

The vaginal bacterial meta-transcriptome

Jean Macklaim

mmacklai@uwo.ca

The University of Western Ontario

Advisors:

Dr. Greg Gloor (ggloor@uwo.ca)

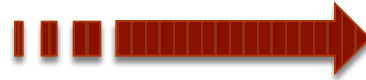
Dr. Gregor Reid (gregor@uwo.ca)



The vaginal microbiota

Normal

Lactobacilli-dominated



Bacterial vaginosis

Lack of lactobacilli

+ *Gardnerella*, *Atopobium*,
Prevotella...

- Bacterial vaginosis (BV) is the most common vaginal disorder of women of child-bearing age: 10-29% of women are affected¹
- No single disease-causing microbe, but a **shift in the microbial community** leading to a dysbiosis
- **Malodor**, discharge, irritation

BV leads to increased risk of:

- Acquisition and transmission of STIs and HIV
- Urinary tract infections
- Complications during pregnancy, and pre-term labour

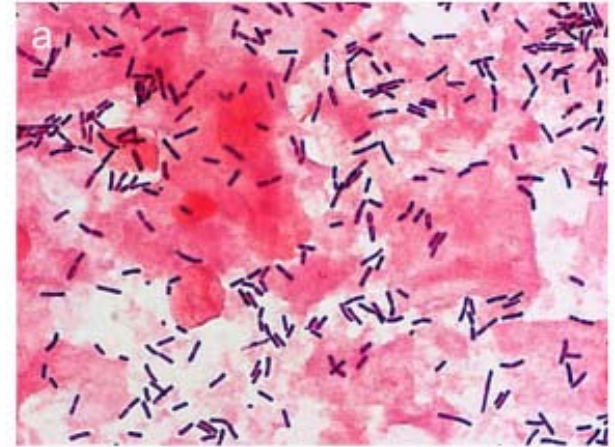
Ranking the vaginal microbiota

Nugent scoring

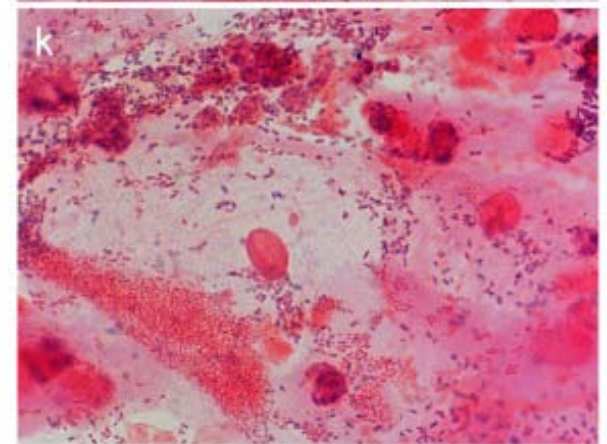
- Presence/abundance of *Lactobacillus*, *Gardnerella*, *Mobiluncus* morphotypes on slide

Amsel (“symptoms”)

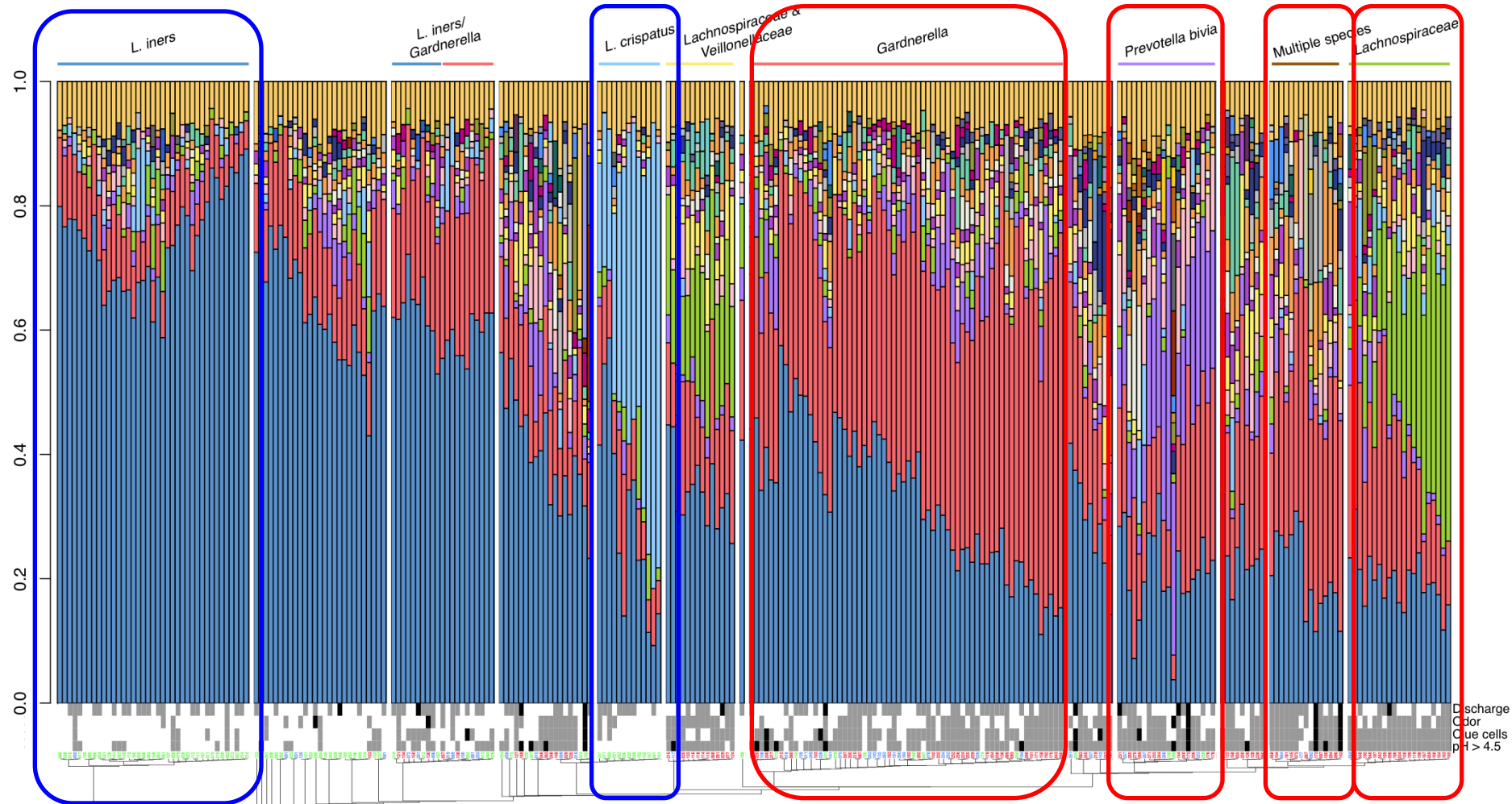
- Discharge
- Maldor
- $pH > 4.5$
- Clue cells



