Assaying ncRNAs using high-throughput transposon mutagenesis in *Salmonella*

Lars Barquist

wellcome trust sanger institute

Salmonella Typhi

- human-adapted pathogen
- ~22 million cases per year; 220,000 deaths

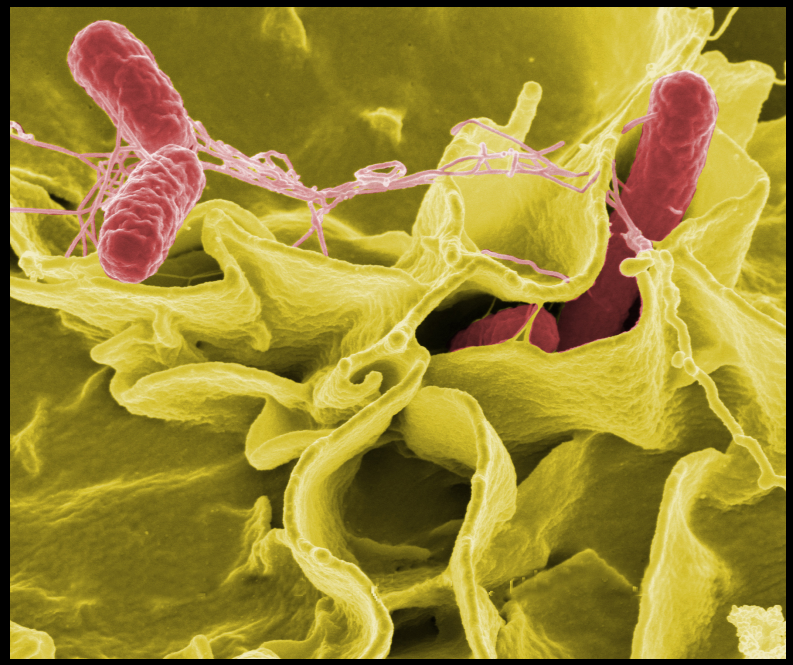
Symptoms: fever, malaise, headche, cough, bloody nose, bradycardia, delirium, diarrhea, constipation, enlargement of the spleen and liver, intestinal hemorrhage, intestinal perforation and septicemia, encephalitis, neuropsychiatric symptoms ("muttering delirium" or "coma vigil"), metastatic abscesses, cholecystitis, endocarditis, osteitis

Also can cause a long term infection of the gall bladder in otherwise healthy individuals

Sources:
Crump *et al.* 2004, *The global burden of typhoid fever*
http://en.wikipedia.org/wiki/Typhoid_fever

Salmonella Typhimurium

- broad host-range
- major cause of gastroenteritis
- model organism for sRNA functional characterization
- causes Typhoid-like disease in mice



Source:NIAID

Problem

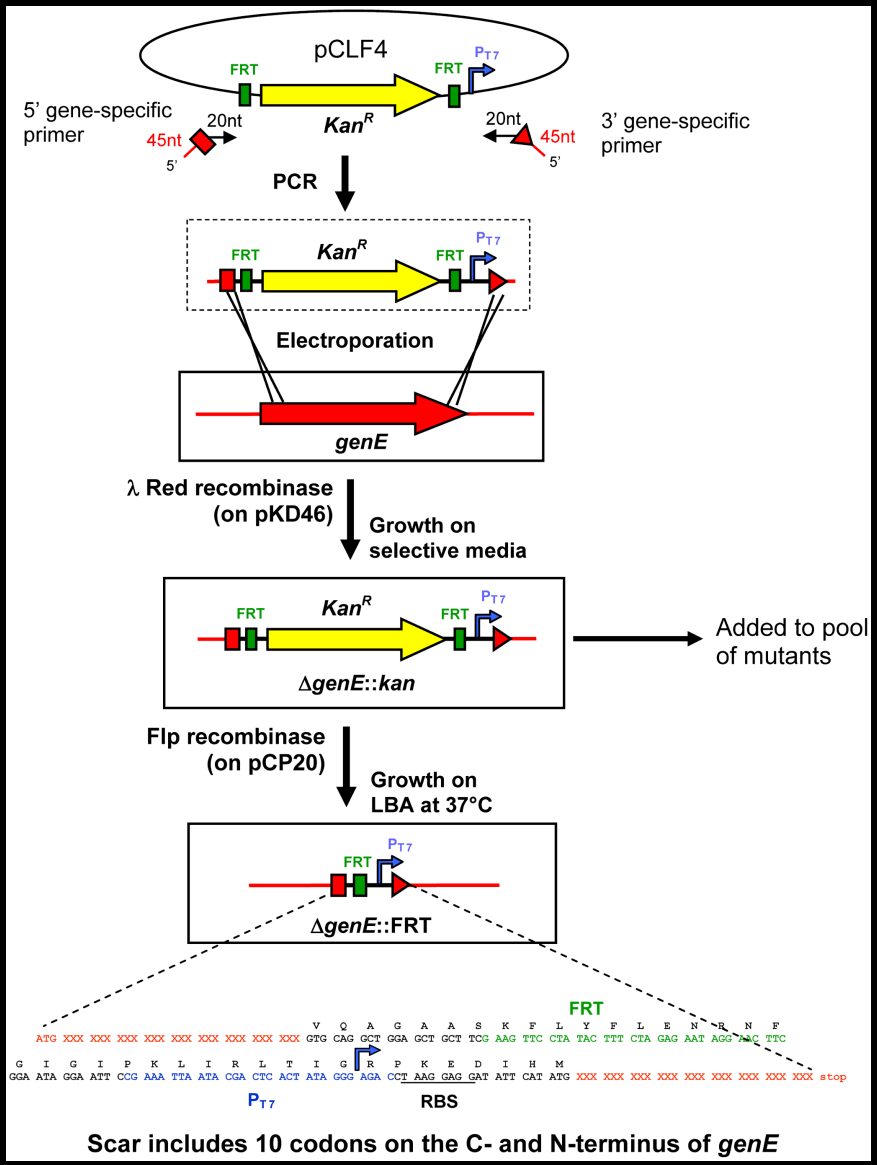
- Lots of genes — only vague ideas of what they do.

