Identifying genomic signatures of niche specialisation in the rumen microbiome

Prof. Chris Creevey

Chair of Computational Biology,
Institute for Global Food Security,
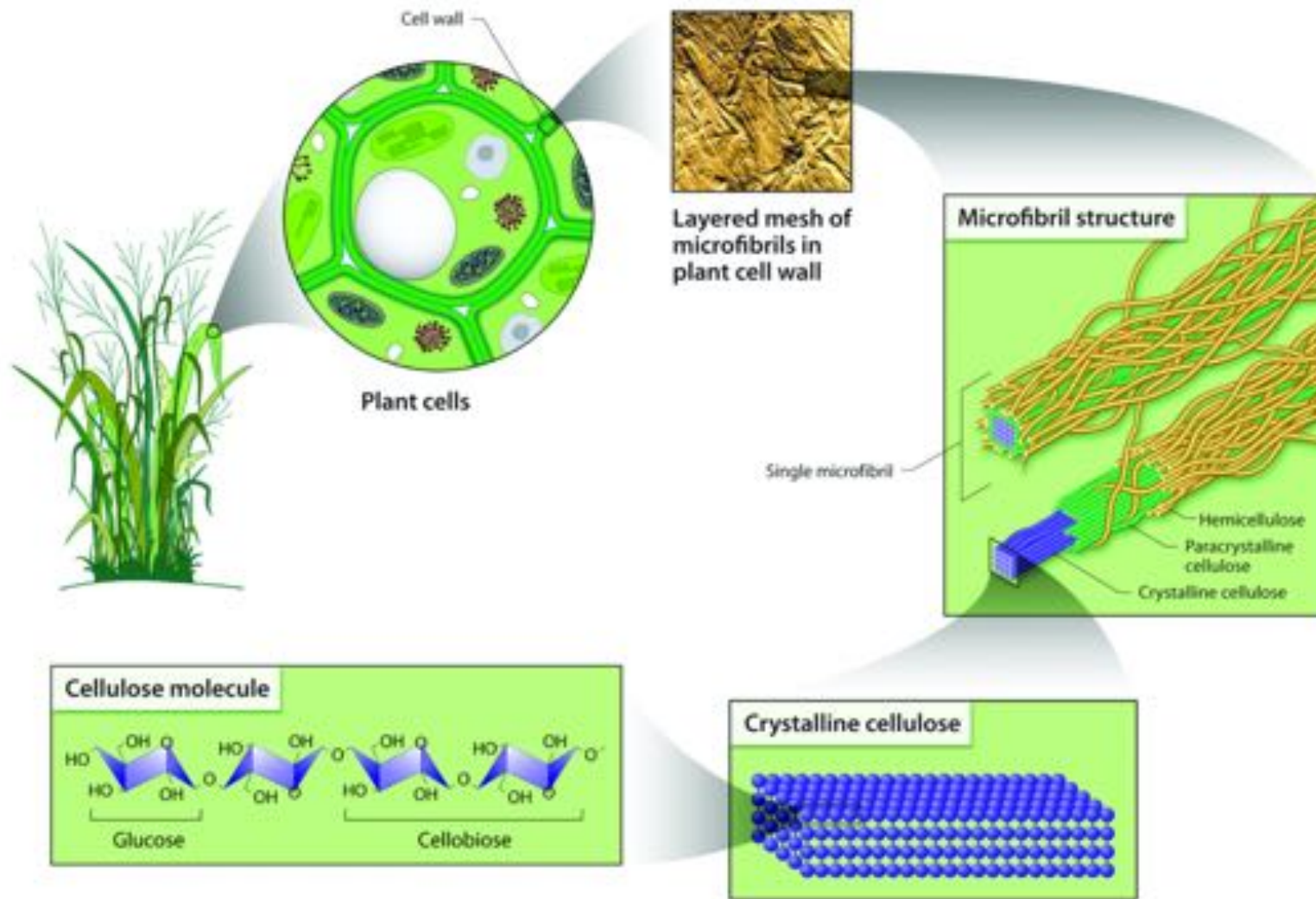School of Biological Sciences,
Queens University Belfast.