Normal



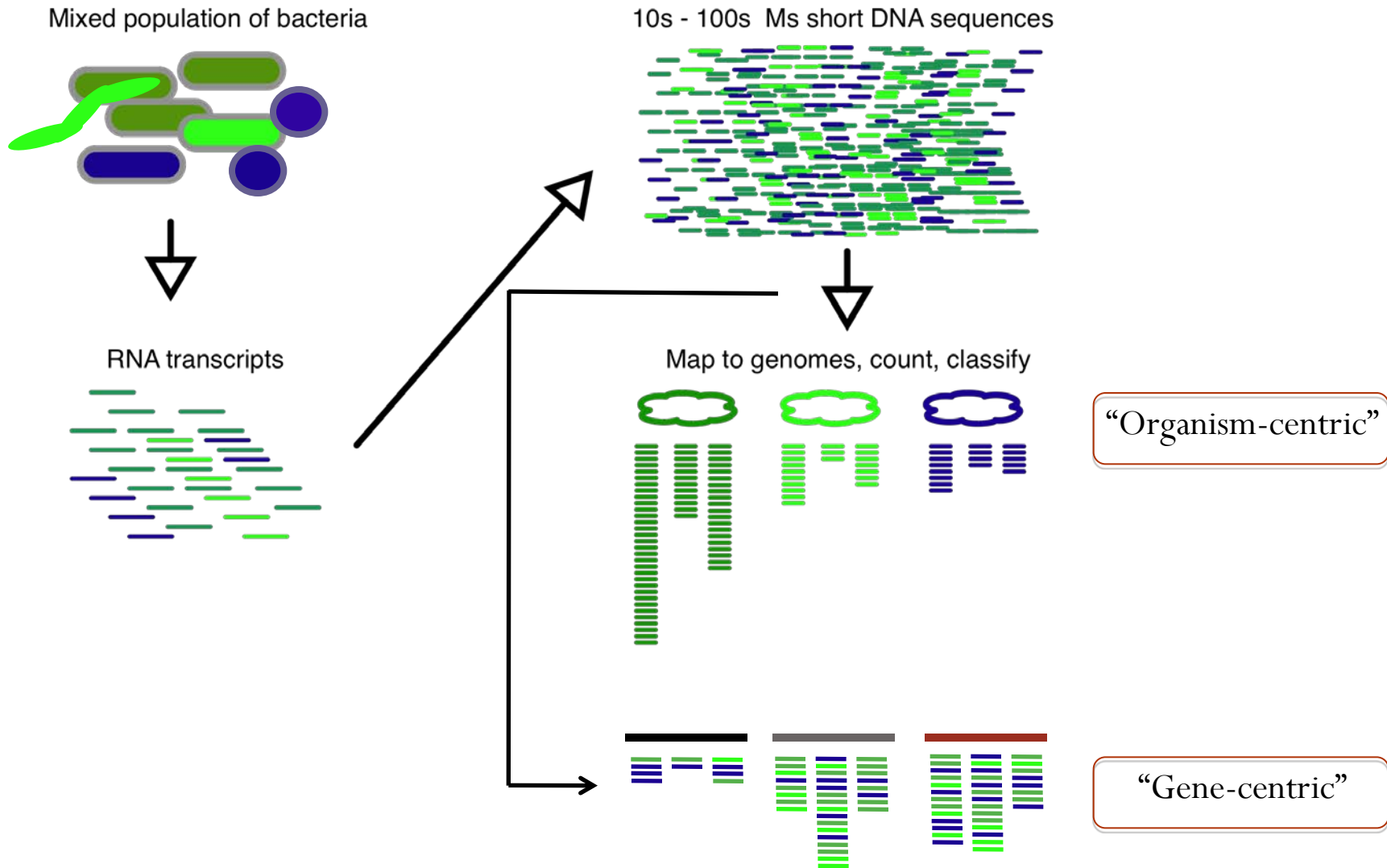
Bacterial vaginosis



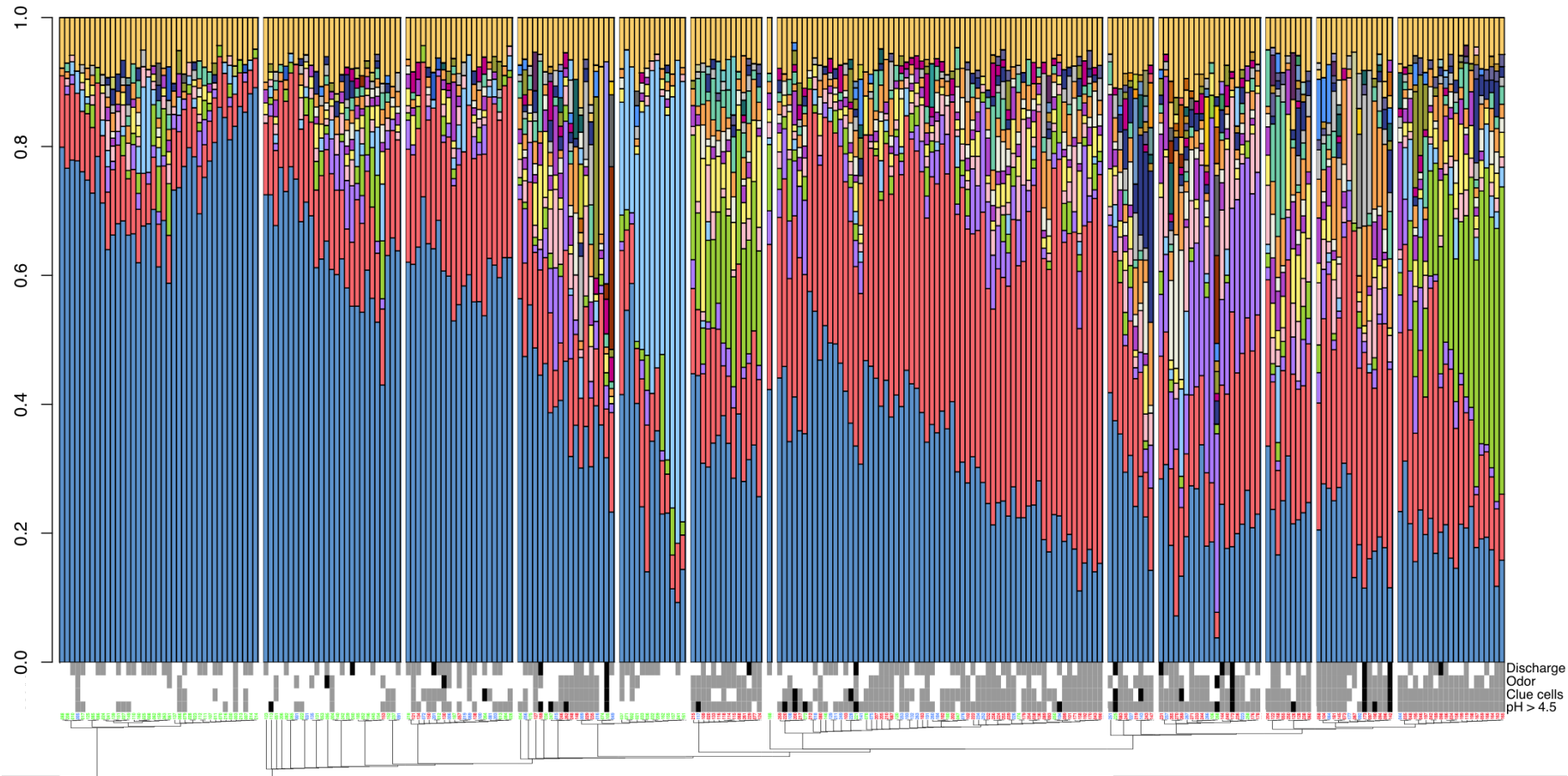
- Bacterial enumeration/composition does not tell us the whole story
- What about functional differences?

Leptotrichia amnionii, OTU 5	Lactobacillus jensenii, OTU 31
Lactobacillus crispatus, OTU 4	Prevotella amniotica, OTU 10
Uncultured Lachnospiraceae, OTU 3	Prevotella timonensis, OTU 9
Prevotella bivia, OTU 2	Atopobium vaginae, OTU 8
Gardnerella vaginalis, OTU 1	Sneathia, OTU 7
Lactobacillus iners, OTU 0	Uncultured Veillonellaceae, OTU 6

Meta-transcriptomics

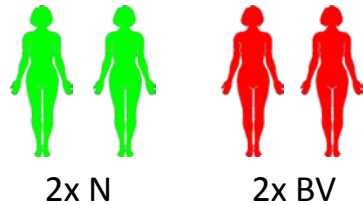


Why is the vaginal microbiota a good system for meta-transcriptomics?



- Relatively low diversity
- Large phylogenetic separation between organisms (sequence variation for mapping)

Vaginal bacterial meta-transcriptional profiling



Vaginal swabs



Extract bacterial RNA



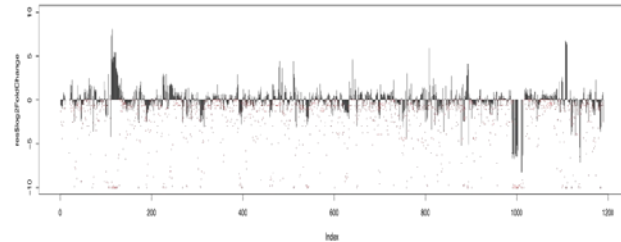
Deplete rRNA (95%+ of sample) to enrich mRNA



Sequence mRNA
(ABI SOLiD 4)

≈ 50,000,000 x50bp
reads/sample

Map sequenced transcripts to reference genomes (Bowtie: colorspace aware)



Compare samples

1. DESeq **R** package (negative binomial distribution)
2. RPKM (reads per kilobase per millions of mapped reads)

How do the bacterial transcriptional profiles differ during bacterial vaginosis and how does this relate to health outcome?

Meta-transcriptome – gene-centric

Collect CDS sequences
for vaginal strains

(N=90)



Meta-transcriptome – gene-centric

Collect CDS sequences
for vaginal strains
(N=90)



Cluster by 95% seq ID to make non-
redundant sequence database
(135,201 non-redundant CDS)



Meta-transcriptome – gene-centric

Collect CDS sequences
for vaginal strains
(N=90)



Cluster by 95% seq ID to make non-
redundant sequence database
(135,201 non-redundant CDS)



Mapping



Functional
assignment/annotation

Meta-transcriptome – gene-centric

Collect CDS sequences
for vaginal strains
(N=90)



Cluster by 95% seq ID to make non-
redundant sequence database
(135,201 non-redundant CDS)



