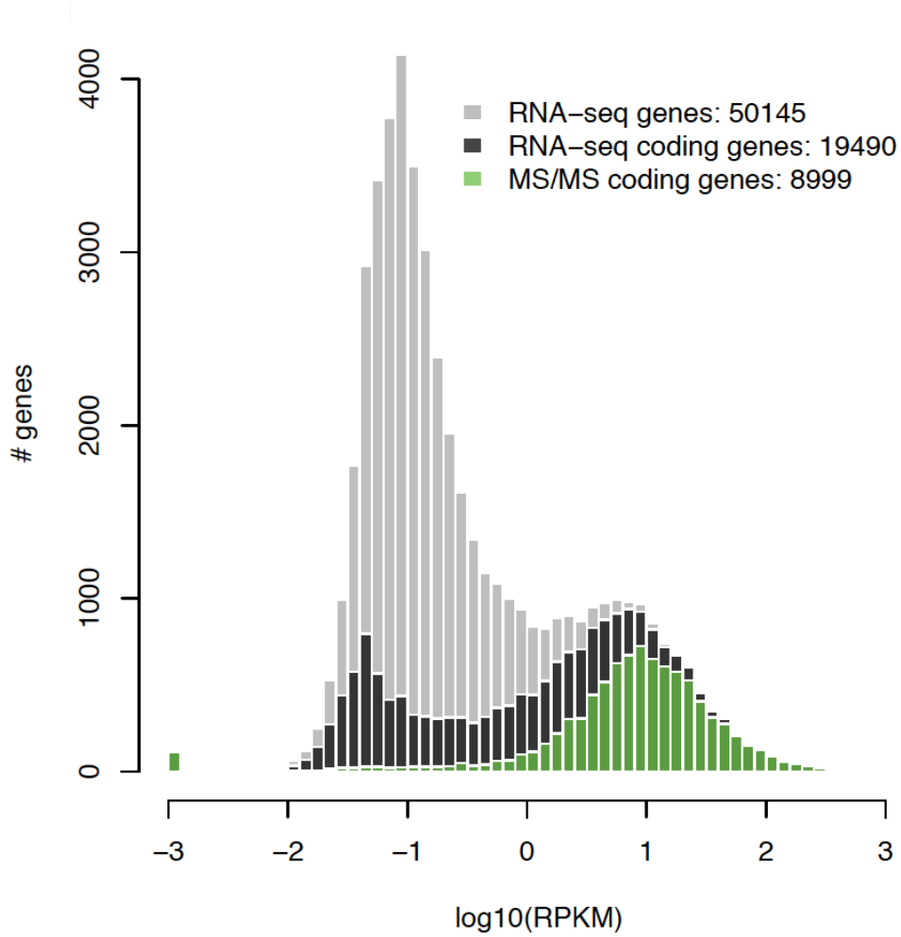


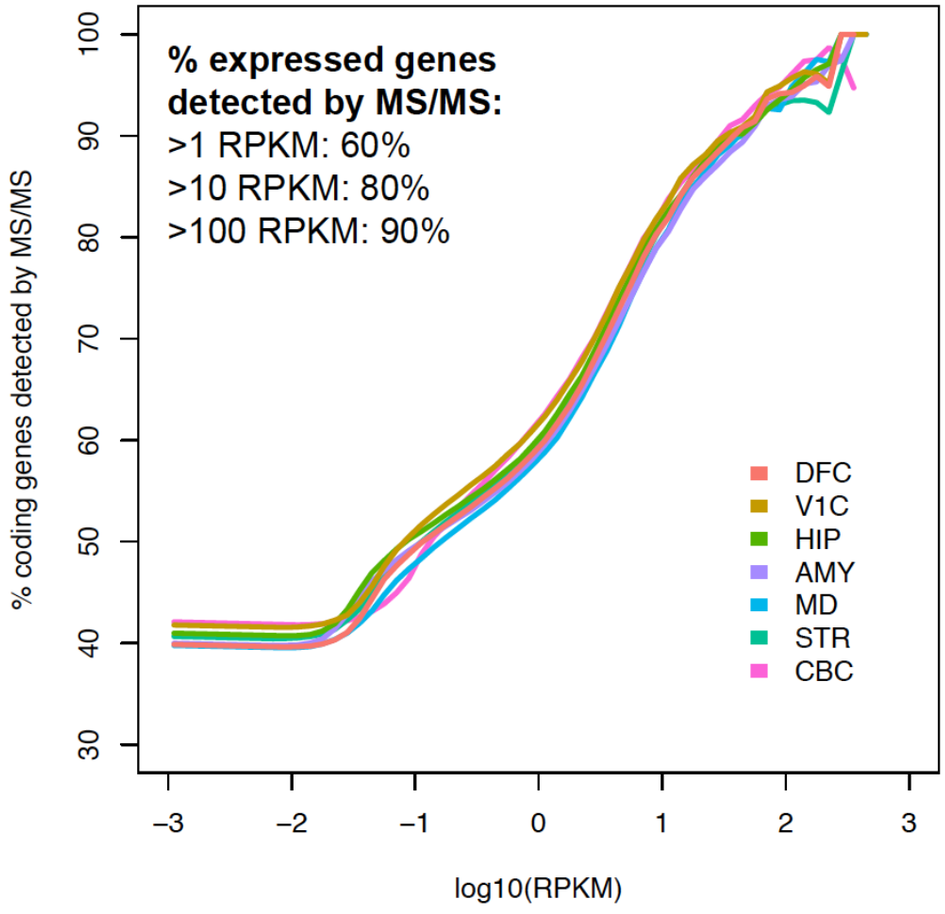
Total protein numbers are comparable to other studies showing ~11,000 proteins (in relatively simpler mixtures)

Most proteins are common to all regions

The cerebellum is a clear outlier

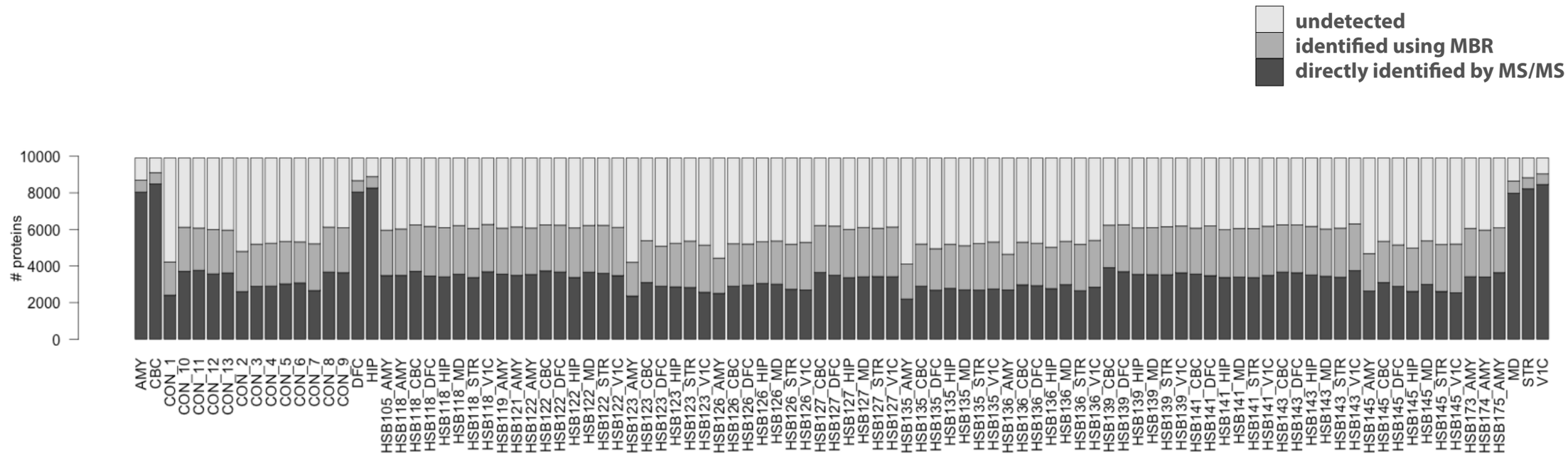
Fractionated regions - Comparison with RNA-seq