Email: c.creevey@qub.ac.uk
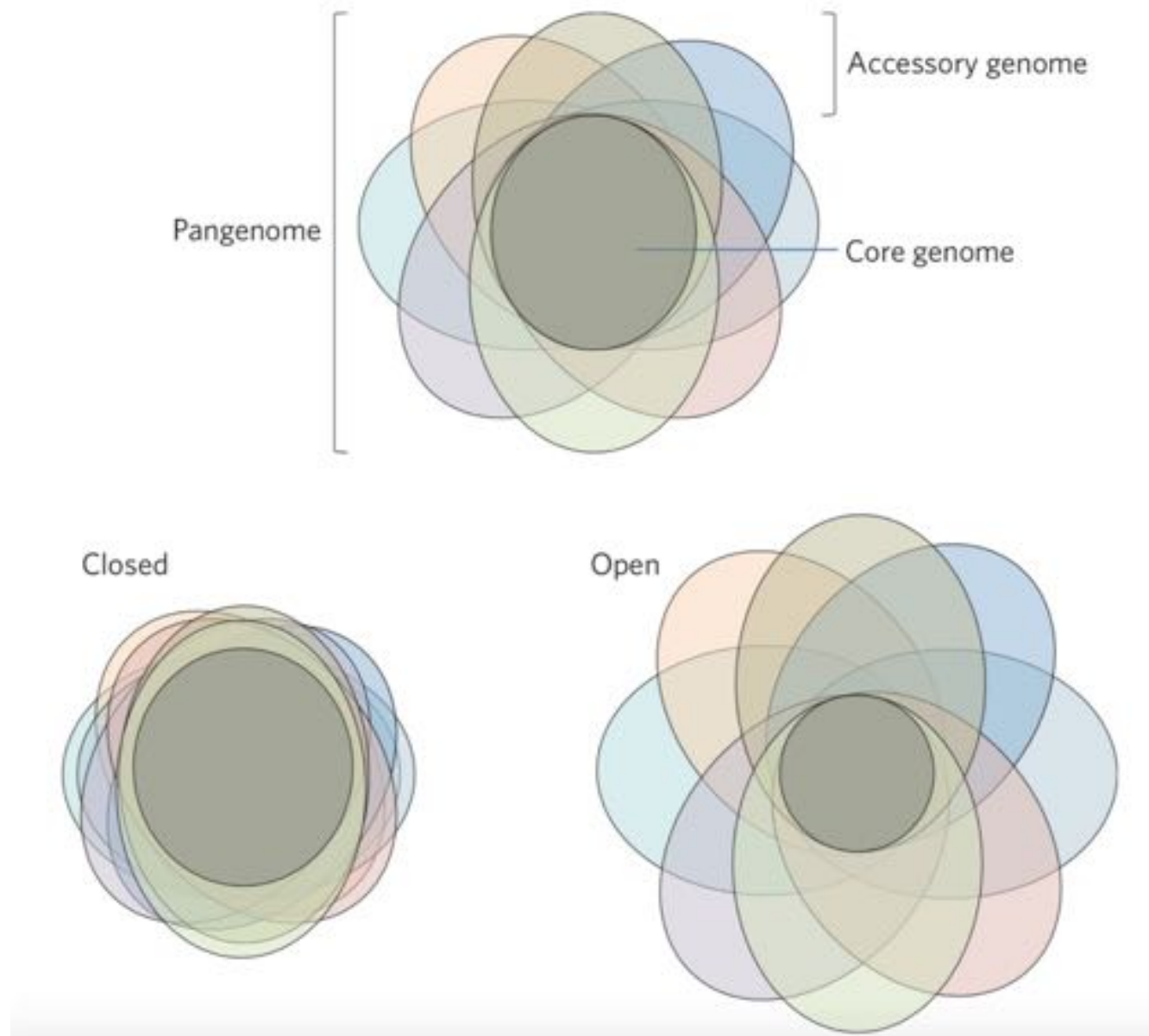www.creeveylab.org
Twitter: @hairy_llama   @creeveylab

QUEEN'S
UNIVERSITY
BELFAST

The Institute For Global Food Security

Does niche specialisation for plant cell wall degradation exist?

Cell wall

Layered mesh of microfibrils in plant cell wall

Plant cells

Microfibril structure

Single microfibril

Hemicellulose
Paracrystalline cellulose
Crystalline cellulose

Cellulose molecule

Glucose

Cellobiose

Crystalline cellulose

What drives Niche specialisation in the Rumen?

McInerney, J.O., McNally, A. and O'Connell, M.J., 2017. Why prokaryotes have pangenomes. *Nature microbiology*, *2*(4), p.17040.

The Conundrum

- Somehow niche specialisation of specific groups of organisms is maintained even under the continued influence of horizontal gene transfer.

- One explanation may be that the horizontal acquisition of a single isoform of a novel gene is not enough to maintain a competitive advantage in fluctuating environmental conditions where a range of isoforms may be required to, for instance, maintain enzymatic activity.

- This could be described as <u>negative frequency dependent selection</u>.

Frequency-dependent selection.

- In **positive** frequency-dependent selection, the fitness of a phenotype increases as it becomes more common.

Ranitomeya imitator

Ranitomeya variabilis

- In **negative** frequency-dependent selection, the fitness of a phenotype decreases as it becomes more common.

Blue tit searches for insect prey using a search image, leaving scarcer types of prey untouched.

Drivers of Niche Specialisation

Dr. Francesco Rubino

- **Can we use genetic information from the microbiome to identify the drivers of niche specialization?**

- We should be able to use adaptive evolution and population genetics approaches developed in other systems to address these questions.

- This will allow us to test whether negative frequency dependent selection drives niche specialization.