Mapping



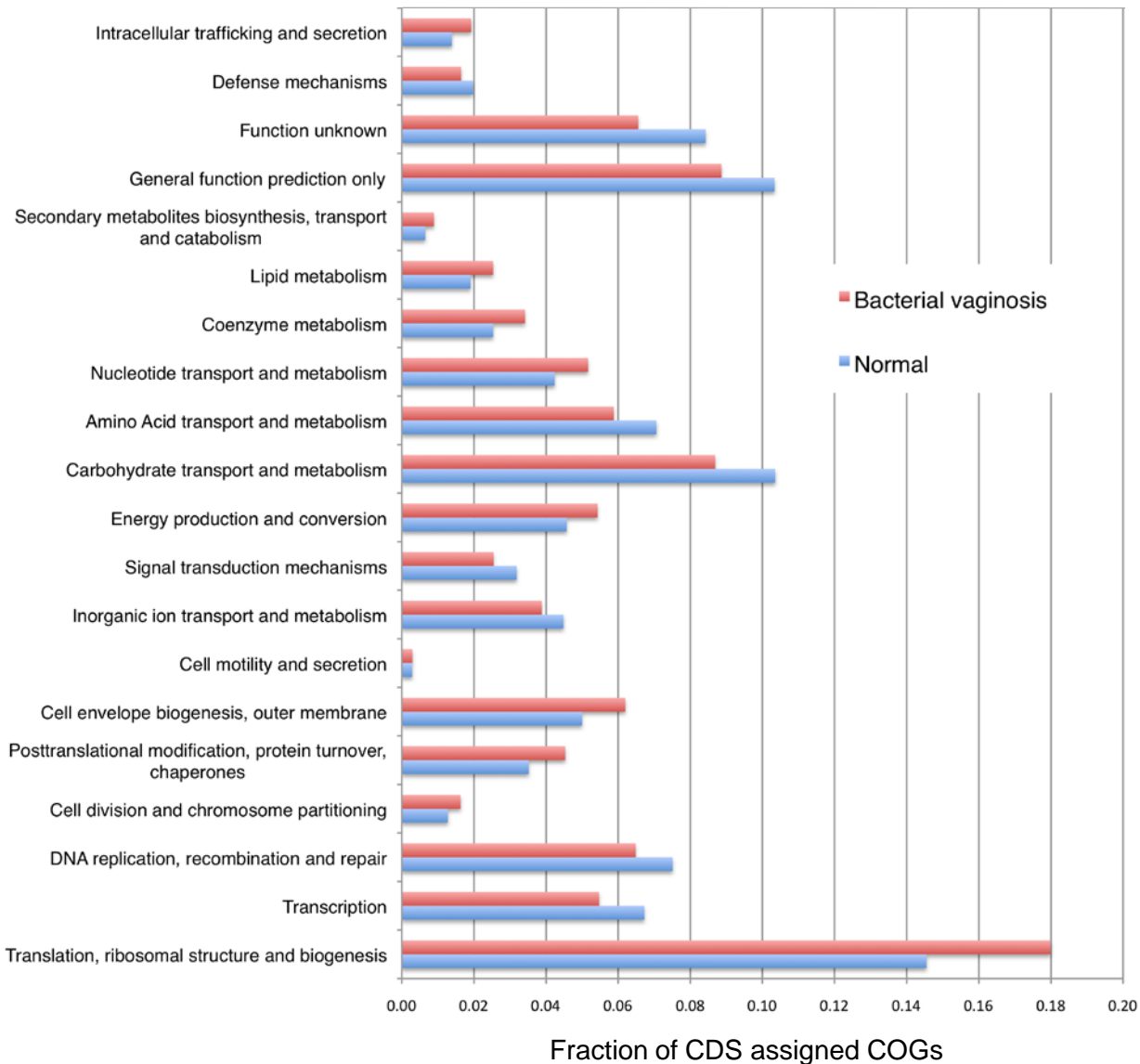
Functional
assignment/annotation

- 20-55% of raw reads mapped per sample: 11mil to 20mil reads

N4	N30	BV27	BV31
2,528	3,234	5,135	5,992

Number of non-redundant CDS mapped at RPKM \geq 10

Functional assignment of the meta-transcriptome



At broad functional level the communities are similar

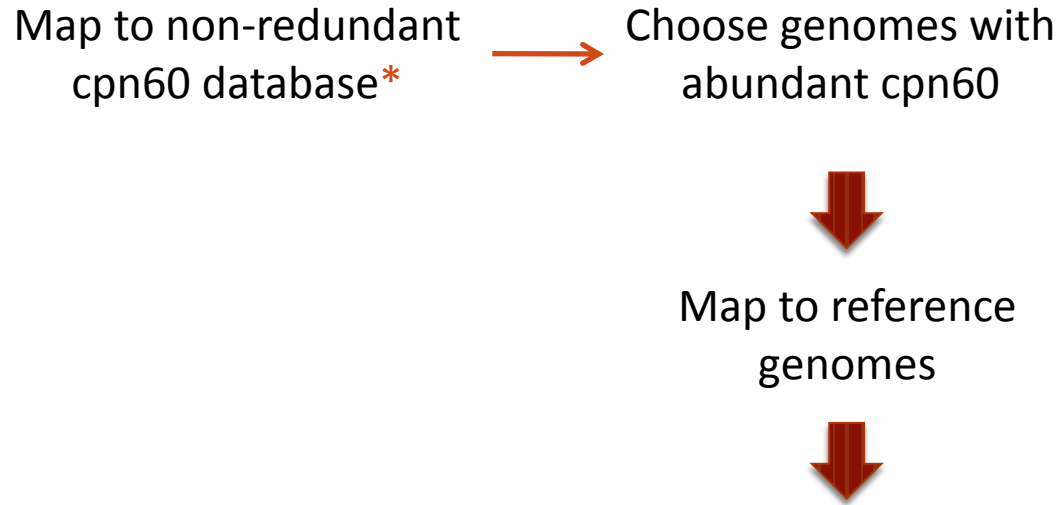
Step 2: consider individual genes and organisms

What organisms are present?

Map to non-redundant
cpn60 database*  Choose genomes with
abundant cpn60



What organisms are present?



What organisms are present?

Map to non-redundant
cpn60 database*



Choose genomes with
abundant cpn60



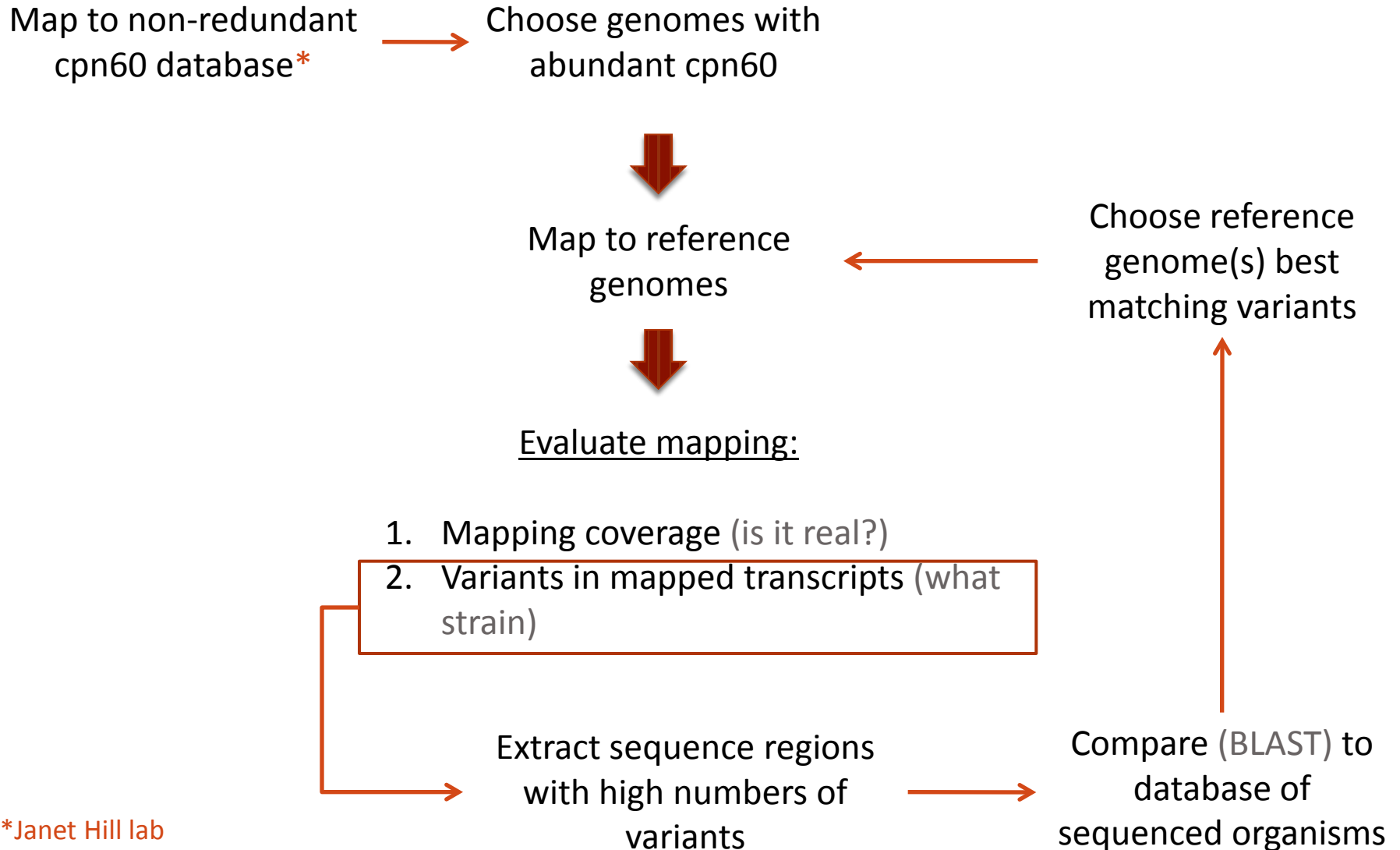
Map to reference
genomes



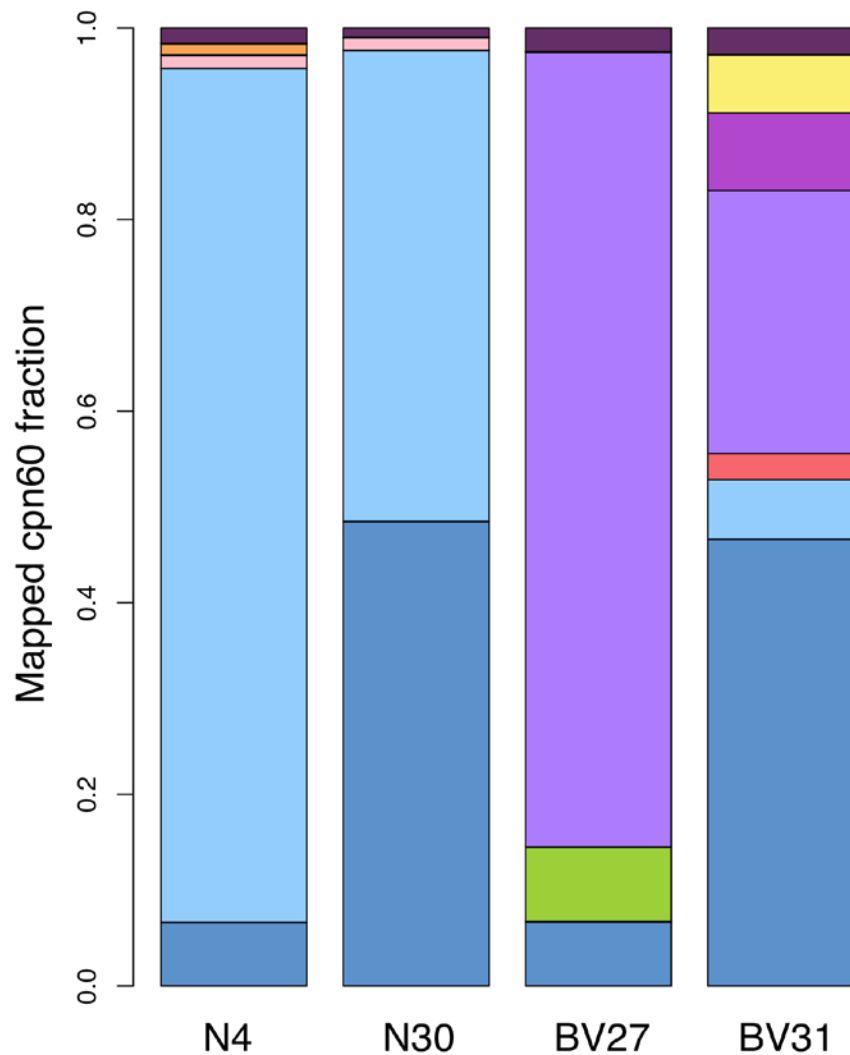
Evaluate mapping:

1. Mapping coverage (is it real?)
2. Variants in mapped transcripts (what strain)

What organisms are present?



What organisms are present?

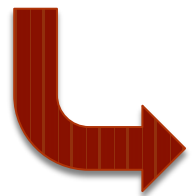


- Represents transcriptionally active organisms
- Eight abundant organisms plus 3 BV-associated organisms (*Atopobium vaginae*, Clostridiales BVAB3, *Peptoniphilus lacrimalis*) chosen for organism-specific transcriptome mapping

Mapping outcome

	N4	N30	BV27	BV31
Raw reads	47,634,967	46,943,184	48,832,687	53,655,422
Mapped reads	23,450,946 (49%)	16,545,092 (35%)	16,625,363 (34%)	14,075,880 (26%)
Unique reads mapped to CDS	10,896,193 (46%)	8,541,748 (52%)	8,378,653 (50%)	5,908,471 (42%)

- 1-9% of reads mapped to human genome
- Of mapped reads, ~50% are 23s or 5s (incomplete/biased rRNA depletion)
 - Few mapping outside predicted CDS
- The rest are inaccessible in colorspace...