Unsurprisingly, coverage improves the more abundant a gene is

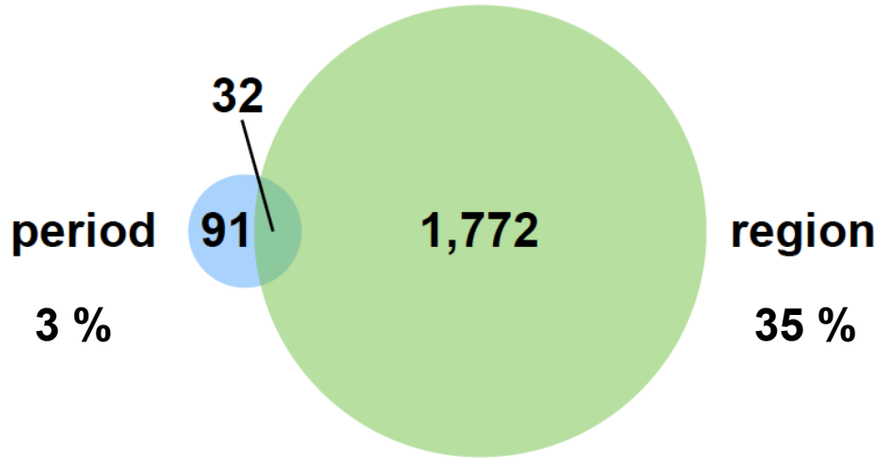
Single shot samples



On average, without matching between runs, we detect 3612 proteins per sample

Using match between runs significantly increases the number of protein IDs in the single shot samples by up to 50%

Single shot data: DEX genes - regions

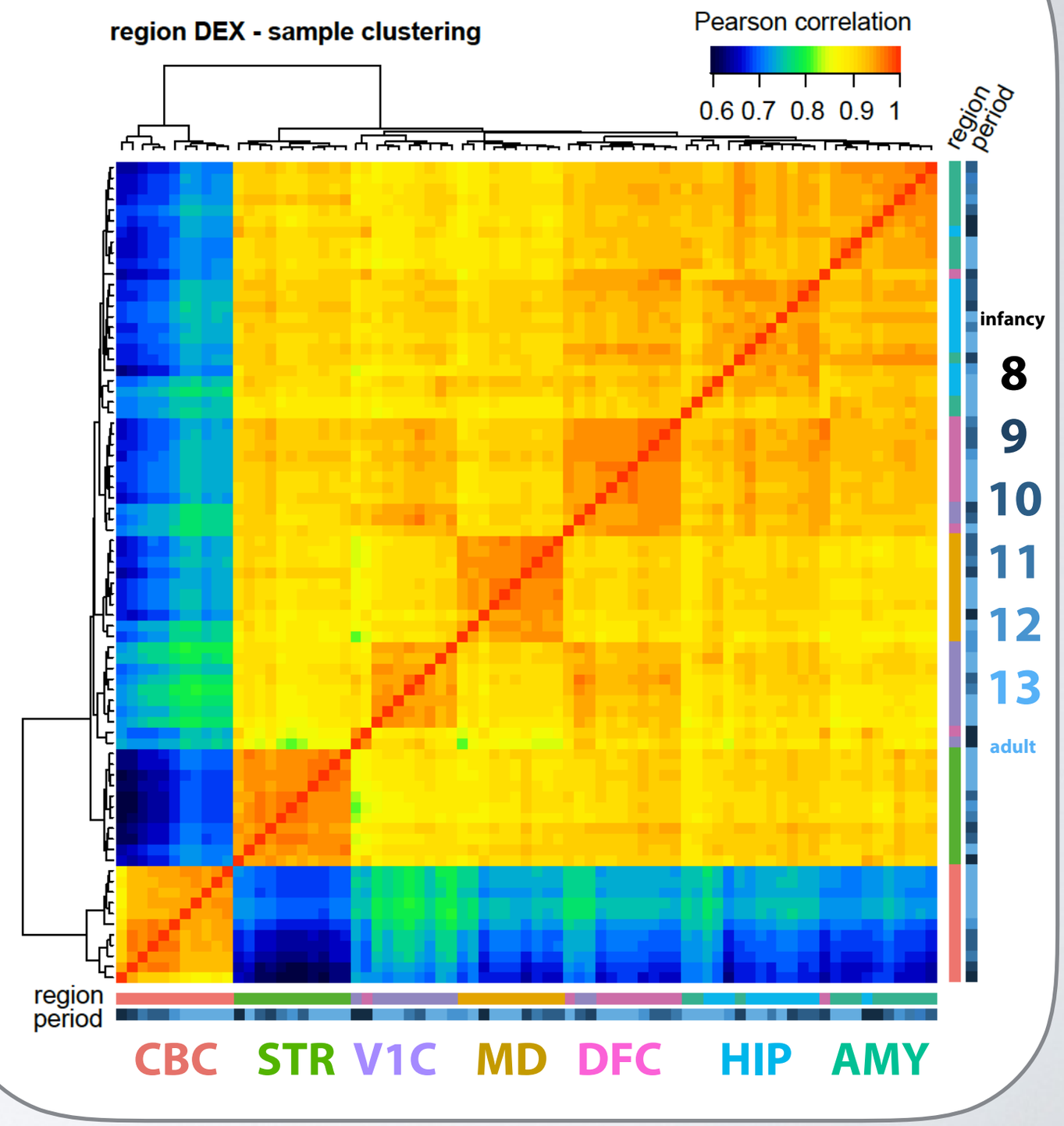


5151 proteins were reliably quantified

Samples were clustered on the basis of DEX genes

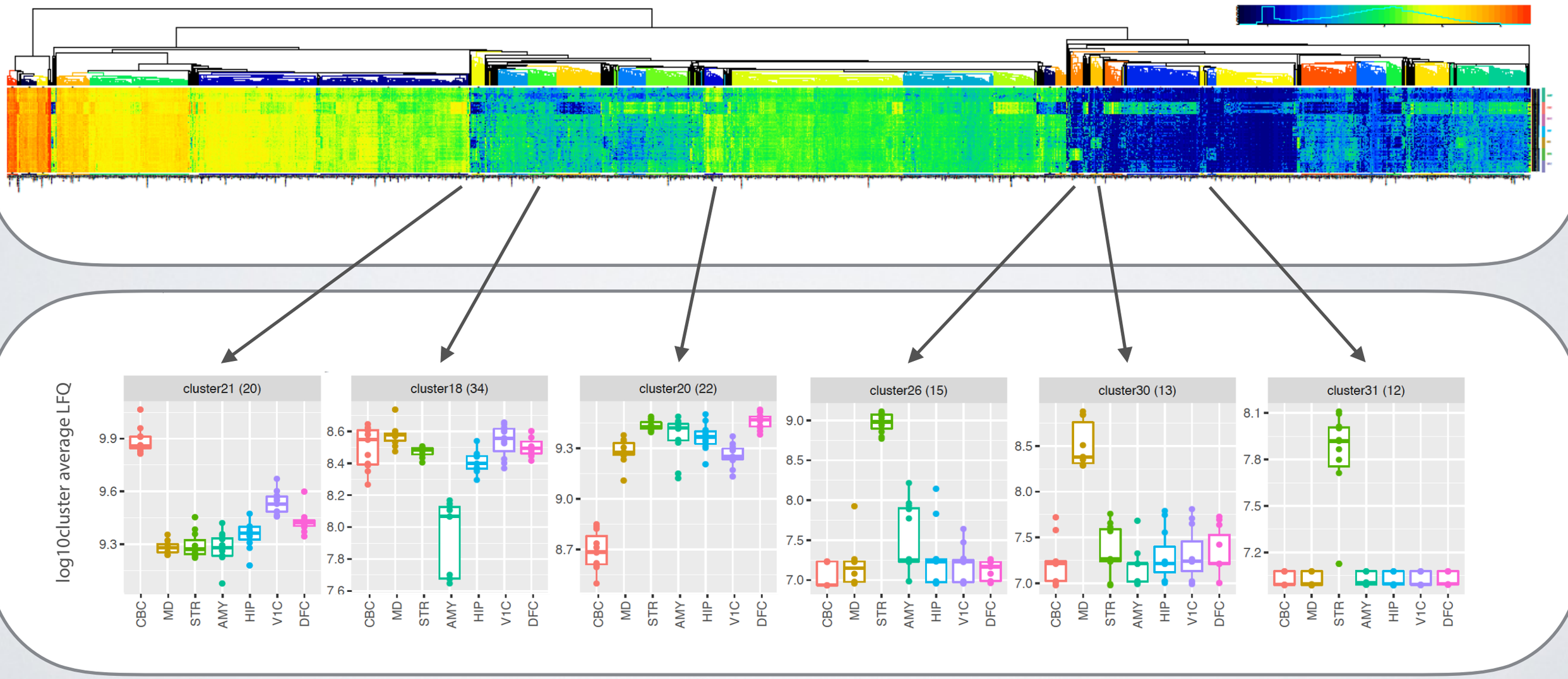
Cerebellum and striatum are clearly defined by their proteome in both development and adulthood

These two regions are markedly more homogenous with regards to cell type than the other 5 regions