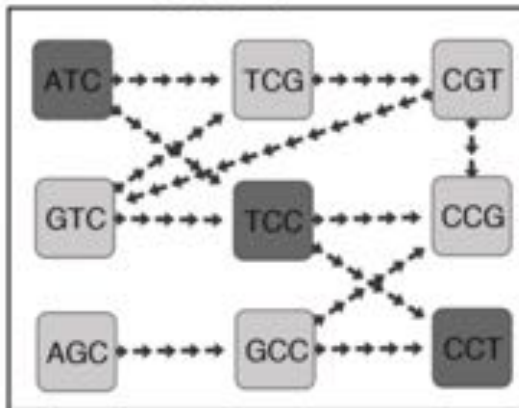# Metagenomic data can result in "average" assemblies

a) Haplotypes and resulting trimers

	Haplotype 1	Haplotype2	Haplotype3	Haplotype4
Sequence:	ATCGT	ATCCG	GTCCT	AGCCT
Trimers:	ATC	ATC	GTC	AGC
	TCG	TCC	TCC	GCC
	CGT	CCG	CCT	CCT

b) Trimer summary

Trimers	Abundance
ATG	2
GTC	1
AGC	1
TCG	1
TCC	2
GCC	1
CGT	1
CCG	1
CCT	2

c) de Bruijn graph of trimers

d) Resulting assembly

Trimers:	ATC
	TCC
	CCT
Sequence:	ATCCT

14 X Rumen samples

Identify SNPs

T	G	G	C	G	C	G	G
T	G	G	C	G	T	G	G
T	G	G	C	G	C	G	G
T	G	G	C	G	T	G	G
T	G	G	C	G	T	G	G
T	G	G	C	G	C	G	G

| T | G | G | C | G | T | G | G |
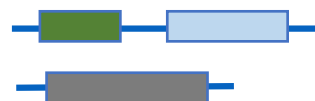
Calculate sequence Diversity/Evolutionary rate
    i.e. pN/pS
Schloissnig *et al* 2012

Prepare DNA, sequence randomly

Align reads to assembly

Map to Metabolic Processes

Assemble Contigs

Predict Genes

Rubino & Creevey Mgkit: Metagenomic Framework For The Study Of Microbial Communities
https://bitbucket.org/setsuna80/mgkit
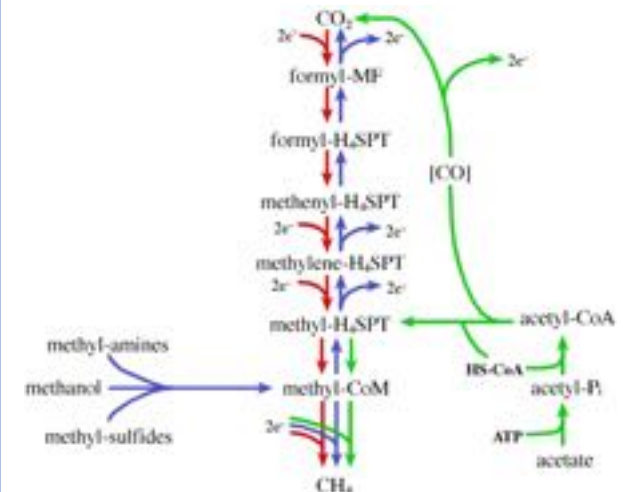
Estimating adaptive diversity estimates from metagenomic data using SNPs

Dr. Francesco Rubino

$$\frac{pN}{pS} = \frac{oN/eN}{oS/eS}$$

oN = number of observed non-synonymous SNPs
eN = number of "expected" non-synonymouse SNPs

oS = number of observed synonymous SNPs
eS = number of "expected" synonymous SNPs

Schloissnig *et al.* (2013) *Genomic variation landscape of the human gut microbiome* Nature **493:**45-50.