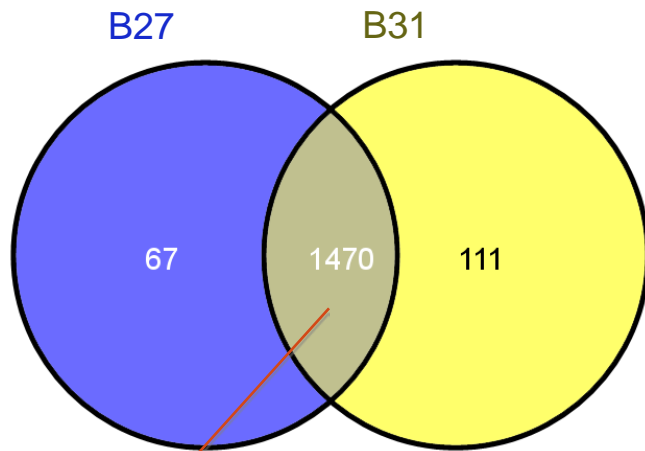
Mapping outcome – 11 reference organisms

Organism	CDS size (bp)	N4	cov	N30	cov	BV27	cov	BV31	cov
<i>Atopobium vaginae</i>	1,224,920	64	0.00	78	0.00	104,298	4.26	10,282	0.42
Clostridiales BVAB3	1,589,129	131	0.00	136	0.00	23,438	0.74	9,206	0.29
<i>Gardnerella vaginalis</i>	1,381,069	1,335	0.05	1,774	0.06	2,482,661	89.88	997,009	36.10
➔ <i>Lactobacillus iners</i>	1,129,392	108,925	4.82	560,553	24.82	1,789,648	79.23	2,644,259	117.07
➔ <i>Lactobacillus crispatus</i>	2,030,618	10,563,439	260.10	7,923,734	195.11	8,731	0.21	1,381,479	34.02
<i>Lactobacillus jensenii</i>	1,445,282	220,576	7.63	52,931	1.83	18,906	0.65	102,993	3.56
<i>Megasphaera</i>	1,541,005	290	0.01	475	0.02	813,353	26.39	81,674	2.65
<i>Peptoniphilus lacrimalis</i>	1,527,734	64	0.00	130	0.00	1,838	0.06	17,387	0.57
<i>Prevotella amnii</i>	2,065,134	1,243	0.03	1,491	0.04	3,080,225	74.58	381,622	9.24
<i>Prevotella disiens</i>	2,515,890	32	0.00	225	0.00	9,640	0.19	69,721	1.39
<i>Prevotella timonensis</i>	2,330,373	94	0.00	221	0.00	45,897	0.98	212,839	4.57
Total mapped reads		10,896,193		8,541,748		8,378,635		5,908,471	

- Uniquely mapped reads

BV organism – *Prevotella amnii*

CDS with RPKM ≥ 10

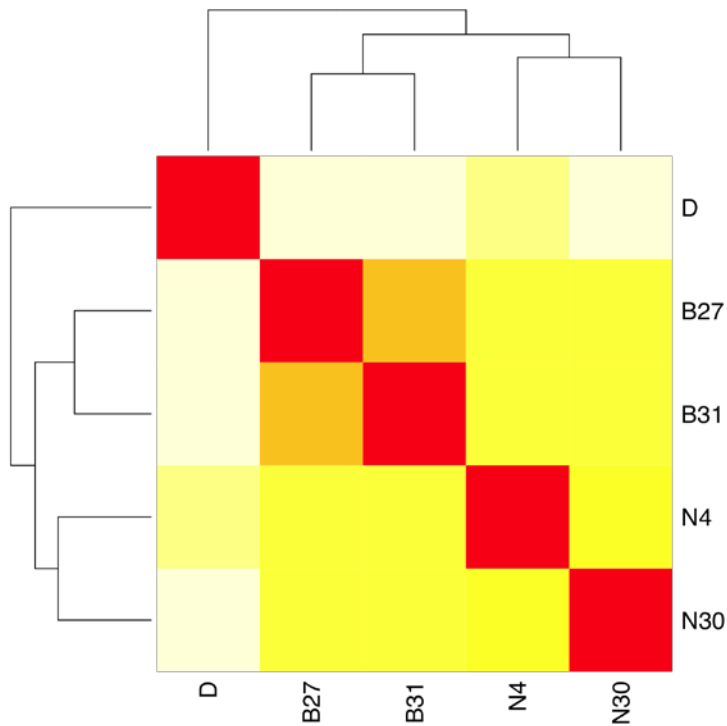


- 90% shared between 2 BV samples
- Represents 72% of all CDS in *P. amnii*

- What transcriptionally active genes have a role in BV pathogenesis?
 - Interactions with host
 - Interactions with other members of the biota (contributions to co-occurrence and competitive exclusion)*

* Andrew Fernandes (38)

Conserved *L. iners* gene expression?



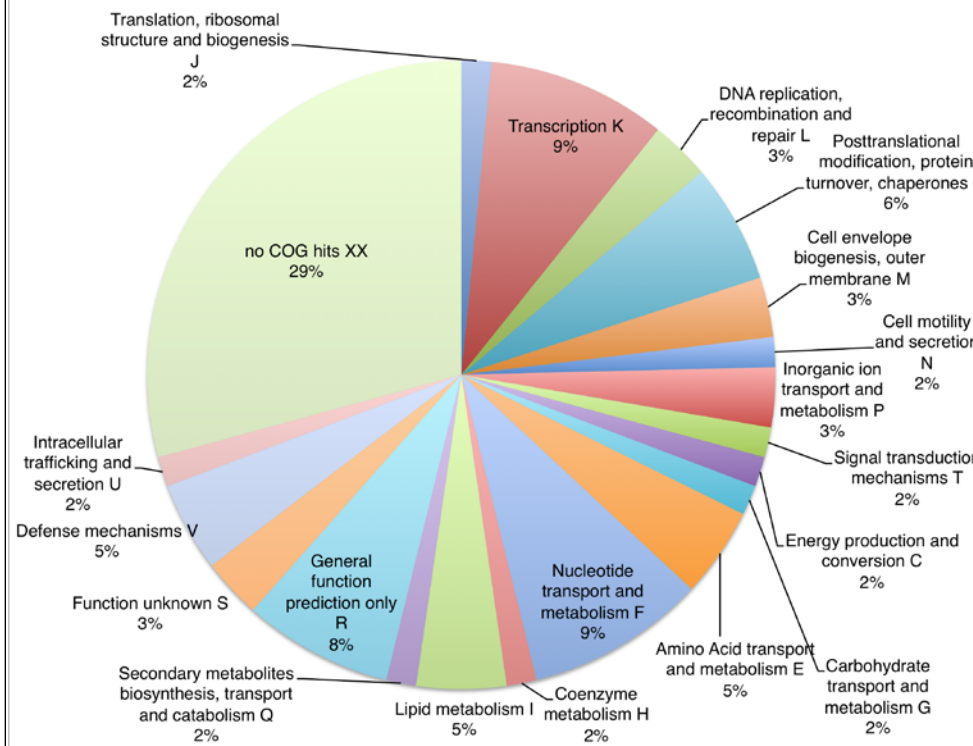
- *L. iners* gene expression in normal biota cluster together as do BV biota
- During BV the gene expression is highly similar despite different biota composition
- *L. iners* cultured in MRS broth (**D**) has a very different profile than *in vivo*

more similar
less similar

How does *L. iners* gene expression function differ in BV compared to normal?

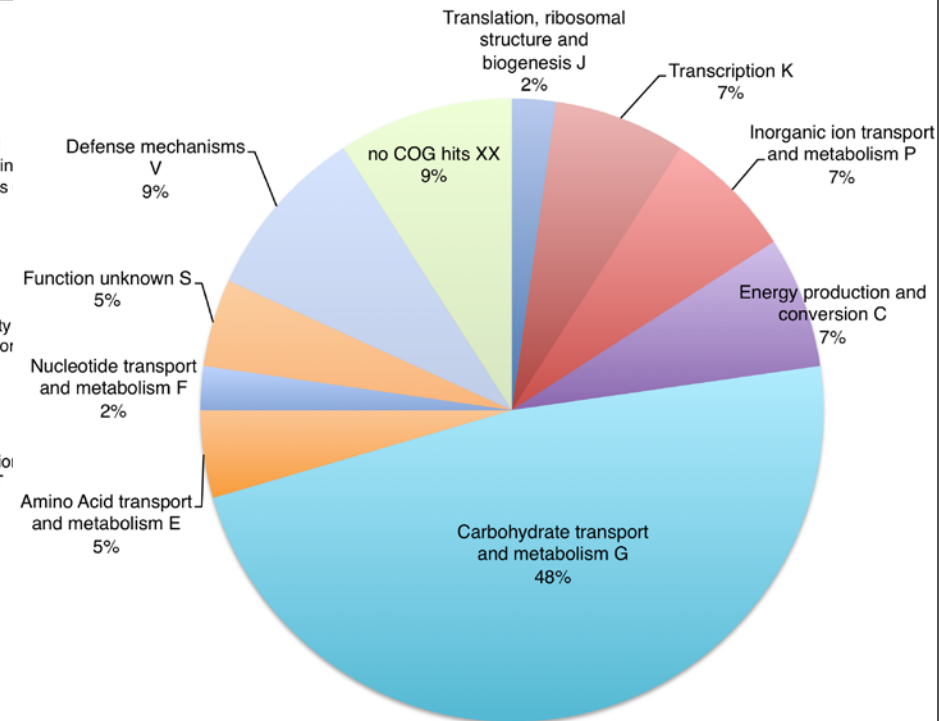
An altered transcriptional profile in *L. iners* during BV