Idea

- Create a library of trackable knockouts
- Use this to determine gene ‘essentiality’
- Pass this library through different environments – competition
- Use this information to generate hypotheses regarding function

Targeted knockouts

Source:
Santiviago *et al.* PLoS Pathogens 2009, *Analysis of Pools of Targeted Salmonella Deletion Mutants Identifies Novel Genes Affecting Fitness during Competitive Infection in Mice*



See also:
Hobbs *et al.* Journal of Bacteriology 2010, *Small RNAs and Small Proteins Involved in Resistance to Cell Envelope Stress and Acid Shock in Escherichia coli: Analysis of a Bar-Coded Mutant Collection*

Baba *et al.* Molecular Systems Biology 2006, *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection*

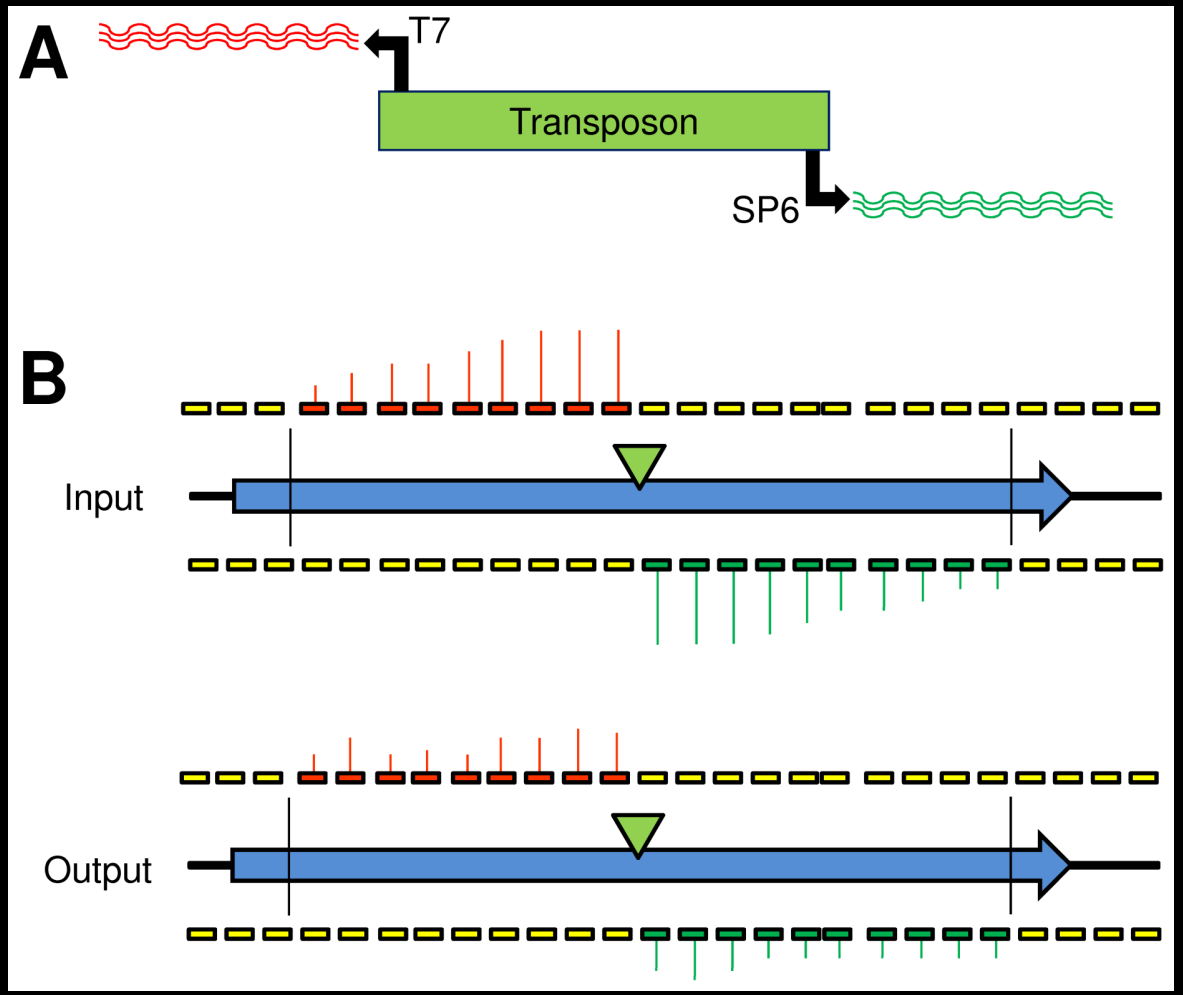
Advantages

- Have the mutant for follow up studies
- Determine gene ‘essentiality’ in a non-competitive environment