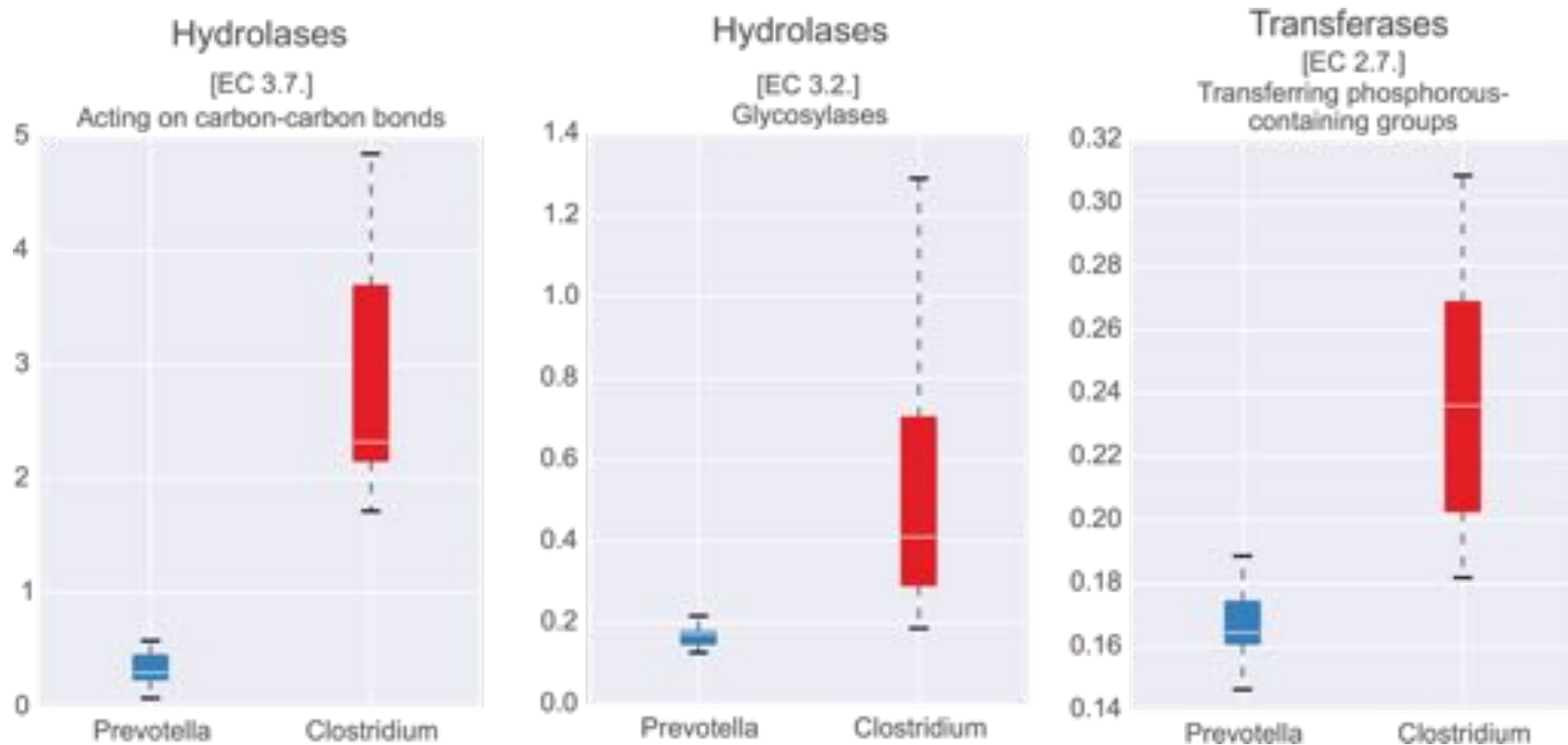# Enzymes categories with a higher adaptive diversity in Prevotella

Dr. Francesco Rubino
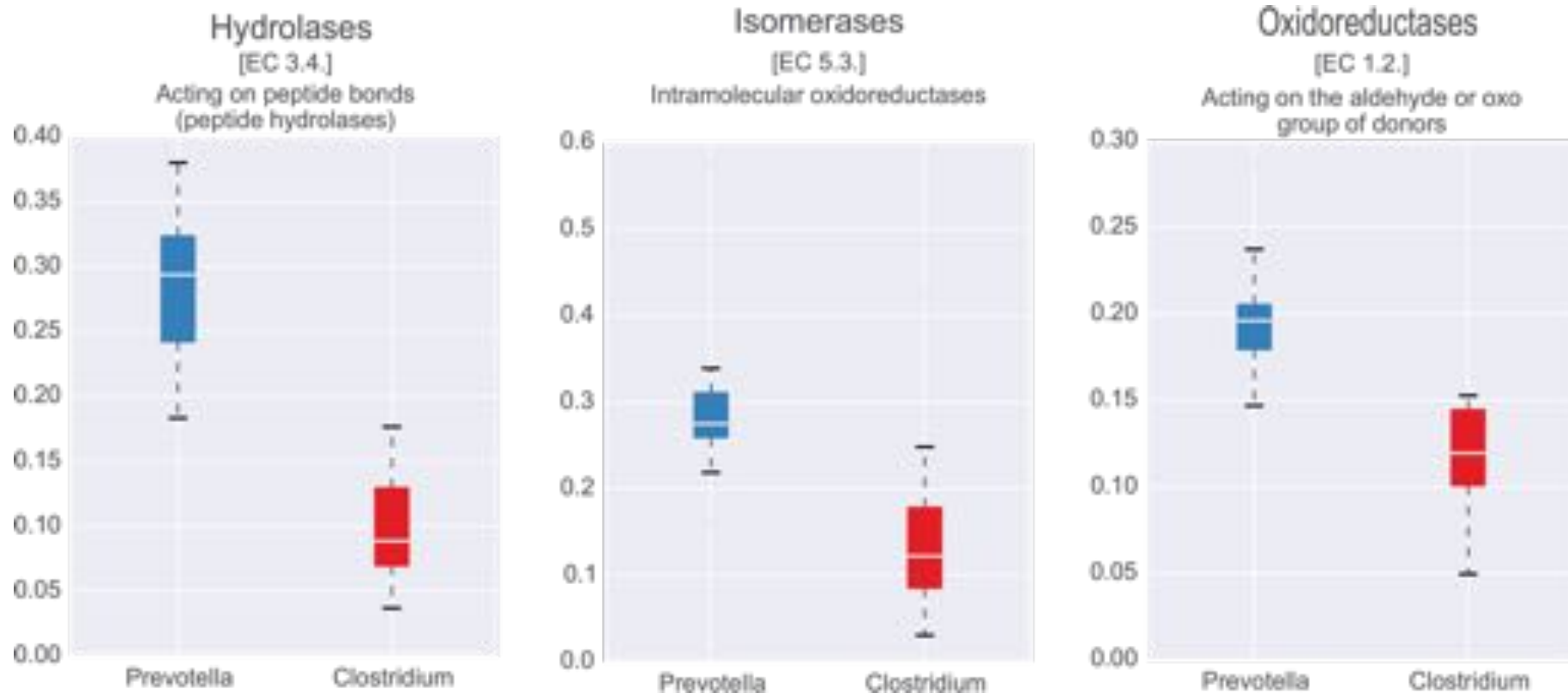
Isomerases in Prevotella may be specialised for the action on xylose (a major component of hemicellulose) to release sugars.