COG categories: *L. iners*' genes expressed in **Normal**



Number of genes: 60

COG categories: *L. iners*' genes expressed in **BV**

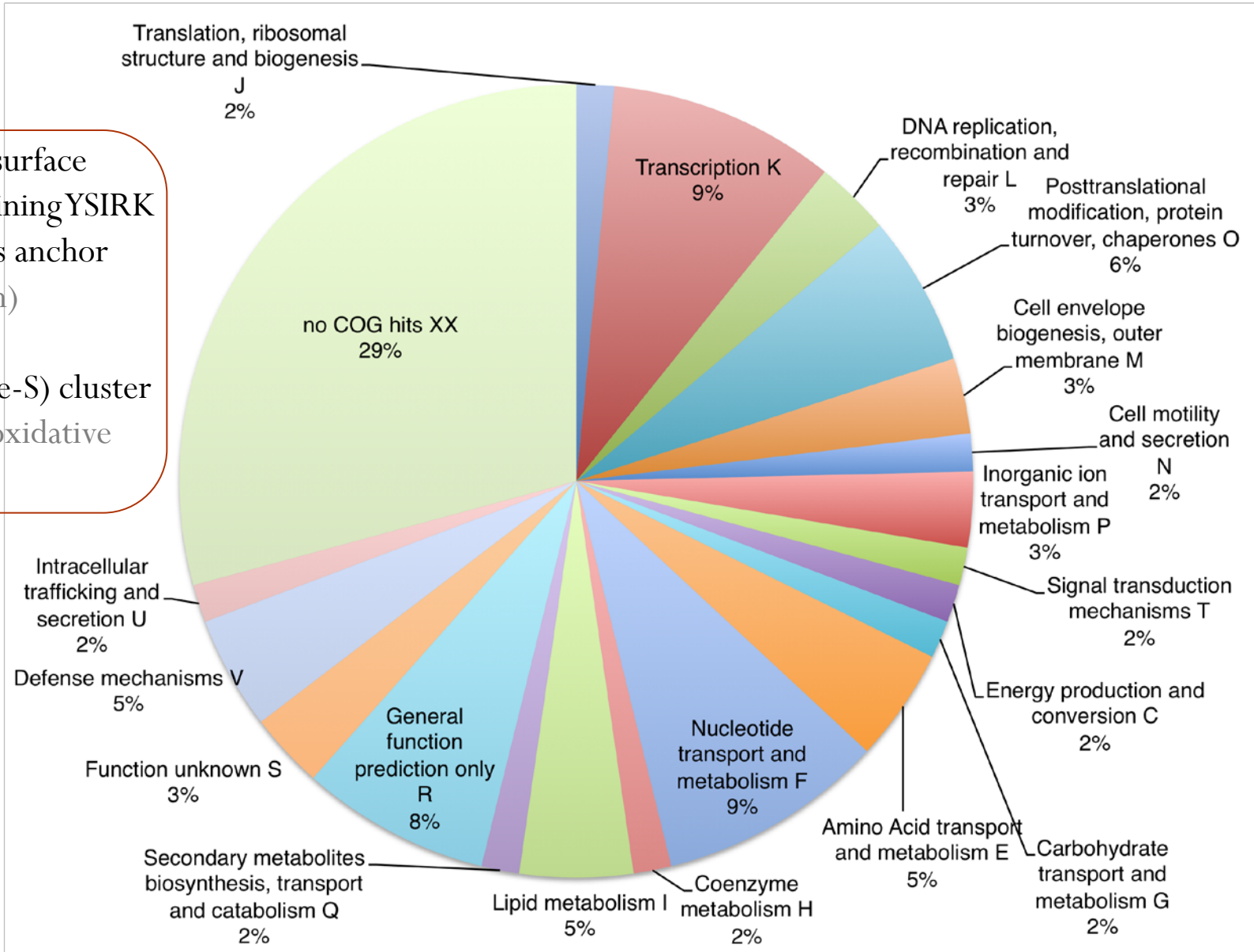


Number of genes: 43

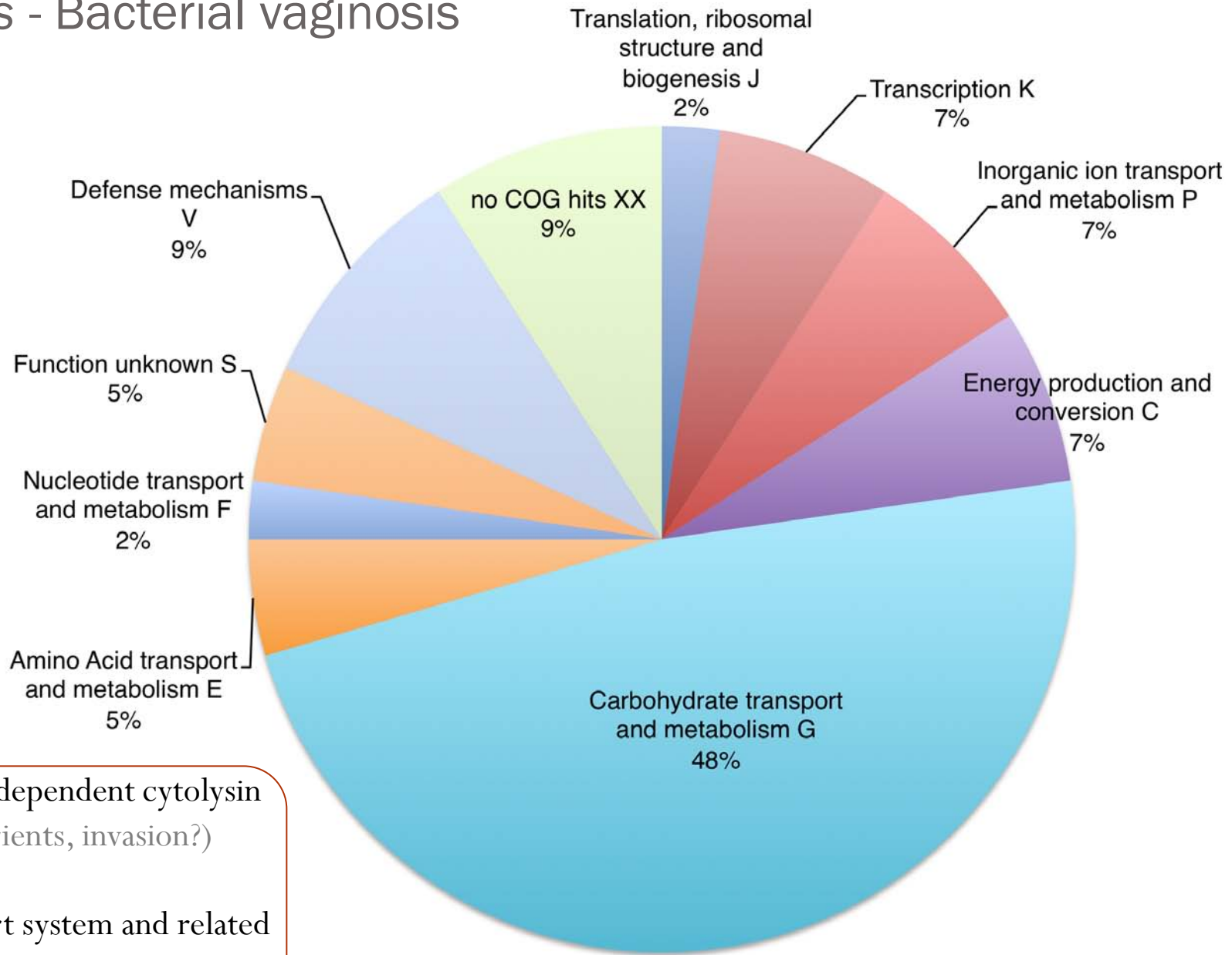
L. iners total CDS = 1190

L. iners - Normal vaginal conditions

- Putative cell-surface protein containing YSIRK and Gram-pos anchor (host adhesion)
- Iron-sulfur (Fe-S) cluster and ferritin (oxidative stress)