Disadvantages

- Labor intensive
- Can only query annotated regions

Transposon mutagenesis



Source:
Chaudhuri *et al.* PLoS Pathogens 2009, *Comprehensive Identification of Salmonella enterica Serovar Typhimurium Genes Required for Infection of BALB/c Mice*

Advantages

- Quick – no gene targeting
- Annotation agnostic

Disadvantages

- Limited density, i.e. non-saturating
- Difficulty locating insertion site
- ‘Essentiality’ determined in a competitive environment

Transposon Directed Insertion-site Sequencing

Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants

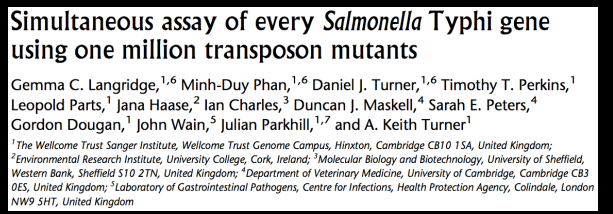
Gemma C. Langridge,[1,6] Minh-Duy Phan,[1,6] Daniel J. Turner,[1,6] Timothy T. Perkins,[1] Leopold Parts,[1] Jana Haase,[2] Ian Charles,[3] Duncan J. Maskell,[4] Sarah E. Peters,[4] Gordon Dougan,[1] John Wain,[5] Julian Parkhill,[1,7] and A. Keith Turner[1]

[1]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; [2]Environmental Research Institute, University College, Cork, Ireland; [3]Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom; [4]Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom; [5]Laboratory of Gastrointestinal Pathogens, Centre for Infections, Health Protection Agency, Colindale, London NW9 5HT, United Kingdom

See also:
van Opijnen *et al.* Nature Methods 2009, *Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms*

TraDIS

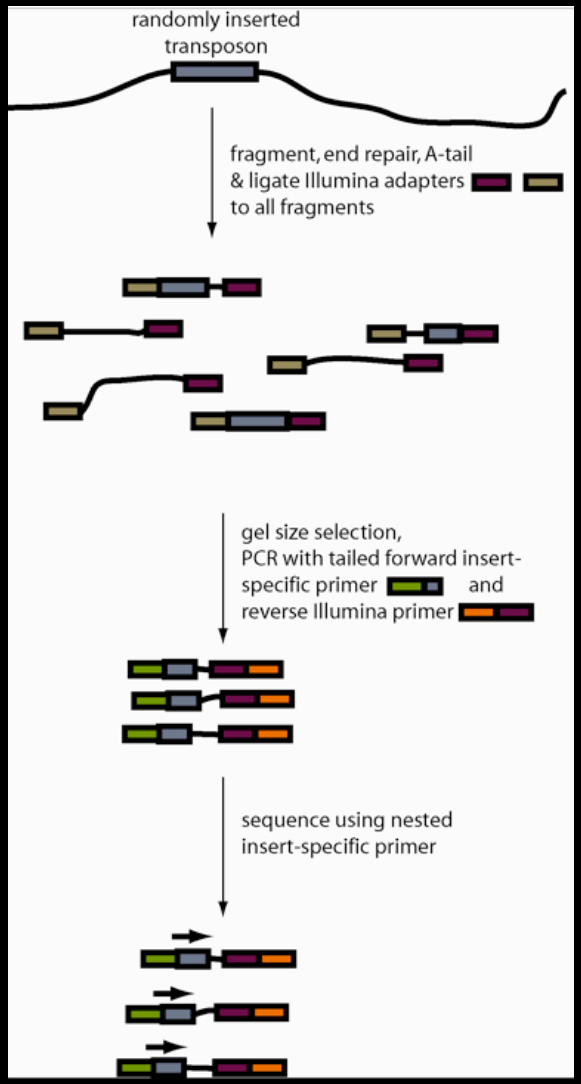
- Electrotransformation with Tn5-derived transposon/transposase complex containing a kanamycin-resistance gene
- 10+ transformations per batch at 42,000 – 146,000 mutants per batch; 13 batches -> 1.1 million mutants.



Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants

Gemma C. Langridge,[1,6] Minh-Duy Phan,[1,6] Daniel J. Turner,[1,6] Timothy T. Perkins,[1] Leopold Parts,[1] Jana Haase,[2] Ian Charles,[3] Duncan J. Maskell,[4] Sarah E. Peters,[4] Gordon Dougan,[1] John Wain,[5] Julian Parkhill,[1,7] and A. Keith Turner[1]

[1]*The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom;* [2]*Environmental Research Institute, University College, Cork, Ireland;* [3]*Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom;* [4]*Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom;* [5]*Laboratory of Gastrointestinal Pathogens, Centre for Infections, Health Protection Agency, Colindale, London NW9 5HT, United Kingdom*