Single shot data: DEX - genes

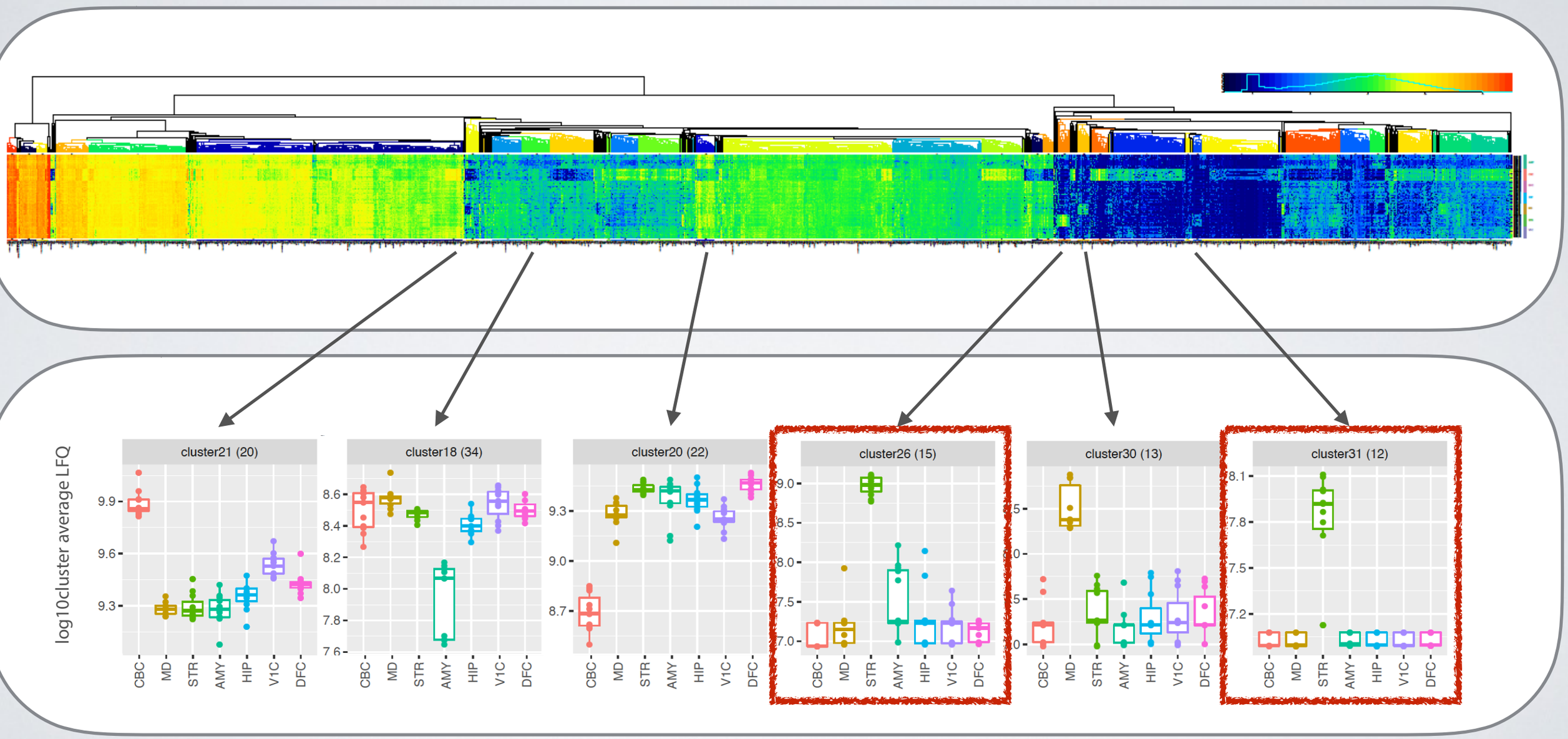
32 region DEX clusters



Single shot data: DEX - genes

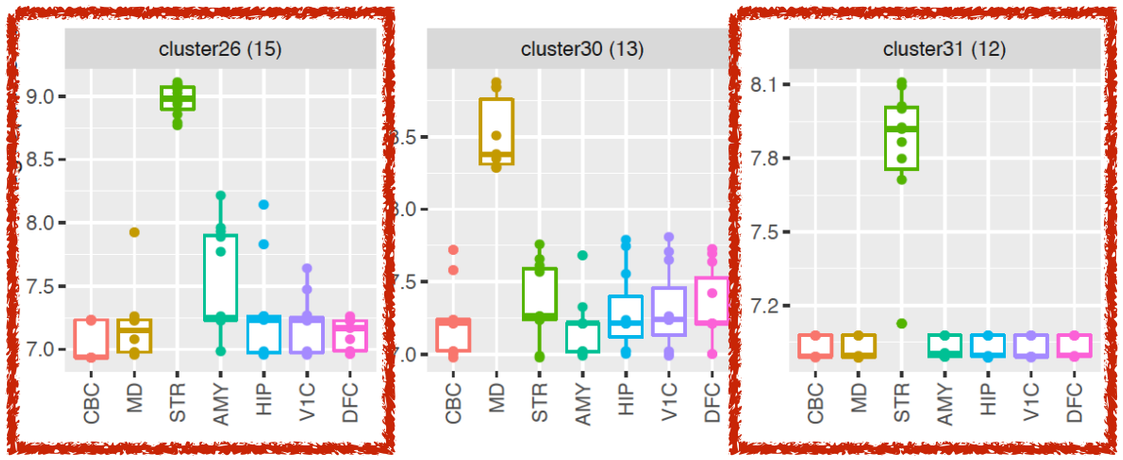
32 region DEX clusters

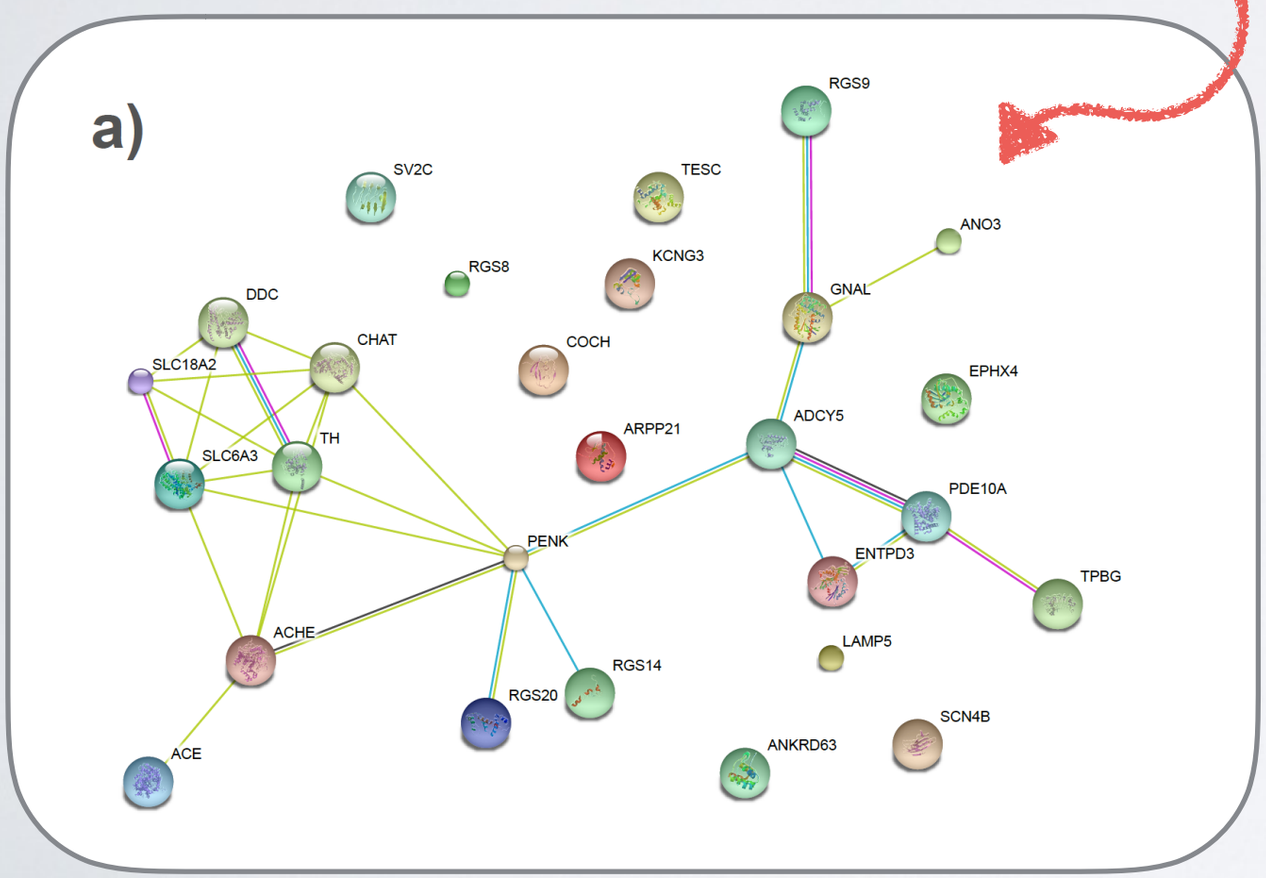
9 period DEX clusters



Striatally enriched clusters

DEX genes - clustering

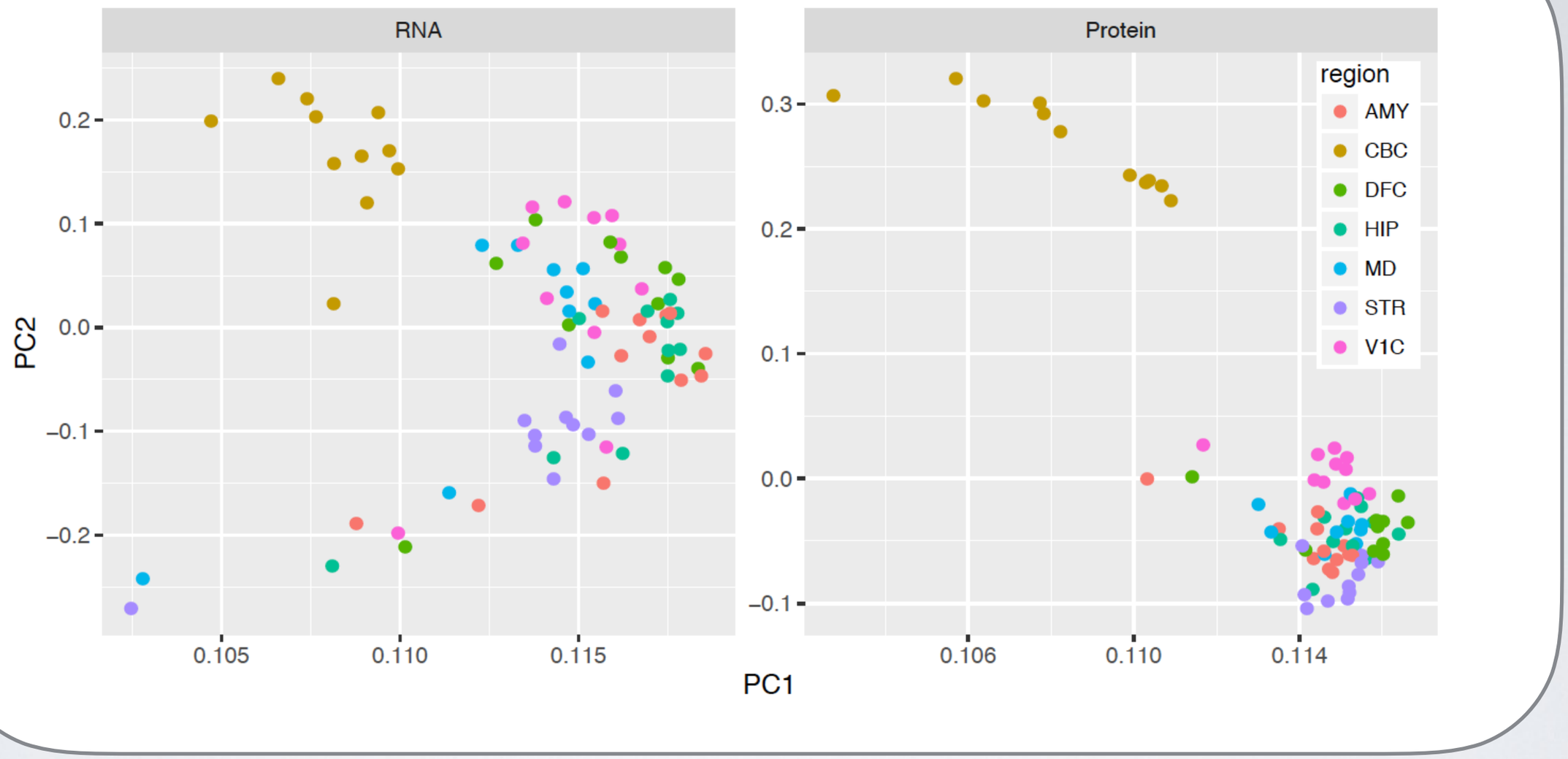




KEGG pathway	pAdj	proteins
Cocaine addiction	2.55E-06	ADCY5, DDC, RGS9, SLC18A2, SLC6A3, TH
Amphetamine addiction	0.0006	ADCY5, DDC, SLC18A2, SLC6A3, TH
Dopaminergic synapse	0.001	ADCY5, DDC, GNAL, SLC18A2, SLC6A3, TH
Parkinson's disease	0.0147	ADCY5, GNAL, SLC18A2, SLC6A3, TH

Proteins from striatal enriched clusters are functionally related, and enriched for appropriate KEGG pathways