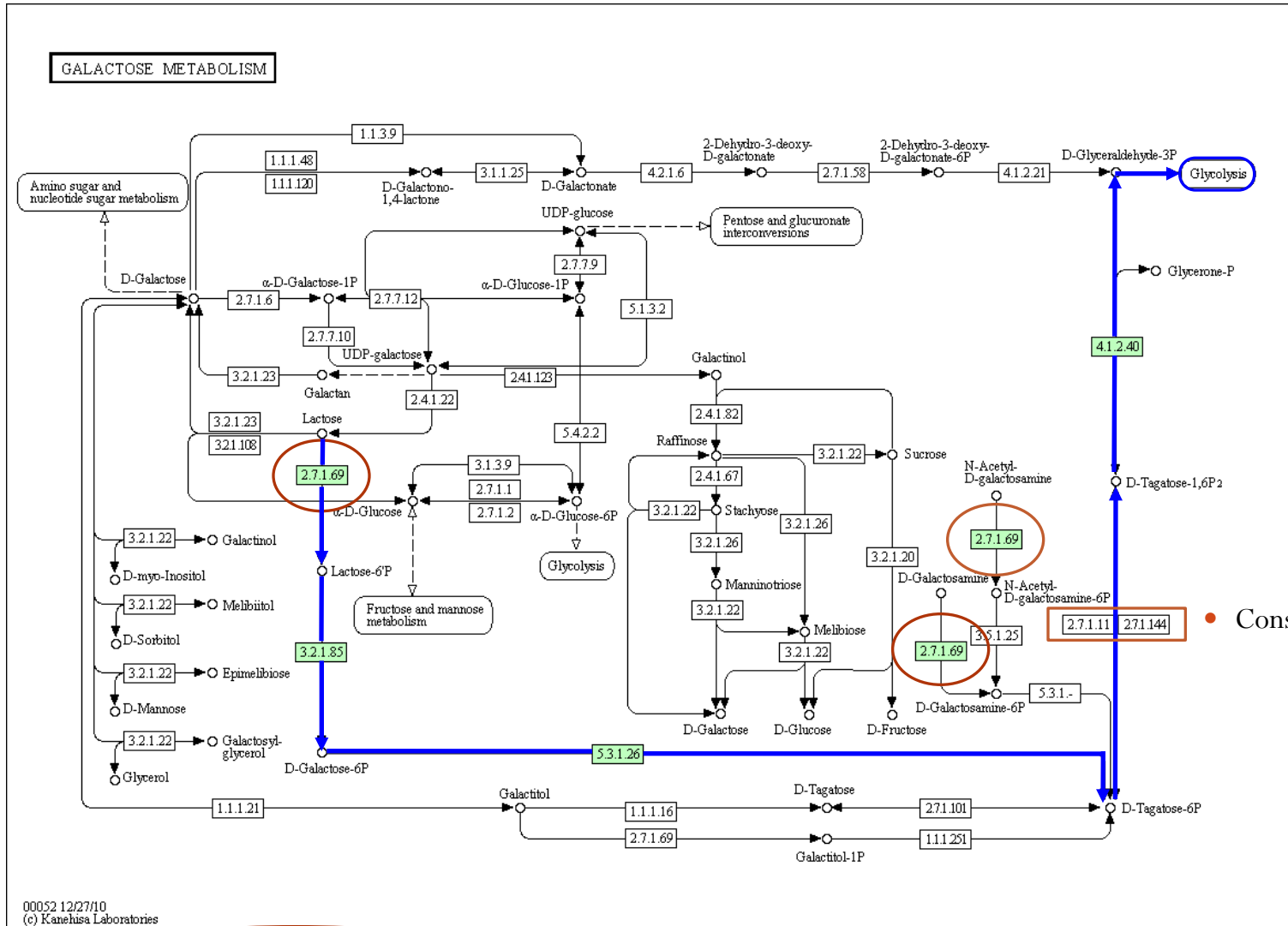


L. iners - Bacterial vaginosis



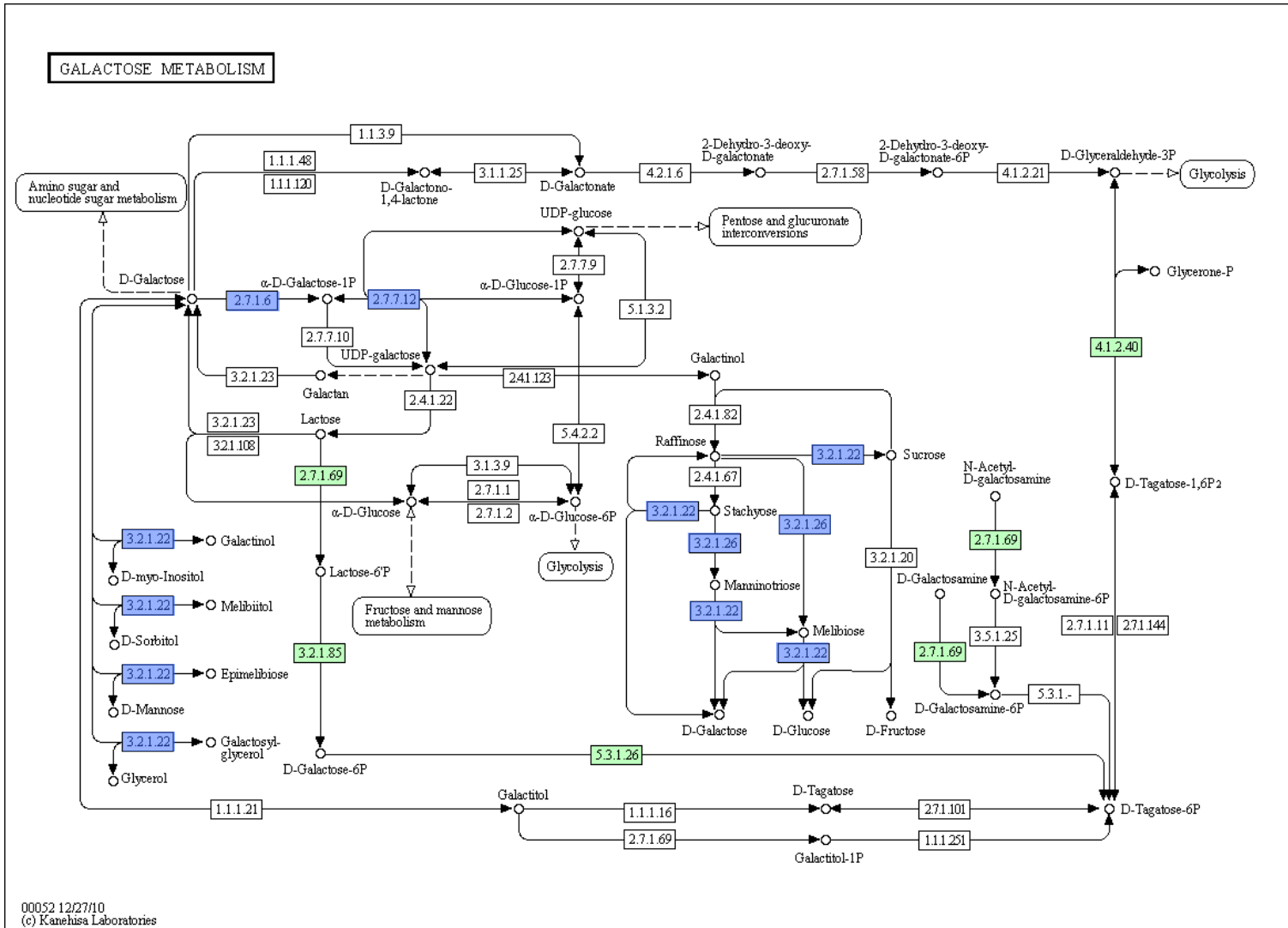
- Cholesterol-dependent cytolysin (acquire nutrients, invasion?)
- PTS transport system and related metabolic enzymes (mucin)

An altered transcriptional profile in *L. iners* during BV



- PTS carbohydrate transporters - Breakdown products of mucin glycoproteins

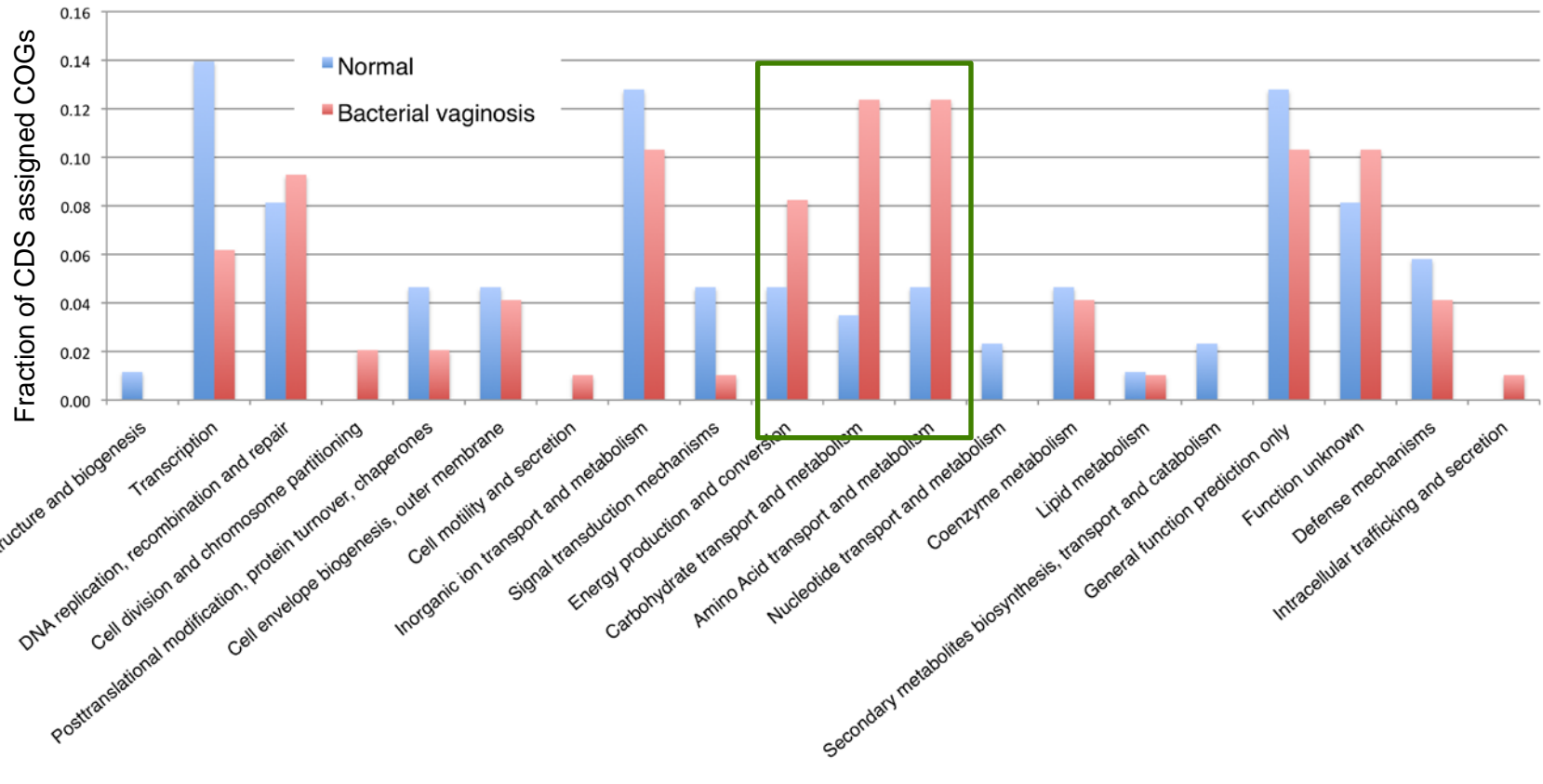
Lactobacillus crispatus under BV conditions