Rubino, Carberry, Waters, Kenny, McCabe & **Creevey,** 2017. *Divergent functional isoforms drive niche specialisation for nutrient acquisition and use in rumen microbiome.* The ISME Journal, 11(4), pp.932-944.

Enzymes categories with a higher adaptive diversity in Clostridium

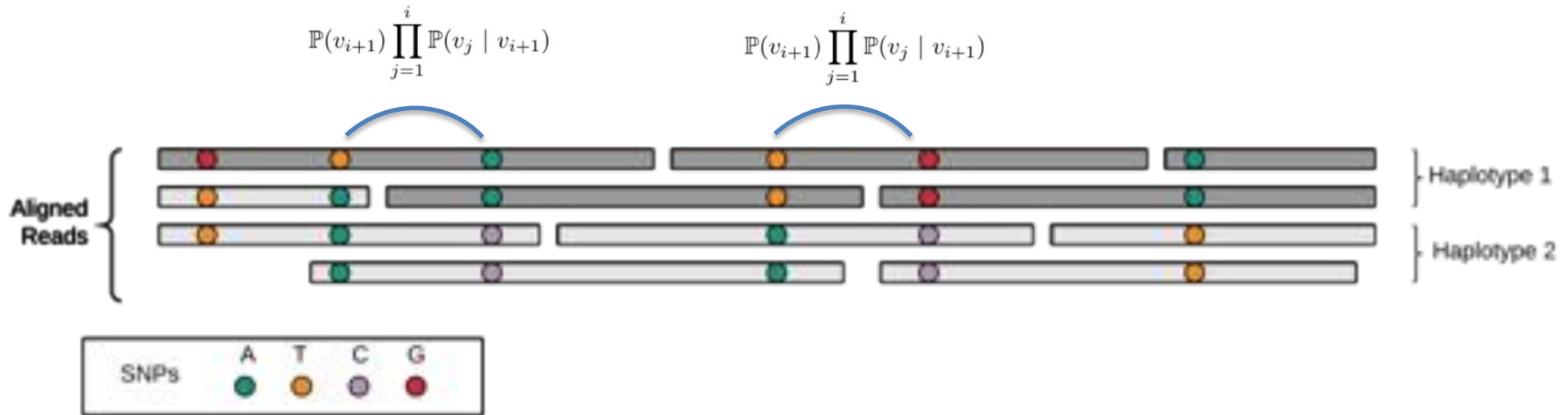

Glycosylases (in particular cellulases – EC 3.2.x) in Clostridium are specialised for the breakdown of Cellulose.

Rubino, Carberry, Waters, Kenny, McCabe & **Creevey,** 2017. *Divergent functional isoforms drive niche specialisation for nutrient acquisition and use in rumen microbiome.* The ISME Journal, 11(4), pp.932-944.

Identifying population-level paths in sequencing reads from metagenomic data

Sam Nicholls

Hansel: A graph-inspired data structure for determining likely chains of sequences from breadcrumbs of evidence