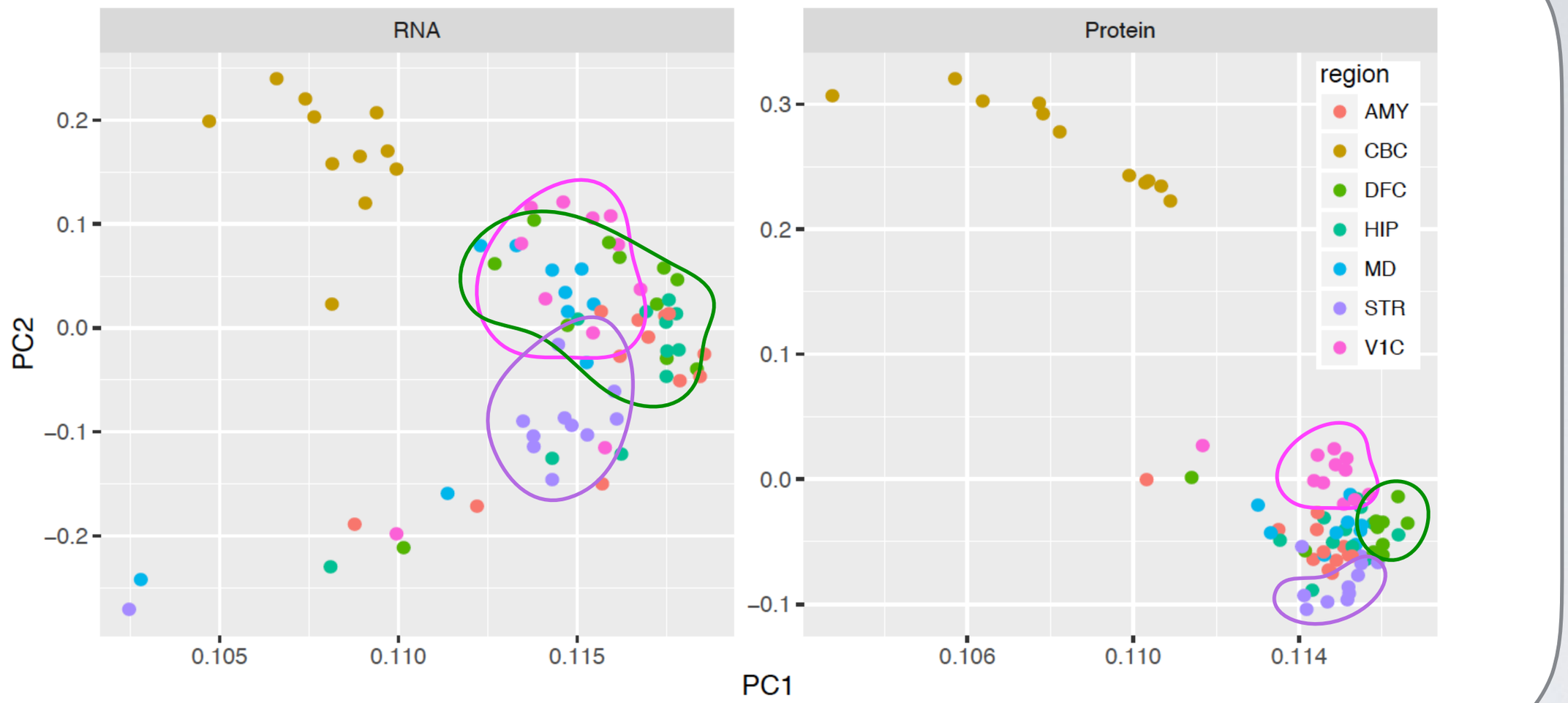
Comparison to RNA-seq



Cerebellum is more clearly separated from the other regions by protein

The other regions are easier to define by protein

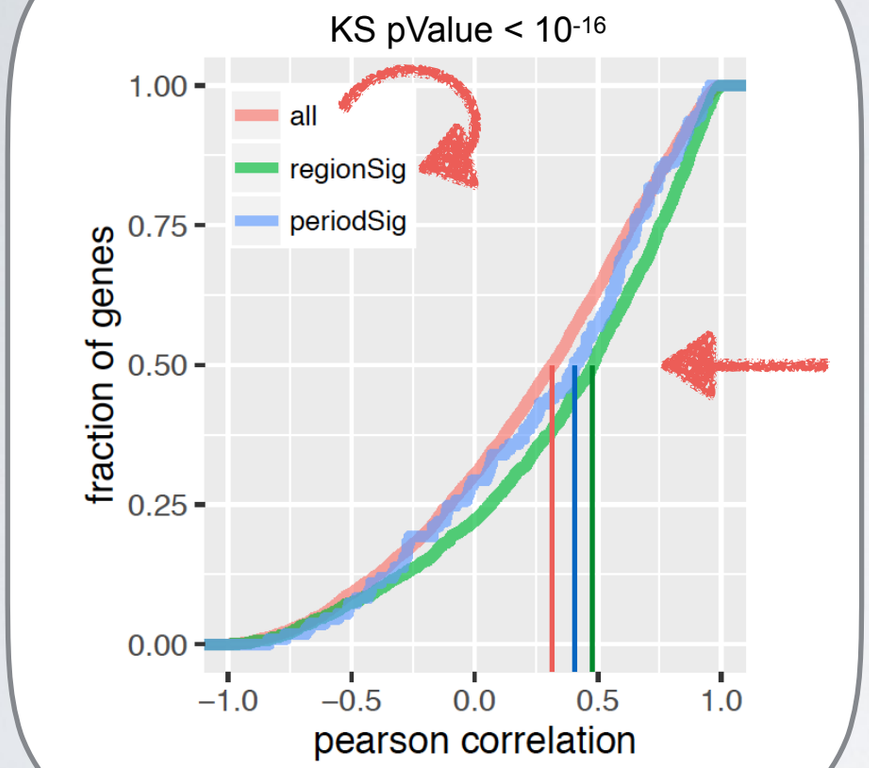
Comparison to RNA-seq



Cerebellum is more clearly separated from the other regions by protein

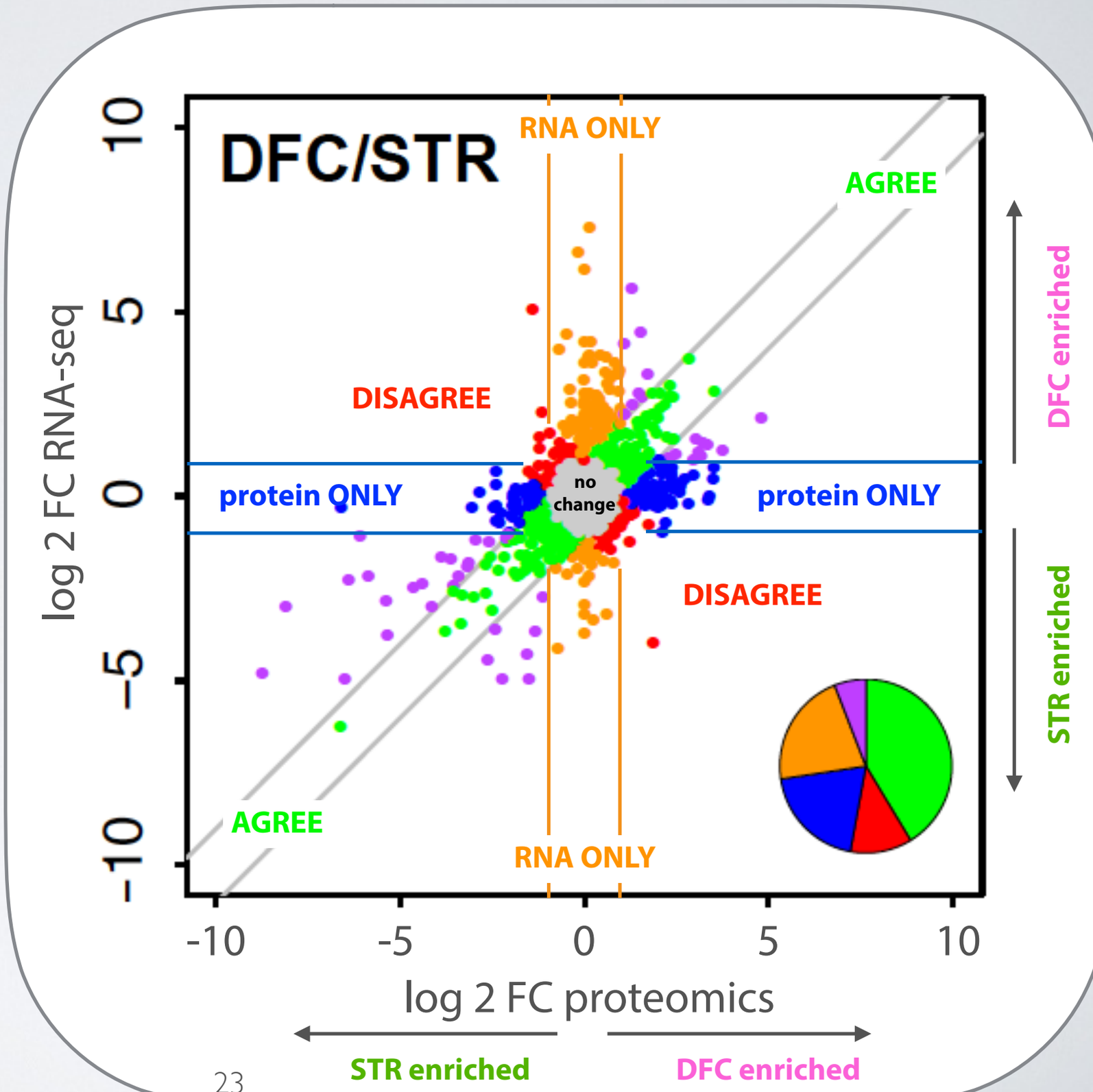
The other regions are easier to define by protein