█ *L. iners*

█ *L. crispatus*

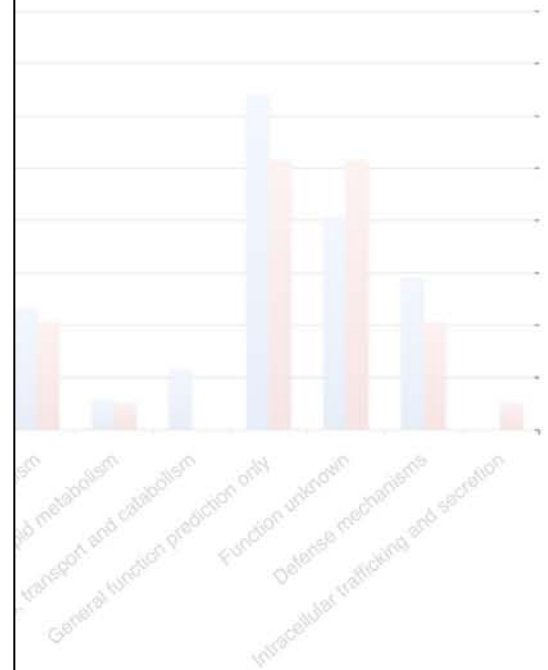
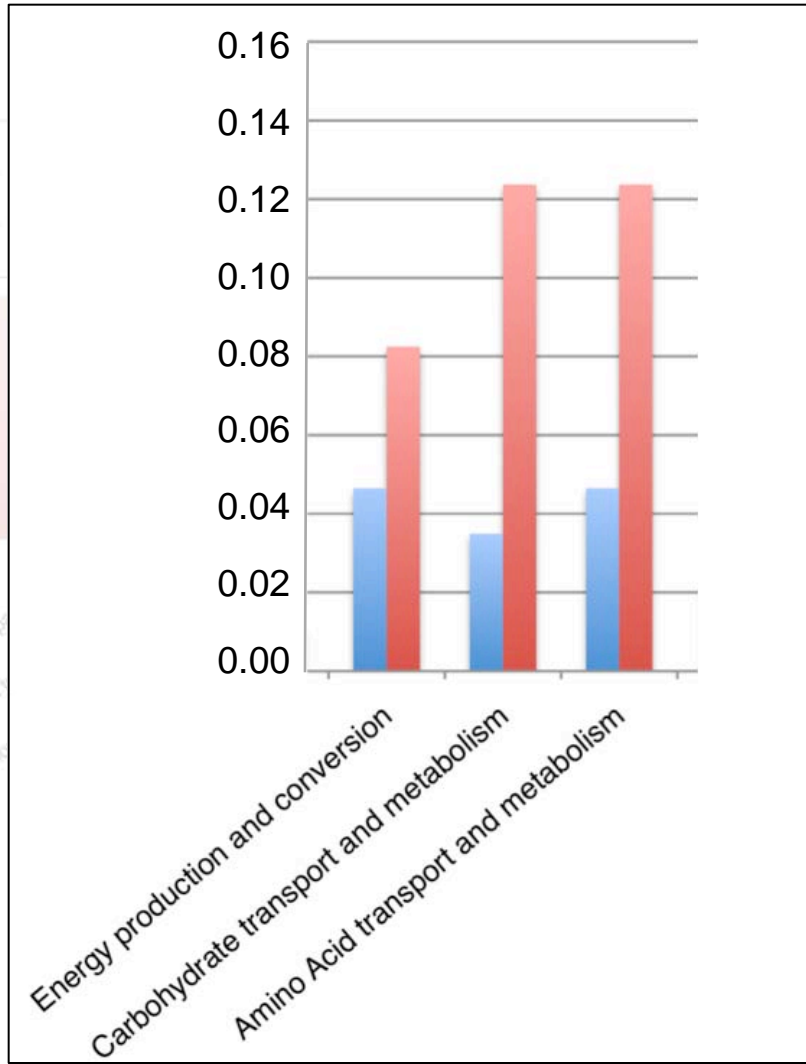
L. crispatus – functional assignment of DE genes



During BV

- Shift to carbohydrate transport and metabolism and energy production (similar to *L. iners*)

L. crispatus – functional assignment of DE genes



During BV

- Shift to carbohydrate transport and metabolism and energy production (similar to *L. iners*)
- Many predicted genes without functional assignment

What next?

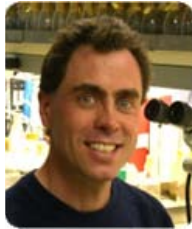
- Dig deeper!
- Validate with qPCR
- In process of recruiting and collecting more samples
- Use RNA-seq to refine gene predictions and genomic architecture (start and stop positions, operon structure)

Summary

- Possible to get high quality data from clinical samples
- *L. iners* is able to differentially express 10% of its gene complement during BV and shows a shift towards carbohydrate utilization
- Carbohydrate availability driving the population changes?
- Organism- and gene-centric analyses paint a different picture

Acknowledgements

Dr. Greg Gloor



Russ Dickson

Dr. Andrew Fernandes

Kevin Chen

Ryan Jung

Charles Yin



Dr. Gregor Reid



Kingsley Anukam

Jordan Bisanz

Ruben Hummelen

Roderick MacPhee

Amy McMillan

Shannon Mifflin

Dr. Wayne Miller

Marc Monachese

Beth Radford

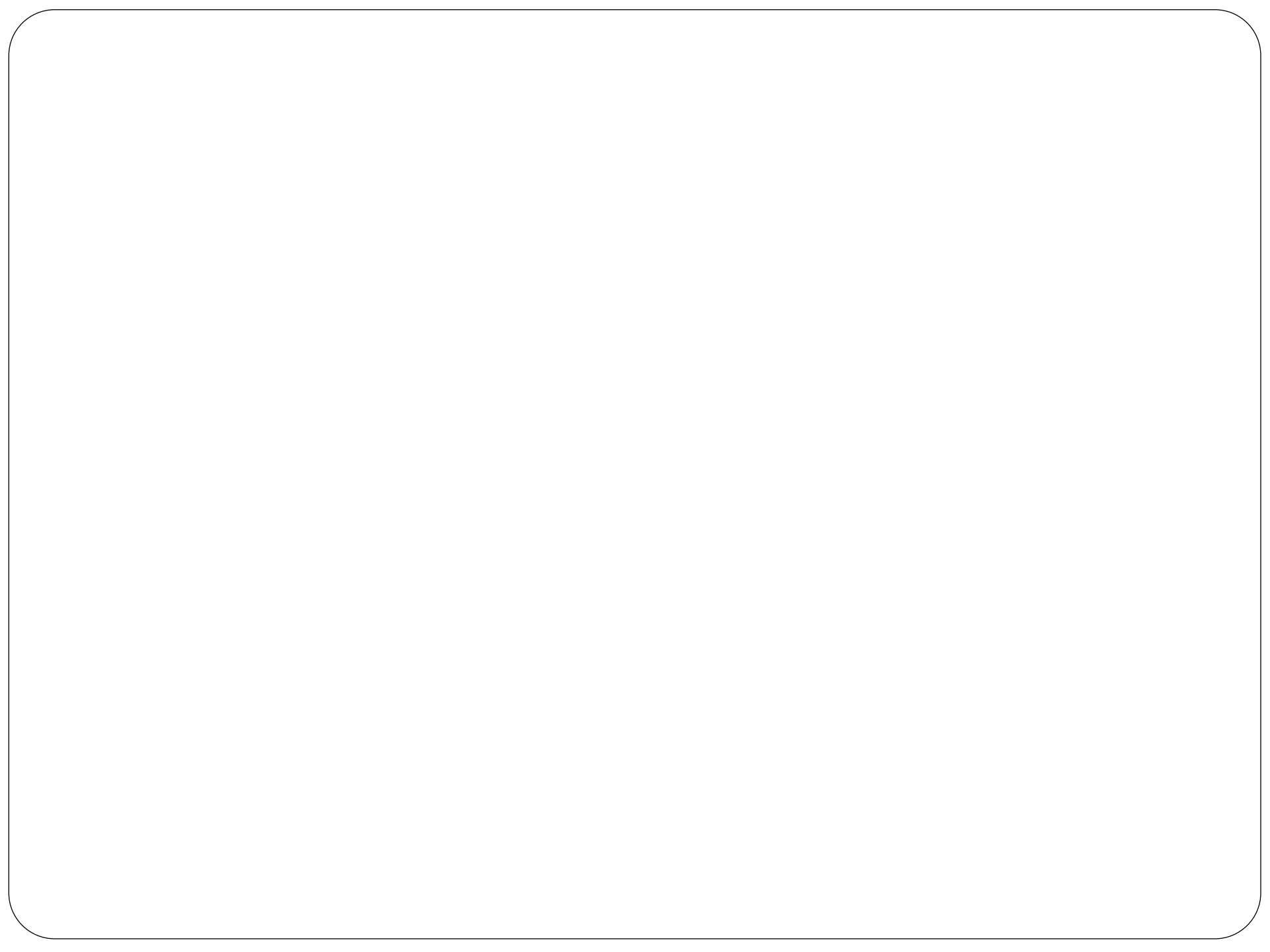
Camilla Urbaniak



**Canadian Research & Development
Centre for Probiotics**

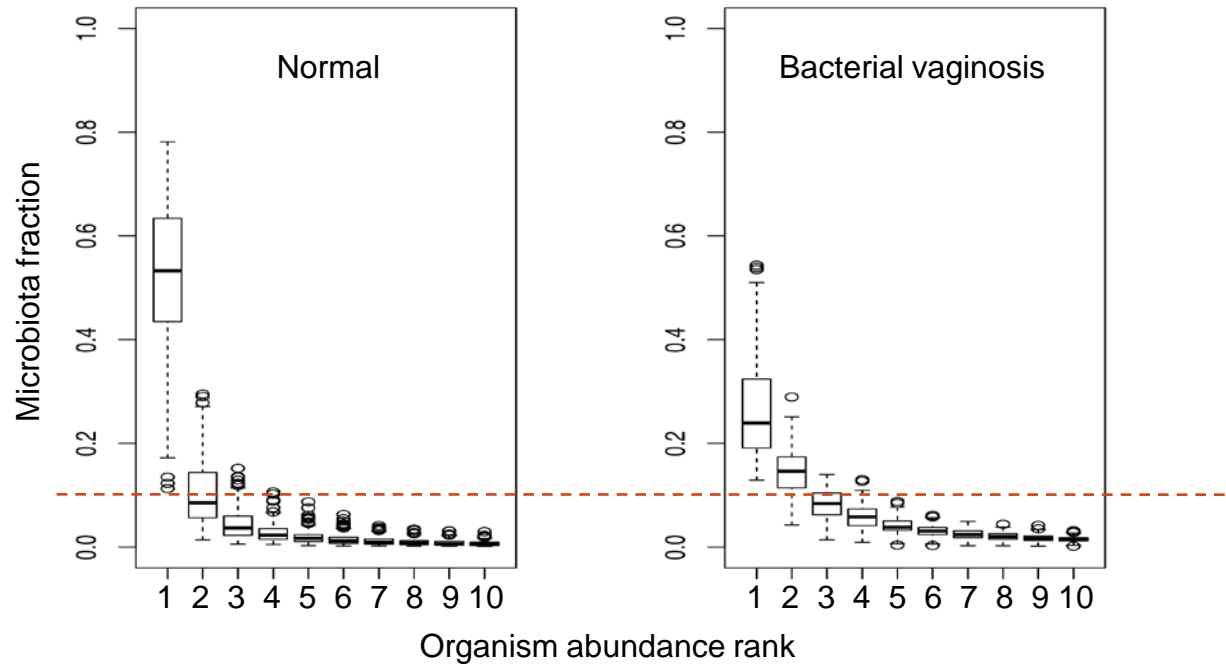
THE RESEARCH INSTITUTE OF LONDON HEALTH SCIENCES CENTRE
AND ST. JOSEPH'S HEALTH CARE, LONDON.





Why is vaginal microbiota a good system for meta-transcriptomics?