https://github.com/SamStudio8/hansel

An algorithm for recovering haplotypes from metagenomes

https://github.com/SamStudio8/gretel

In vitro validation of predicted haplotypes

(a) G31 (b) G90 (c) G123 (d) G152 (e) G251

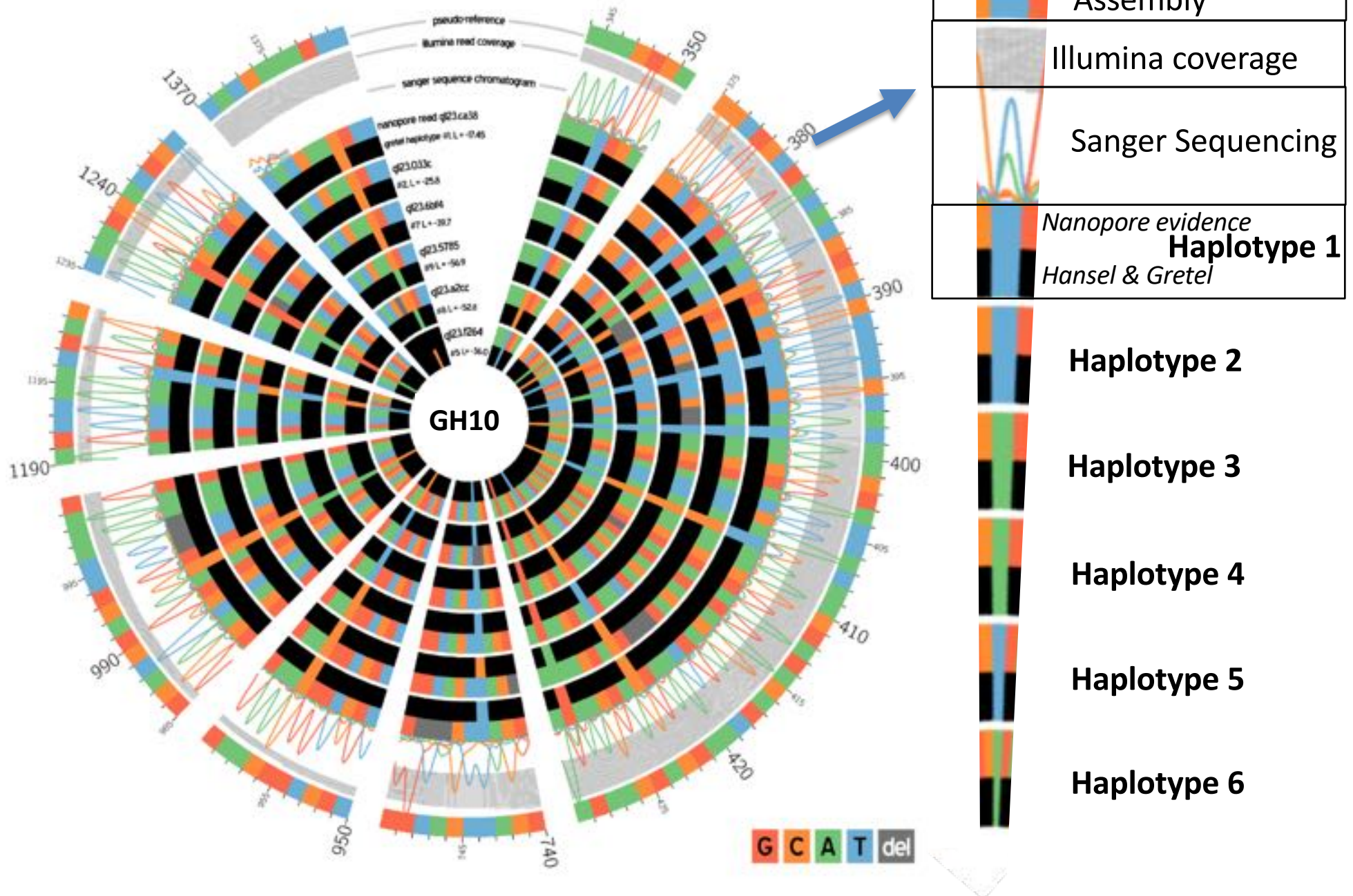(a) Careful loading of prepared library (b) MinION in use, reporting real-time status to connected laptop (excited person for scale) (c) A MinION flow cell

Butyrivibrio proteoclasticus
Glycosyl hydrolase family 10 haplotypes

The Hungate Collection

Cultivation and sequencing of rumen microbiome
members from the Hungate1000 Collection

Rekha Seshadri[1,9] , Sinead C Leahy[2,8,9] , Graeme T Attwood[2], Koon Hoong Teh[2,8], Suzanne C Lambie[2,8],
Adrian L Cookson[2], Emiley A Eloe-Fadrosh[1], Georgios A Pavlopoulos[1], Michalis Hadjithomas[1],
Neha J Varghese[1], David Paez-Espino[1] , Hungate1000 project collaborators[3], Rechelle Perry[2],
Gemma Henderson[2,8], Christopher J Creevey[4], Nicolas Terrapon[5,6] , Pascal Lapebie[5,6], Elodie Drula[5,6],
Vincent Lombard[5,6], Edward Rubin[1,8], Nikos C Kyrpides[1], Bernard Henrissat[5-7], Tanja Woyke[1] ,
Natalia N Ivanova[1], William J Kelly[2,8]