RNA - protein comparison

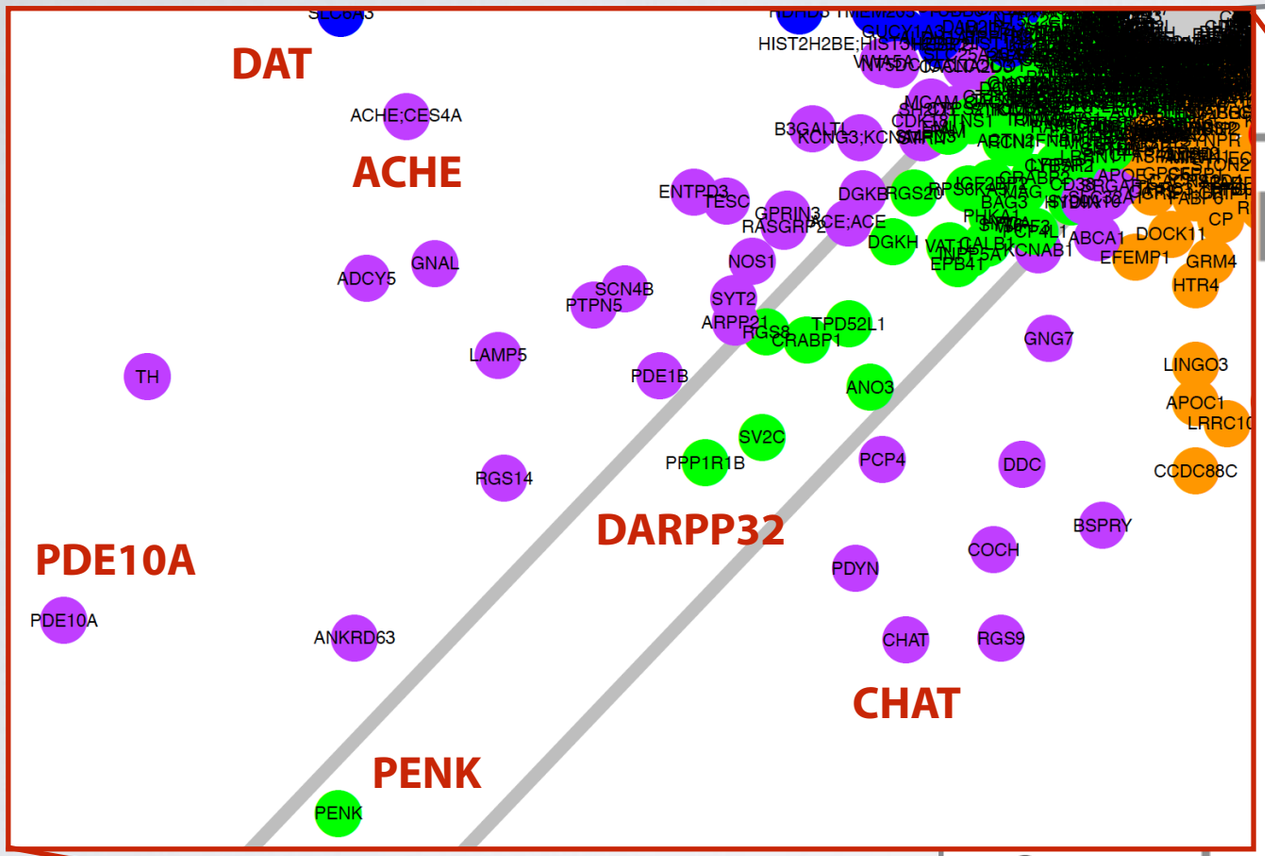


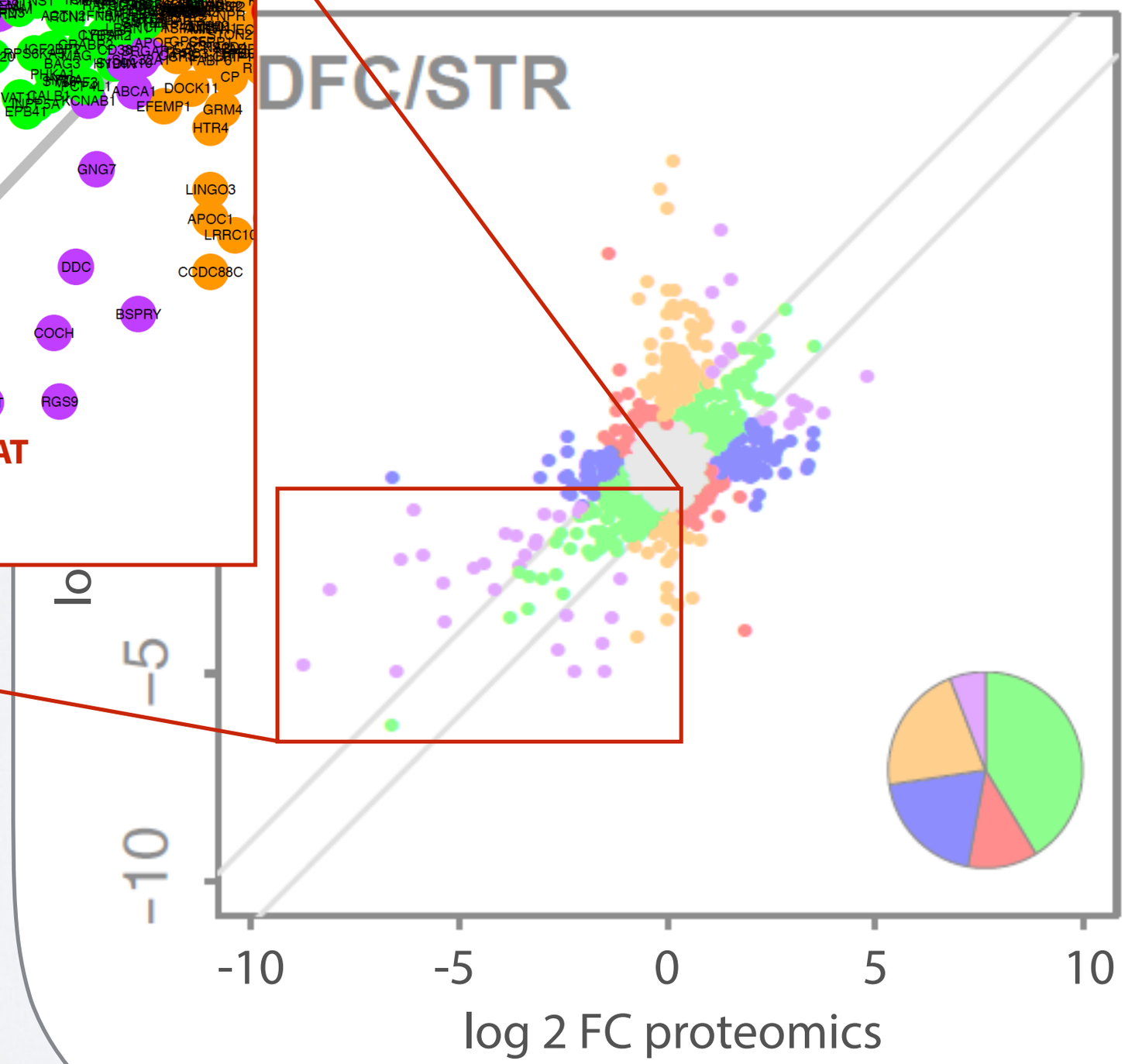
Region DEX genes significantly more correlated between RNA and protein

Trend towards period expression (too few genes)



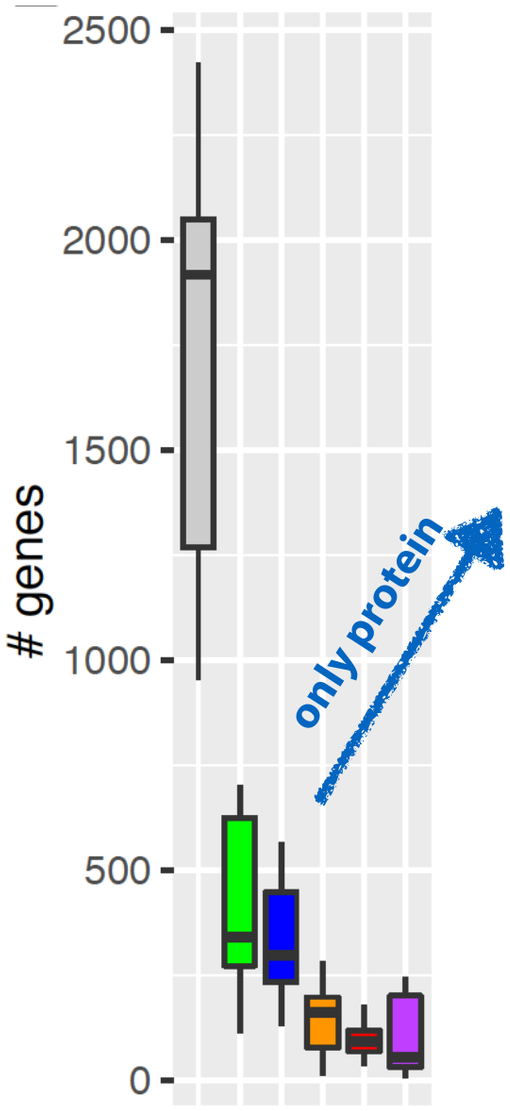
RNA - protein comparison





Ontological analysis

Do these groups of genes relate to ontology in any way?