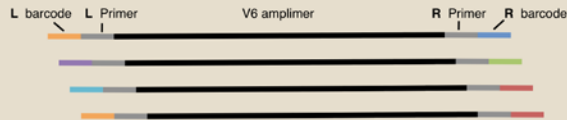
- Relatively low diversity
- Usually 1-4 organisms dominating



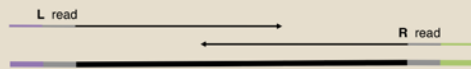
- Large phylogenetic separation between organisms (sequence variation for mapping)
- Bacterial enumeration/composition does not tell us the whole story

Method – 16s (V6) community profiling

Acquire genomic DNA from microbial environment and PCR amplify V6 region of 16s rRNA gene using pairs of primers containing unique sequence (barcodes) for each clinical sample



Sequence PCR amplicons using Illumina paired-end sequencing resulting in overlapping sequence pairs



Use computational biology tools to quality-filter and overlap paired reads into complete V6 sequence

Cluster identical V6 sequences into identical sequence units (ISUs). Cluster ISU sequences by similarity into operational taxonomic units (OTUs) using 95% sequence identity

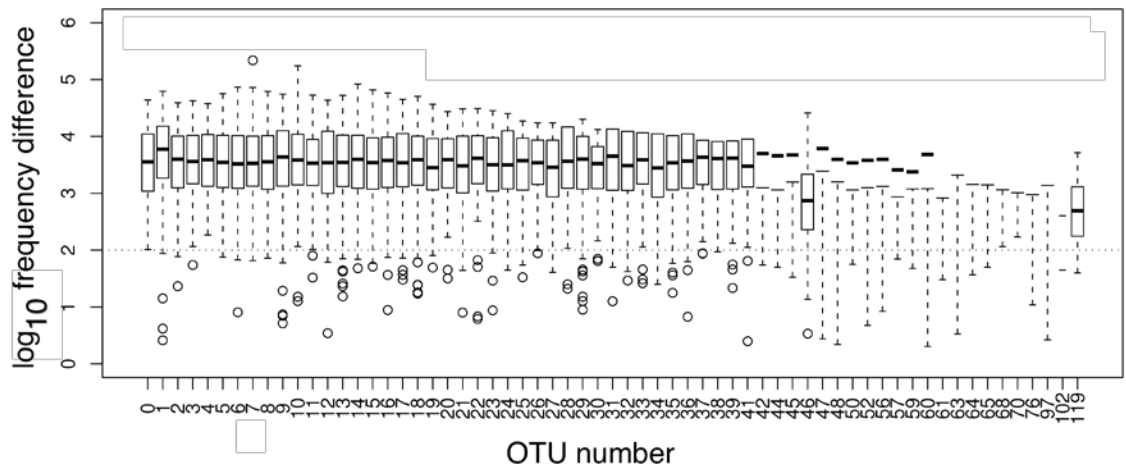
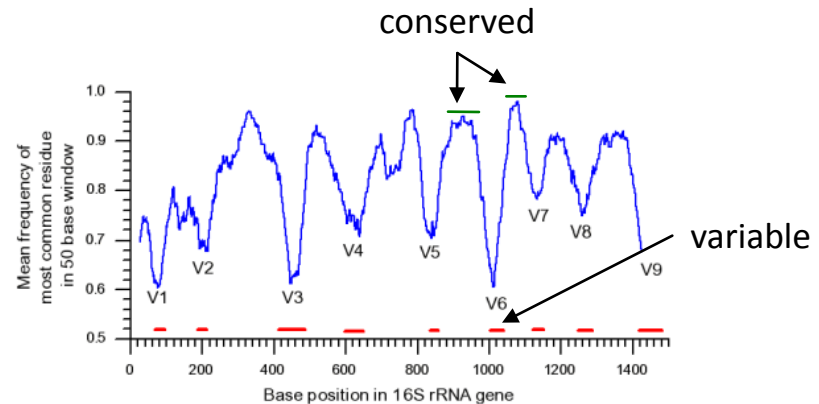


Assign taxonomy using BLAST:

- 1) Refseq bacteria (fully sequenced genomes)
- 2) Filtered RDP dataset
- 3) NCBI non-redundant excluding environmental samples

Associate OTU frequencies with clinical samples using barcode information

DATA WORKFLOW



Studies and samples

	Tanzania	Brazil	Toronto	London (post-menopausal)	London (pre-menopausal)
Cohort	HIV ⁺ with and without BV	<u>Study 1</u> With and without BV <u>Study 2</u> With and without VVC	Pregnant (~28 weeks)	With and without VVA (dryness)	With and without BV
Treatment	Oral metronidazole + Oral probiotic capsule or placebo	Oral antibiotic/antifungal + Oral probiotic or placebo	Progesterone (prevention of pre-term labour)	None	None
Sampling time points	<ul style="list-style-type: none"> • Day 0 • 2 weeks • 5 weeks • 15 weeks • 25 weeks 	<ul style="list-style-type: none"> • Day 0 (before treatment) • Day 28 (after treatment) 	<ul style="list-style-type: none"> • 28 weeks gestation • ?? 	<ul style="list-style-type: none"> • Day 0 • 2 weeks • 4 weeks • 6 weeks • 8 weeks • 10 weeks 	Single samples
Total no. samples	272	325	69	90	21

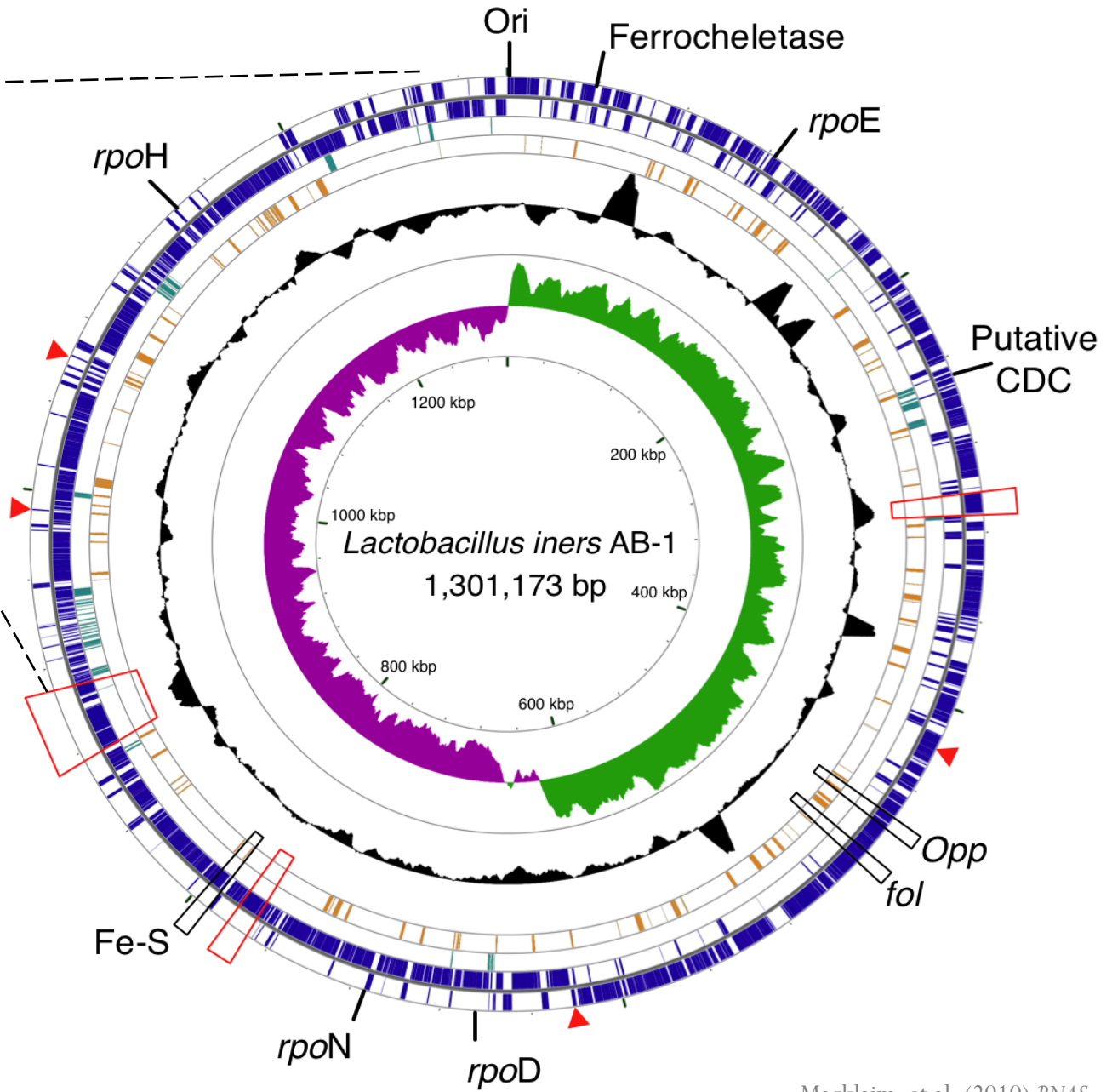
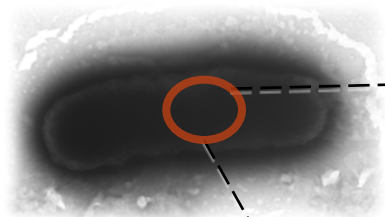
Total vaginal microbiota: 777

VVA – Vulvovaginal atrophy (dryness)

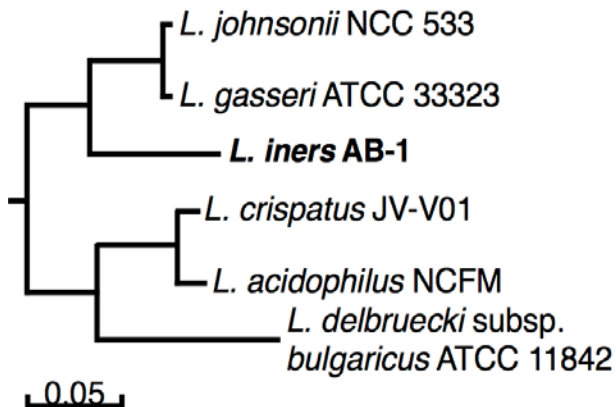
BV – Bacterial vaginosis

VVC – Vulvovaginal candidiasis (yeast)

Probiotic capsule: *Lactobacillus rhamnosus* GR-1 and *Lactobacillus reuteri* RC-14



Reduced genome



Size	CDS	HA(n)	HA(%)
1.99	1821	2	0.11
1.89	1755	2	0.11
1.28	1190	65	5.46
2.07	2209	10	0.45
1.99	1862	4	0.21
1.86	1562	24	1.54

- Smallest *Lactobacillus* genome sequenced to date (reduced across 13/20 functional categories)
- Highest proportion of genes acquired from foreign sources (via horizontal gene transfer)