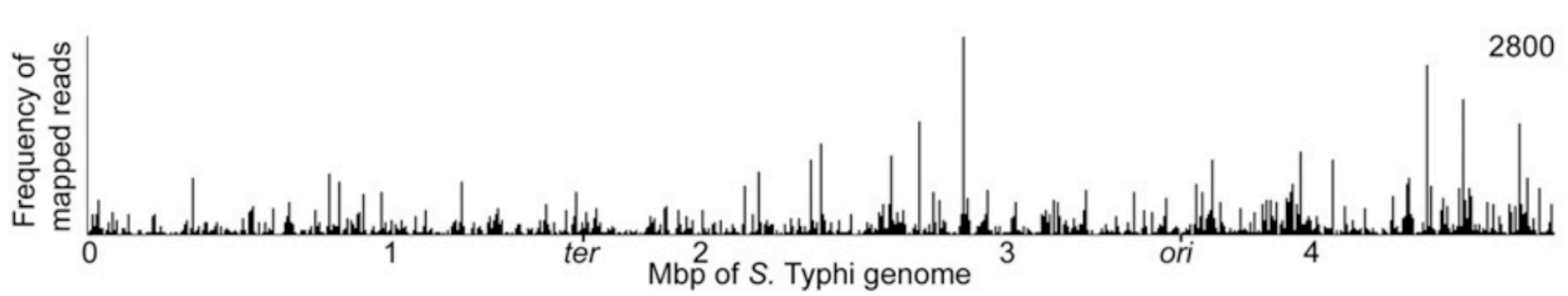


Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants

Gemma C. Langridge,[1,6] Minh-Duy Phan,[1,6] Daniel J. Turner,[1,6] Timothy T. Perkins,[1] Leopold Parts,[1] Jana Haase,[2] Ian Charles,[3] Duncan J. Maskell,[4] Sarah E. Peters,[4] Gordon Dougan,[1] John Wain,[5] Julian Parkhill,[1,7] and A. Keith Turner[1]

[1]*The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom;* [2]*Environmental Research Institute, University College, Cork, Ireland;* [3]*Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom;* [4]*Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom;* [5]*Laboratory of Gastrointestinal Pathogens, Centre for Infections, Health Protection Agency, Colindale, London NW9 5HT, United Kingdom*

TraDIS



370,000 unique transposon insertion sites (higher in newer libraries)

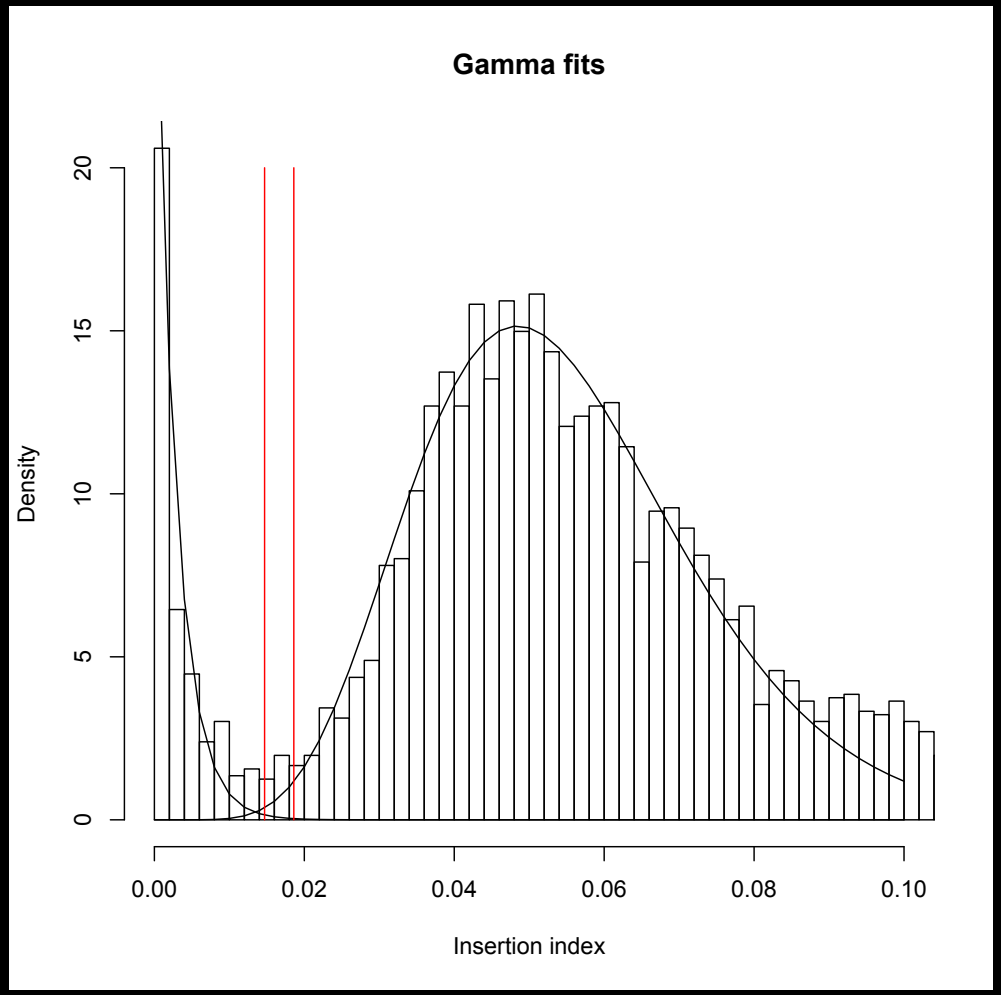
1% chance of a 60 base region not having an insert (assuming Poisson-distributed)

Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants

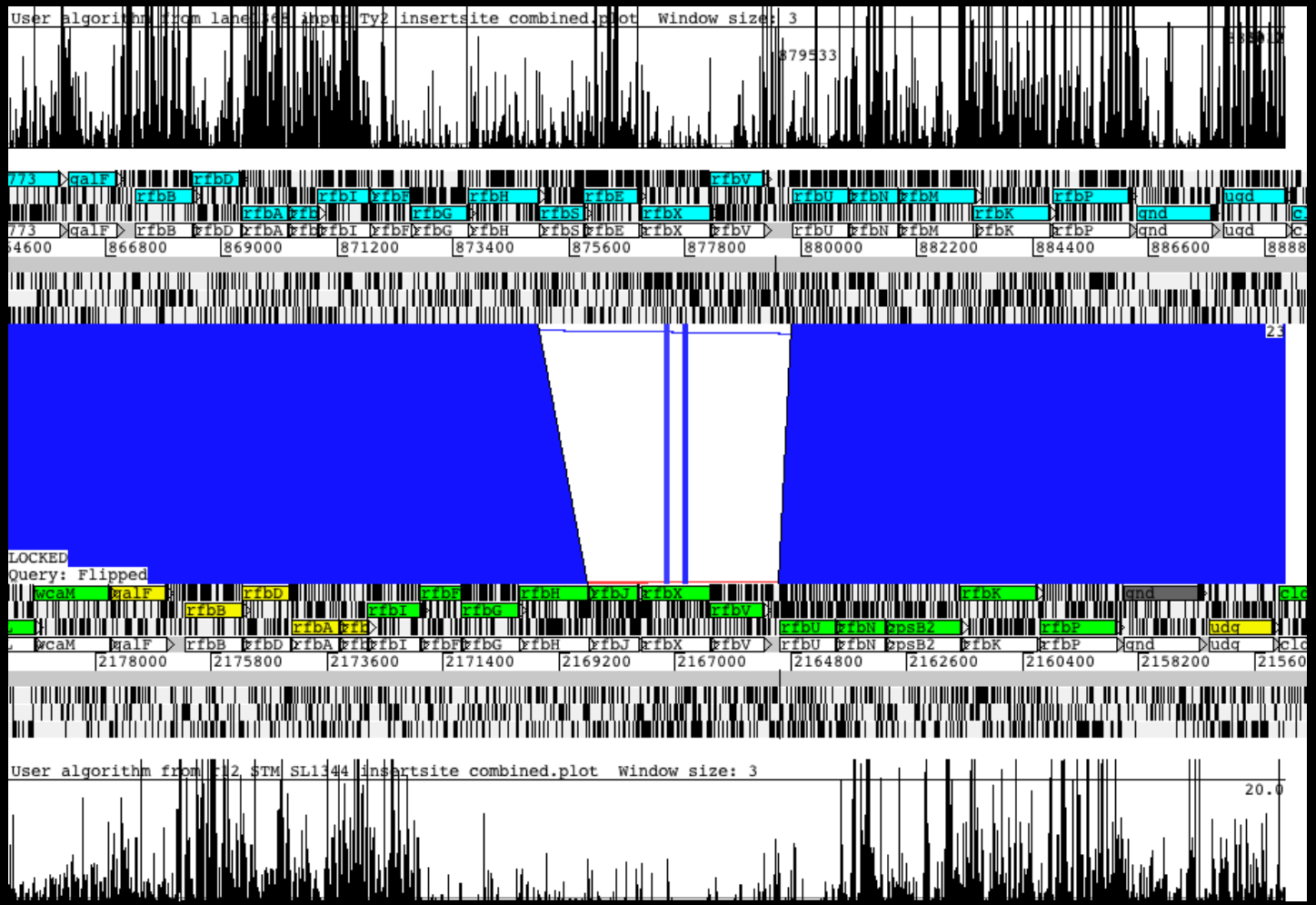
Gemma C. Langridge,[1,6] Minh-Duy Phan,[1,6] Daniel J. Turner,[1,6] Timothy T. Perkins,[1] Leopold Parts,[1] Jana Haase,[2] Ian Charles,[3] Duncan J. Maskell,[4] Sarah E. Peters,[4] Gordon Dougan,[1] John Wain,[5] Julian Parkhill,[1,7] and A. Keith Turner[1]

[1]*The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom;* [2]*Environmental Research Institute, University College, Cork, Ireland;* [3]*Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom;* [4]*Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom;* [5]*Laboratory of Gastrointestinal Pathogens, Centre for Infections, Health Protection Agency, Colindale, London NW9 5HT, United Kingdom*

Calling gene ‘essentiality’