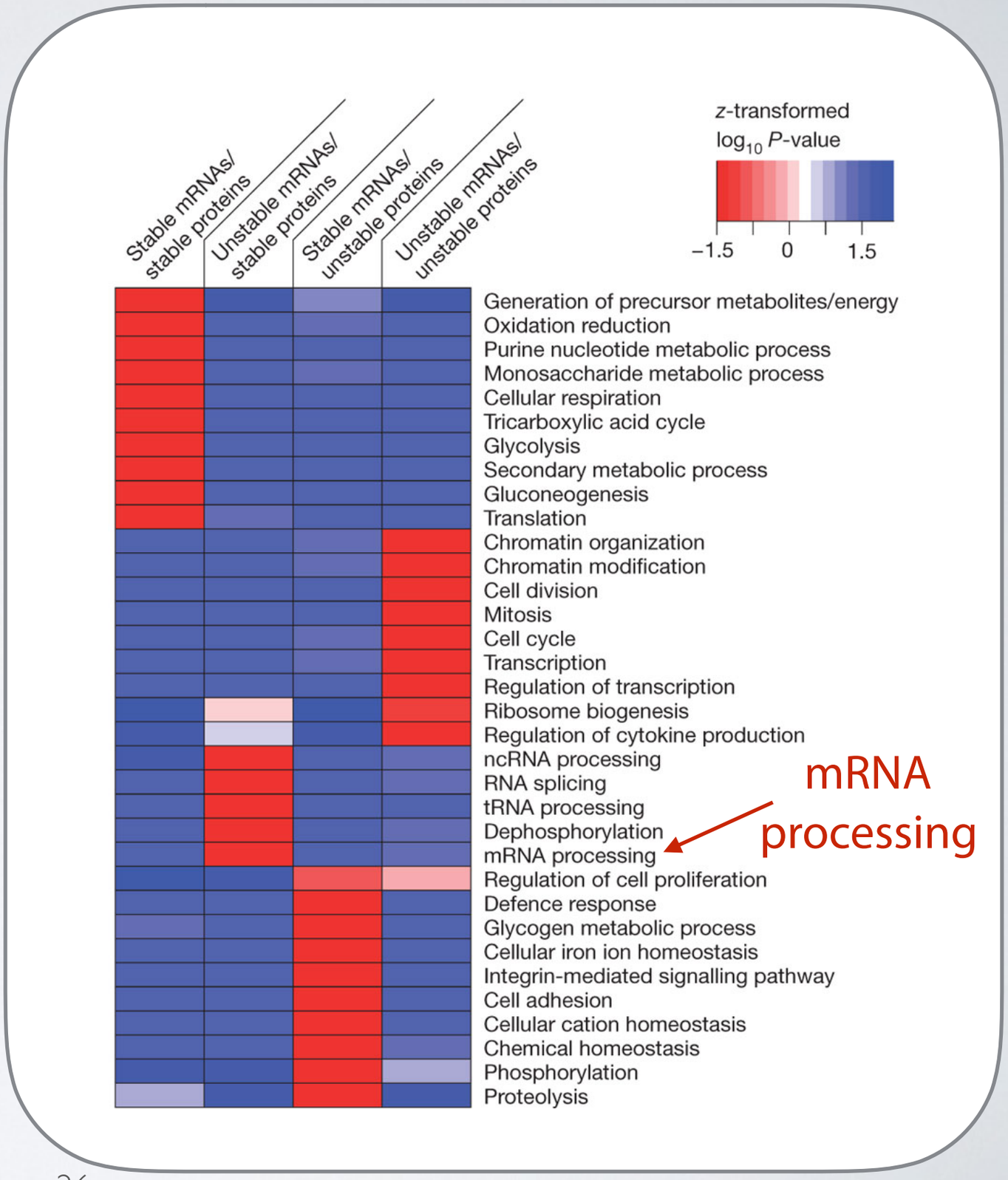


Term	pAdj	Enrichment
RNA processing	3.31E-28	3.23
mRNA processing	3.31E-28	3.58
RNA splicing	3.31E-28	3.57
RNA splicing, via transesterification	3.31E-28	3.7
mRNA splicing, via spliceosome	3.31E-28	3.69
RNA metabolic process	3.31E-28	1.95
nucleic acid metabolic process	3.31E-28	1.88
gene expression	3.31E-28	1.78
nucleobase-containing compound	3.31E-28	1.65

nucleoplasm	8.58E-29	2.19
nuclear part	8.58E-29	1.93
nuclear lumen	8.58E-29	2.01
nucleus	8.58E-29	1.57

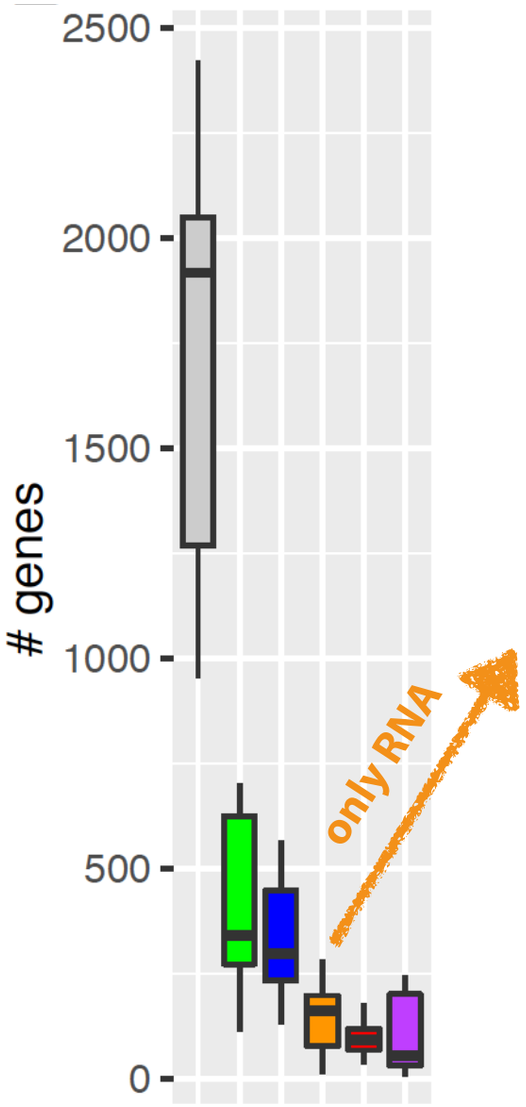
Protein only changes

- Overall enrichment for nuclear proteins driven by cell body density
- RNA processing terms more significant than other nuclear terms
  - may be a reflection of relative stability
- Other interesting proteins appear in region comparisons with more similar cytoarchitecture



Schwanhausser et al., Nature, 2011

RNA only changes

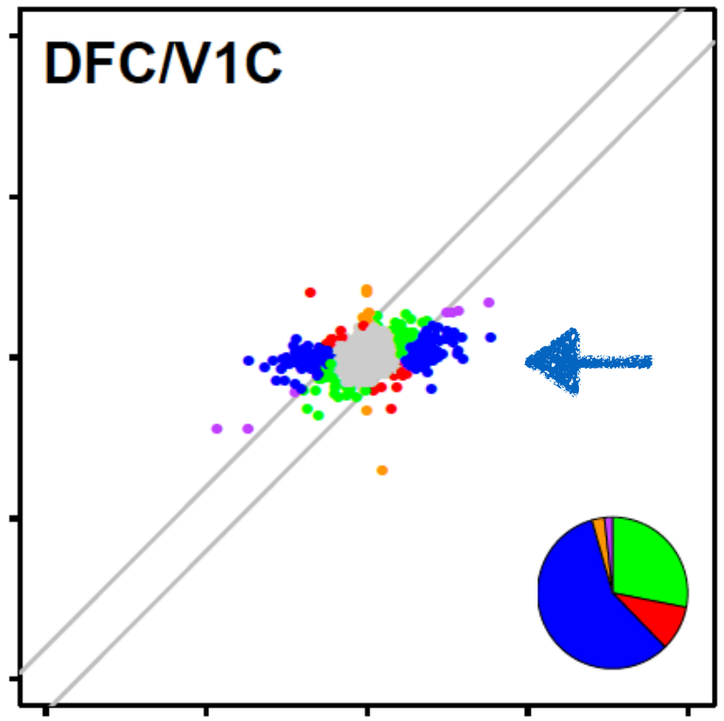


Term	pAdj	Enrichment
signaling	0.00071	1.39
cell communication	0.0008	1.38
metal ion transport	0.0016	2.33
ion transmembrane transport	0.0017	2.09
ion transport	0.0027	1.84
neurological system process	0.0027	2.23
transmembrane transporter complex	0.039	2.29
plasma membrane region	0.04	1.73
transporter complex	0.041	2.24
ion channel complex	0.042	2.36
dense core granule	0.05	6.78

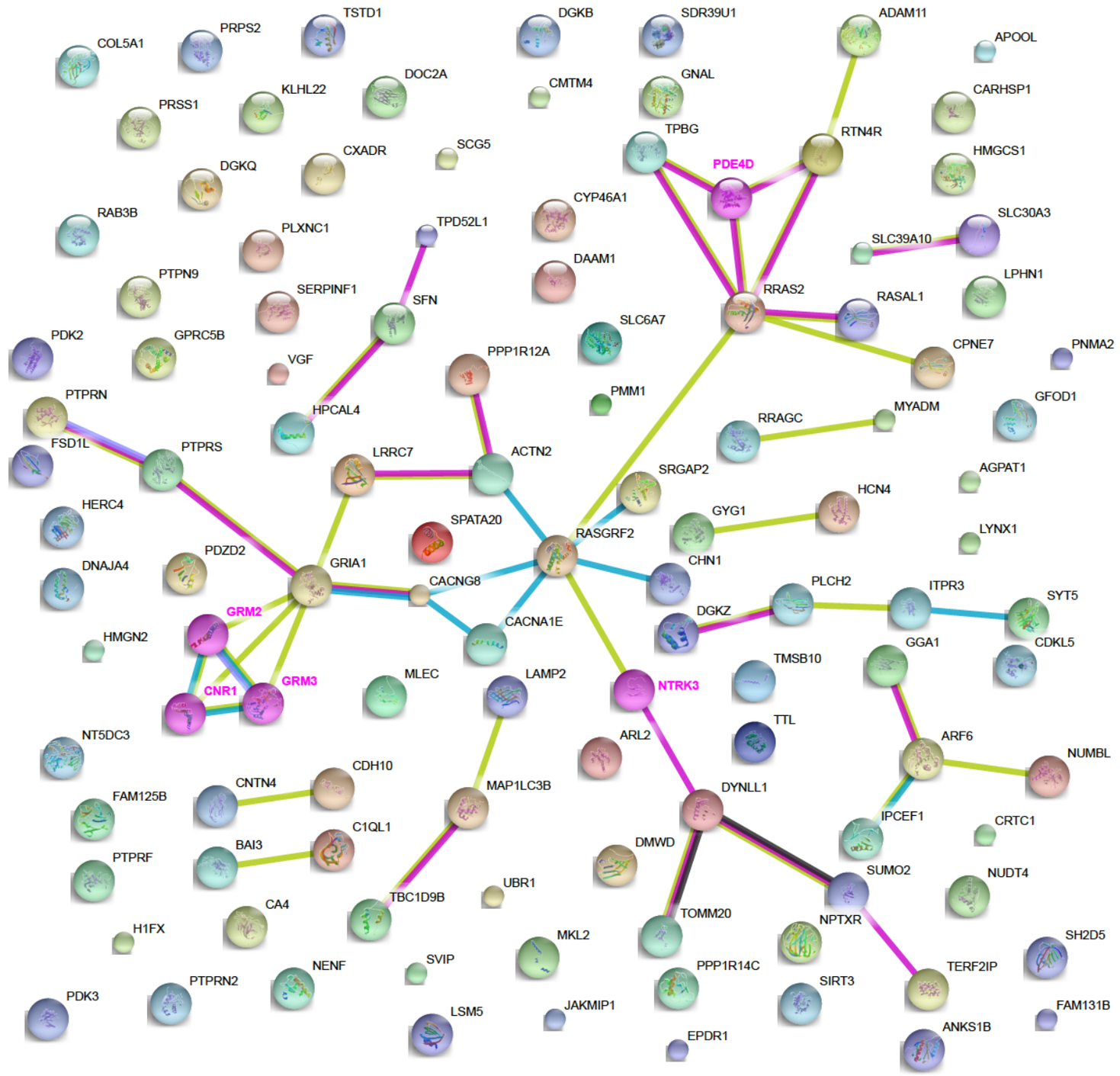
RNA only changes

- Tend to be membrane based, signalling proteins
- These proteins may be synthesized in one region, but transported to another
- Could have important implications for targeting knockdowns, for example