Microbial community
composition data from the Global
Rumen Census overlaid with the
16S rRNA gene sequences (red
dots) from the Hungate genomes.
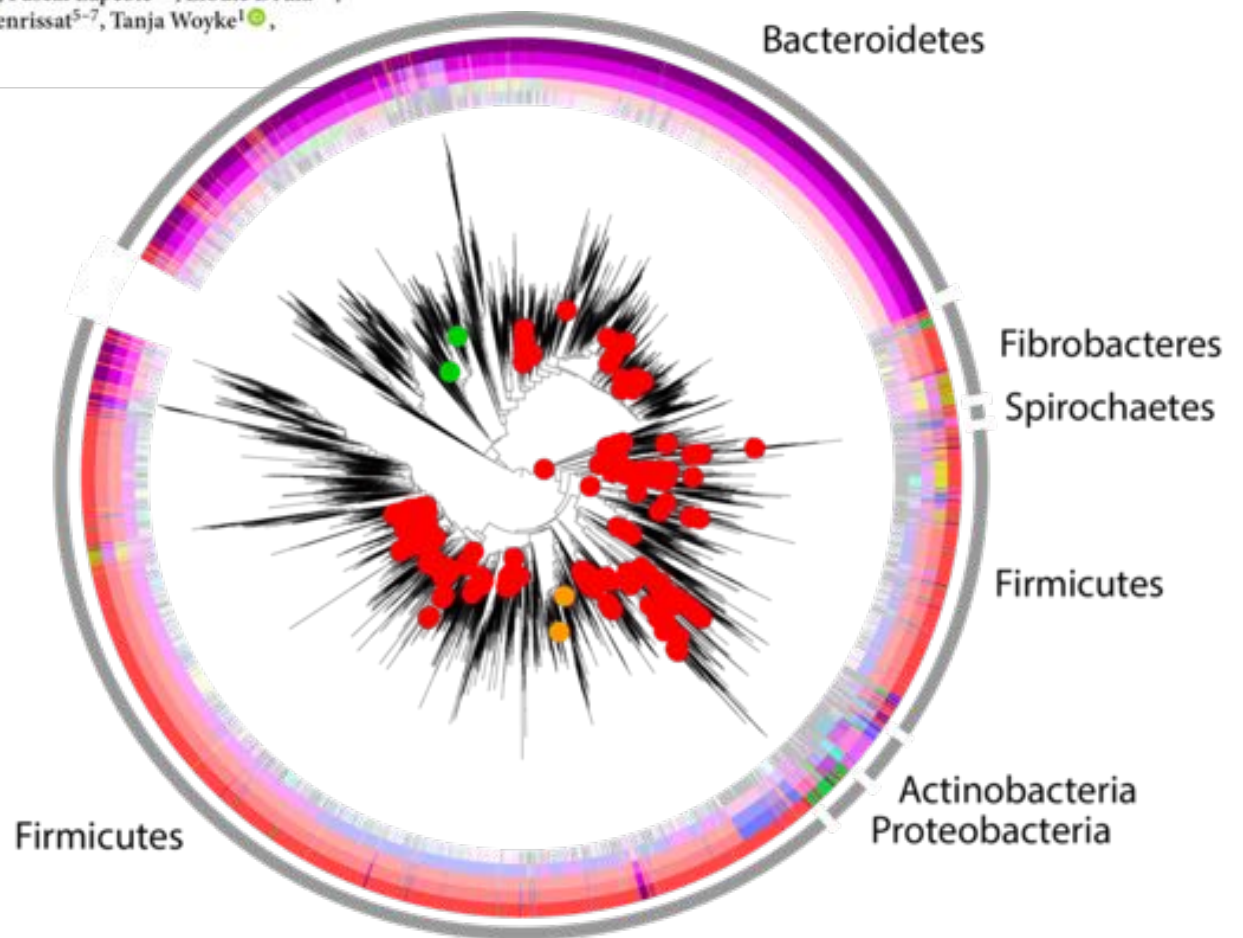
Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome

Weibing Shi,[1,2] Christina D. Moon,[3] Sinead C. Leahy,[3] Dongwan Kang,[1,2]
Jeff Froula,[1,2] Sandra Kittelmann,[3] Christina Fan,[1]
Dragana Gagic,[3] Henning Seedorf,[3] William J. Kel
Carrie Sang,[3] Priya Soni,[3] Dong Li,[3] Cesar S. Pina
Peter H. Janssen,[3] Feng Chen,[1,2] Axel Visel,[1,2,4] Z
Graeme T. Attwood,[3] and Edward M. Rubin[1,2]

[1]Department of Energy, Joint Genome Institute, Walnut Creek, California 94
Laboratory, Berkeley, California 94720, USA; [3]AgResearch Limited, Grassla
Zealand; [4]School of Natural Sciences, University of California, Merced, Cal
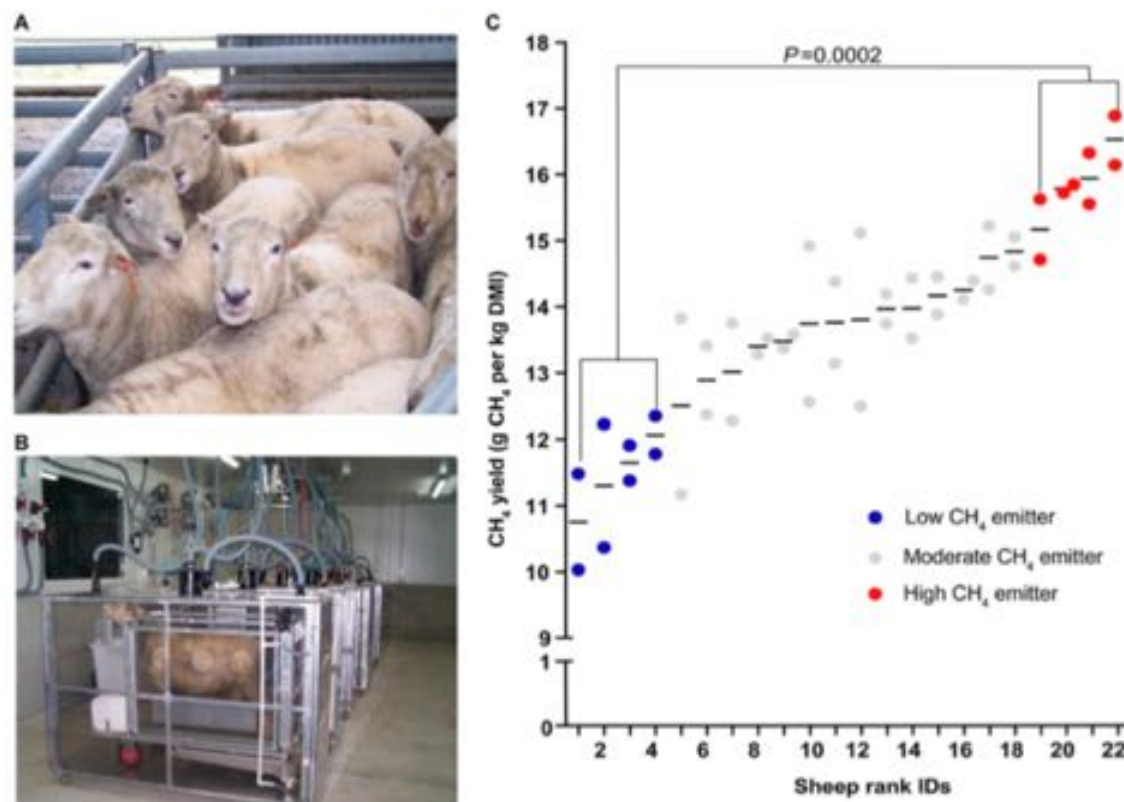
Figure 1. The measurement of CH_4 yields in sheep. (A) New Zealand sheep used for this study. (B) CH_4 yields from the sheep in grams of CH_4/kg dry matter intake (DMI) were measured using open-circuit respiration chambers (http://www.globalresearchalliance.org). (C) CH_4 yield measurements from 22 sheep (each with two time points) sorted by mean values. Four high (red) and four low (blue) emitters are selected for further study. P-value indicates the statistical significance of the differences in CH_4 yield between the two selected groups.