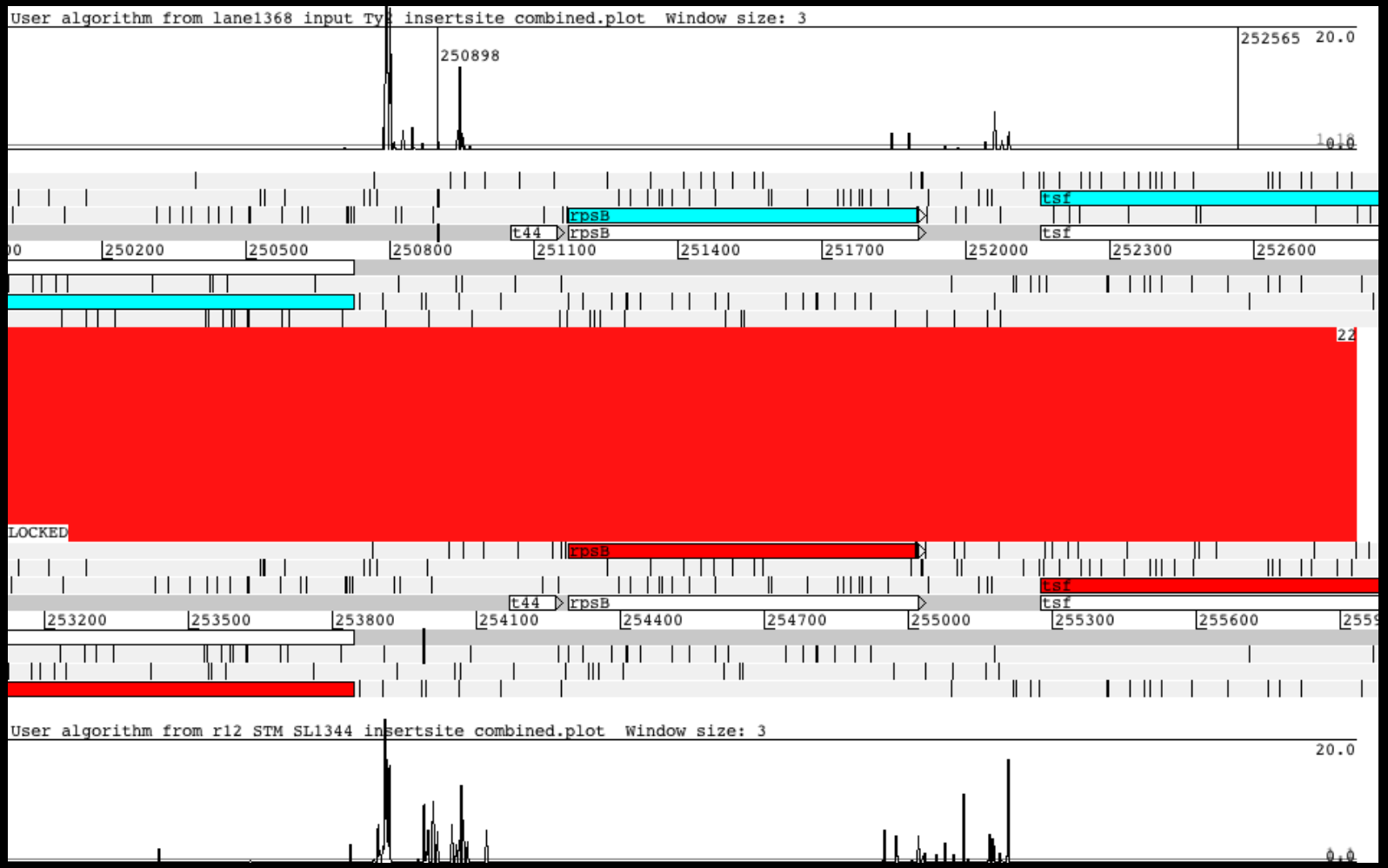
Comparing Genomes



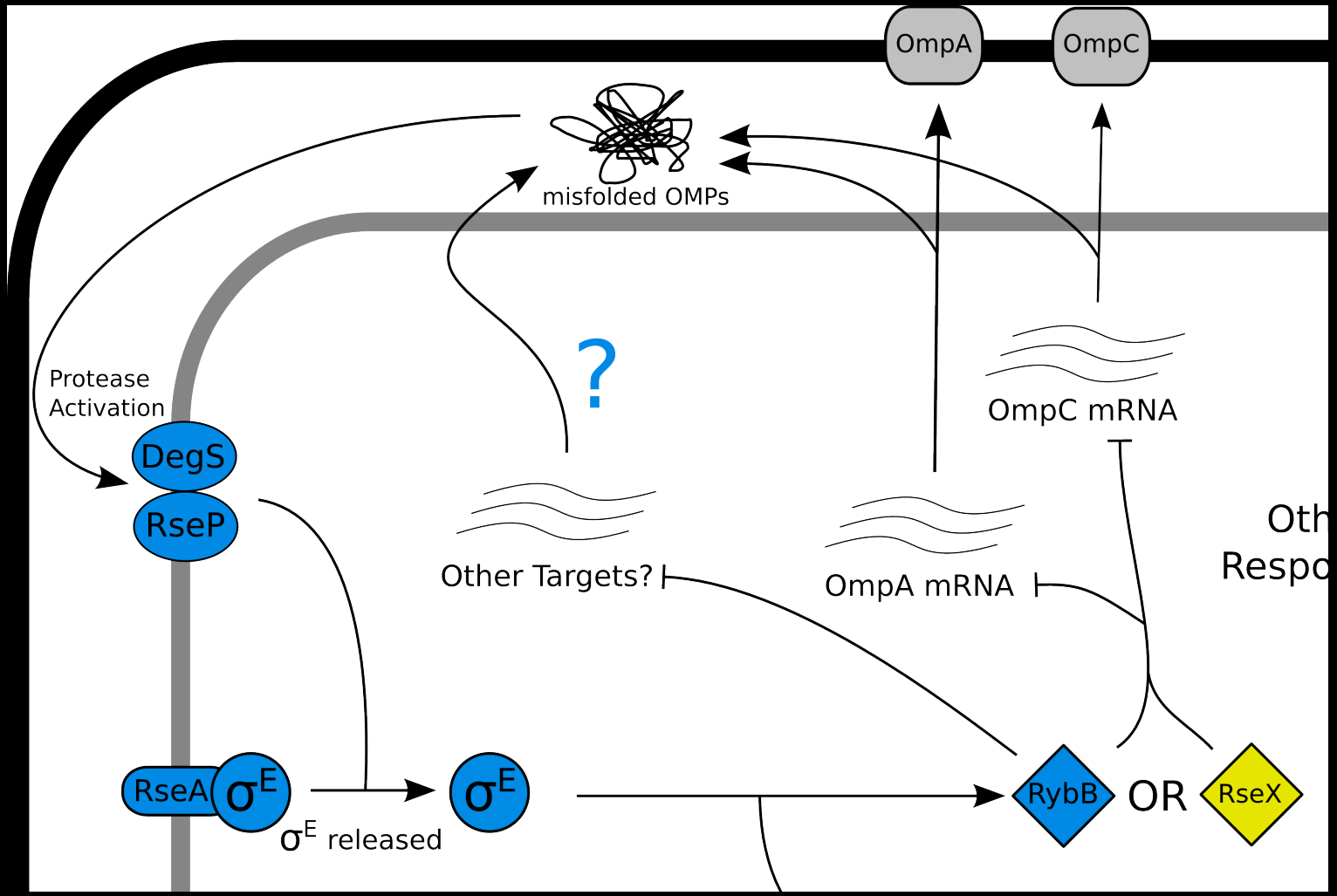
Applicable to ncRNAs?