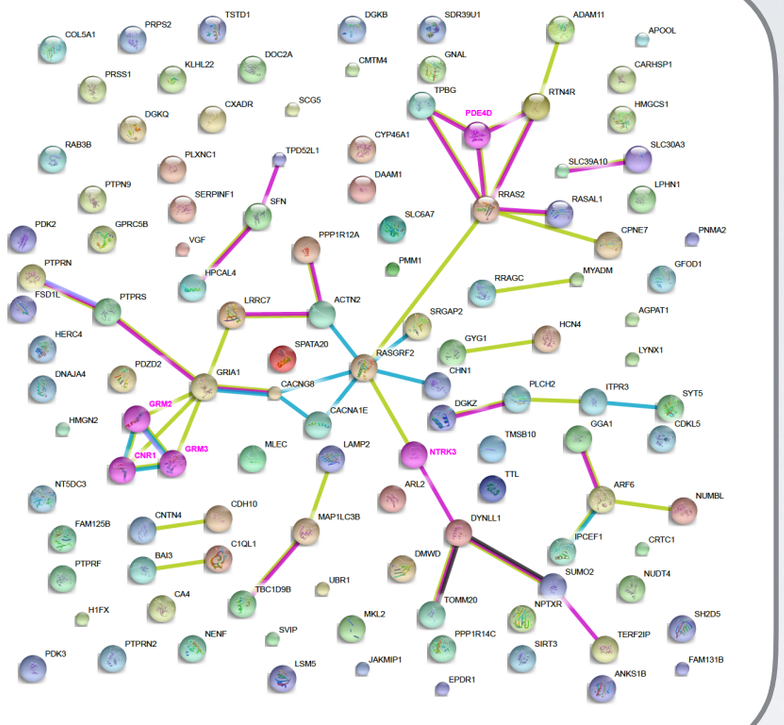
Comparison of areas with similar cytoarchitecture



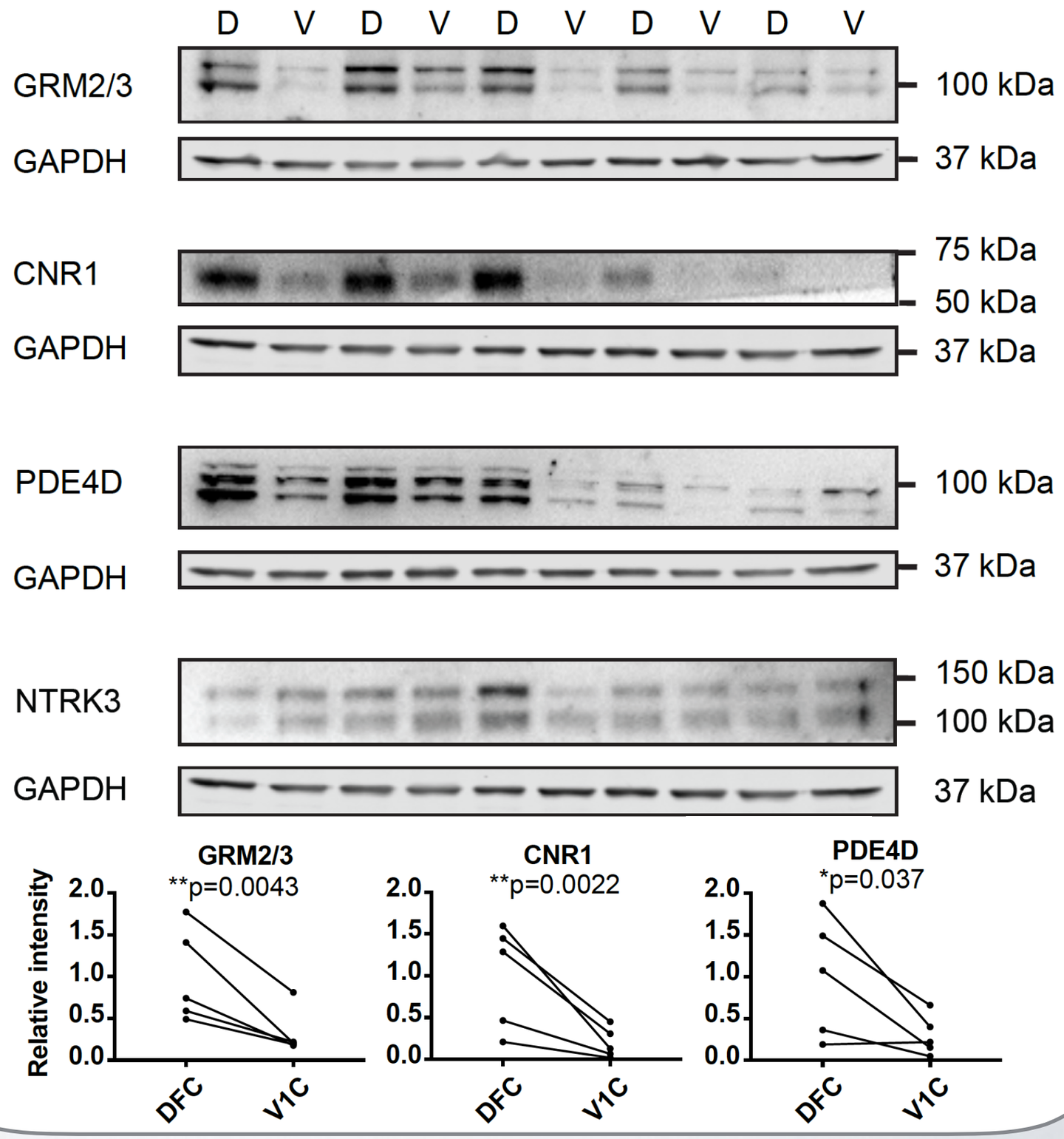
Proteins enriched in DFC vs V1C are functionally related



Comparison of areas with similar cytoarchitecture



Proteins enriched in DFC compared to V1C are functionally related



Talk outline

- Why proteomics?
- BrainSpan/psychENCODE project samples
- Proteomic methods
- Results
- **Conclusions**

Conclusions

- Proteomic level data can reveal differences between regions not seen at the RNA level
- This approach is most powerful in comparing regions with similar cytoarchitecture
- Proteomic data is still biased towards abundant proteins. Future work should consider how best to simplify mixtures to increase the sensitivity of the technique:
  - multiplexing samples for fractionation by liquid chromatography
  - sub cellular fractionation
  - tissue specific references to increase identification rates/identify isoforms
  - cell-type specific proteomics

How can you use this data?

- Increase the number of IDs in your label free experiments through matching between our fractionated libraries
- Use these spectra to select peptides for targeted proteomics
- For assessing protein stoichiometry
- Useful tool for looking at RNA/protein agreement for a new gene of interest
- We also have mouse vs human comparison data - is protein expression similar between humans and your animal model?
- We are hoping to extend this work to isoform level coverage soon

Data can be found in the supplementary tables here:
https://www.nature.com/articles/s41593-017-0011-2

Thanks!

Nairn Lab - Yale
Angus Nairn
Becky Carlyle
Rob Kitchen
Shannon Leslie
Dan Li
Veronica Musante
Xuehong Shang
Rashaun Wilson

Gerstein Lab - Yale
Rob Kitchen
Jing Zhang
Joel Rozowsky
Mark Gerstein

Sestan Lab
Nenad Sestan
Mihovil Pletikos
Andre Sousa

Yale / NIDA Proteomics
TuKiet Lam
Jean Kanyo
Ted Voss
Navin Rauniyar

Funding: psychENCODE consortium, Yale NIDA Neuroproteomics Center, NARSAD Young Investigator