Shi, W., Moon, C.D., Leahy, S.C., Kang, D., Froula, J., Kittelmann, S., Fan, C., Deutsch, S., Gagic, D., Seedorf, H. and Kelly, W.J., 2014. Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. *Genome research*, *24*(9), pp.1517-1525.

Summary

Ecological and evolutionary concepts can be applied to microbiomes, but require adjusting for the unique properties of these systems.

Suggests that niche specialisation and homogeneity in the rumen microbiome is driven by an interplay between HGT, pangenomes and frequency-dependent selection.

Provides a set of hypotheses that can be tested and a mechanism for identifying functions under selection in novel datasets or microbiomes.

Acknowledgements

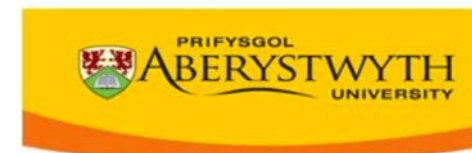**@Creeveylab present and past members:**
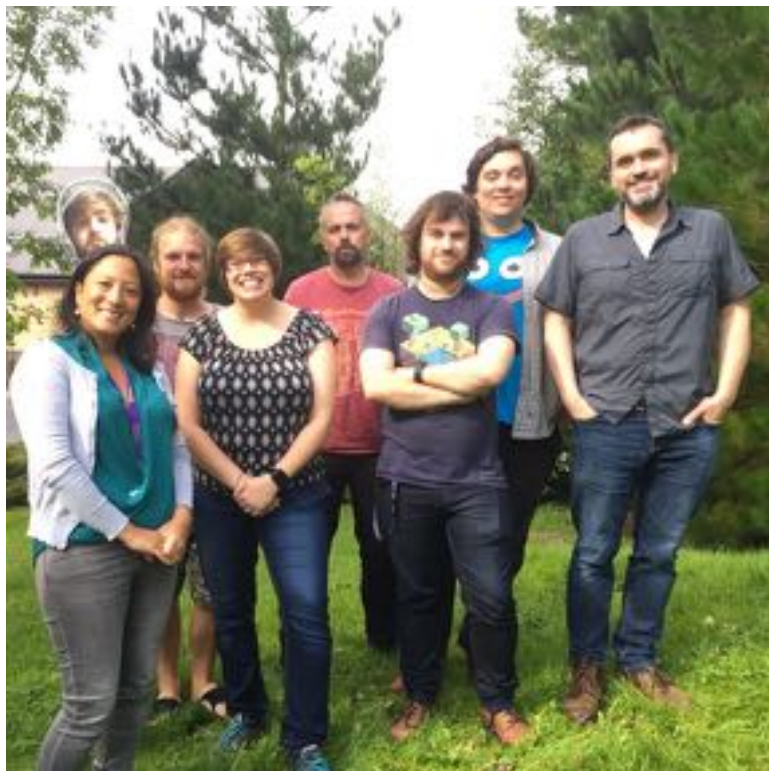
Karen Siu Ting
Sam Nicholls
Ben Thomas
Nick Dimonaco
Jess Friedersdorff
Toby Wilkinson
Martin Hughes
Max Friedersdorff
Tom Hitch
Francesco Rubino
Bruce Moran
Anthony Doran

Our collaborators:

Sharon Huws	Hilario Mantovani	Sinead Waters
Alison Kingston-Smith	James McInerney	Graeme Attwood
Jamie Newbold	Mary O'Connell	Matthias Hess
Joan Edwards	Kai Hilpert	Diego Morgavi

Contact: c.creevey@qub.ac.uk   *@hairy_llama*
           www.creeveylab.org   *@creevey_lab*