- Ready-made sanity check: tRNAs
- 40 anti-codons, 80 – 90 tRNAs
- Assume tRNAs uniquely covering a codon should be required for growth
- PPV of ~80%, FPR of <4%, worst case across two independent libraries

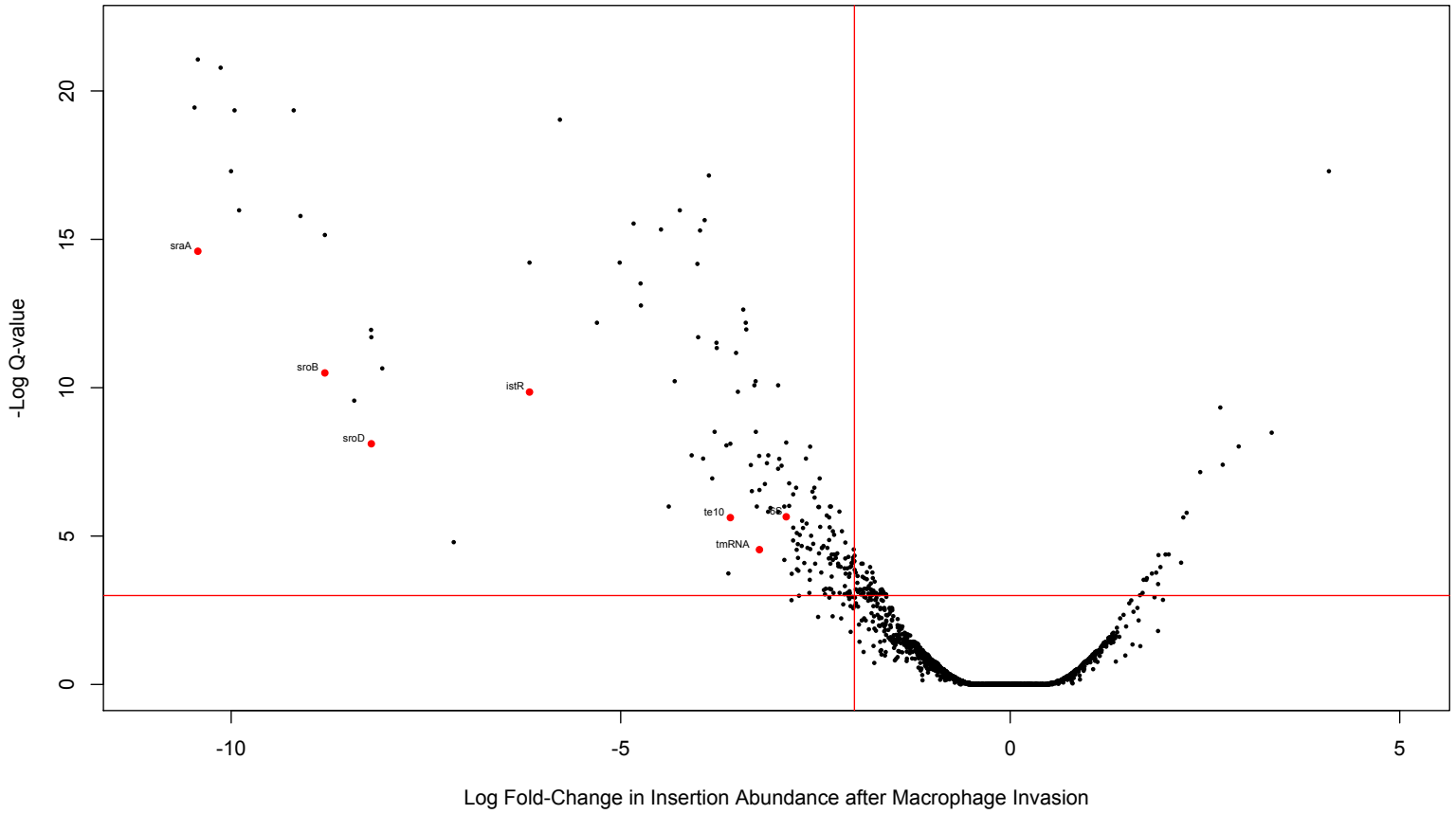
Ribosomal Leaders



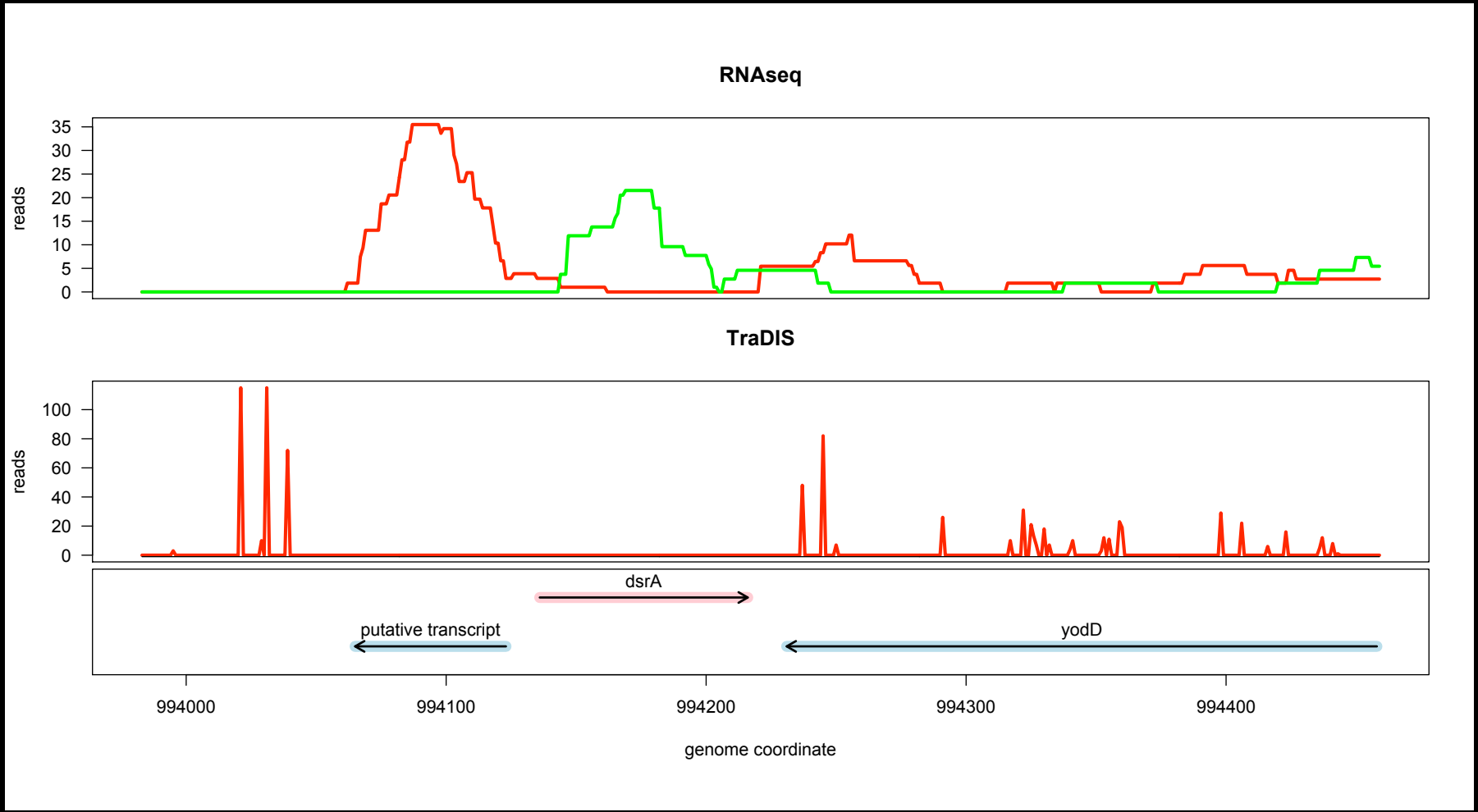
Differences in sRNA networks?



Comparative Conditional TraDIS



Combining HTS techniques



Summary

- TraDIS provides a method to rapidly generate hypotheses for gene function (not without caveats)
- Annotation agnostic
- More data in the near future – more organisms, more conditions, matched RNA-seq

Acknowledgements

Gemma Langridge
Leopold Parts
Amy Cain
Christine Boinett
Keith Turner
Nicholas Thomson
Julian Parkhill

Alex Bateman
Xfam, WTSI
United Kingdom

John Wain
University of East Anglia
United Kingdom

Paul Gardner
University of Canterbury
New Zealand

Pathogen Genomics, WTSI
United Kingdom

wellcome trust
sanger
institute