A circular phylogenetic tree diagram with multiple concentric rings. The outermost ring consists of colored vertical bars (red, green, blue) representing different taxa. The inner rings are composed of fine blue lines connecting the tips of the branches. The entire diagram is set against a background of radial gray lines.

SAGE 2016-17

2nd semester

Philipp Engel

Outline of today's session

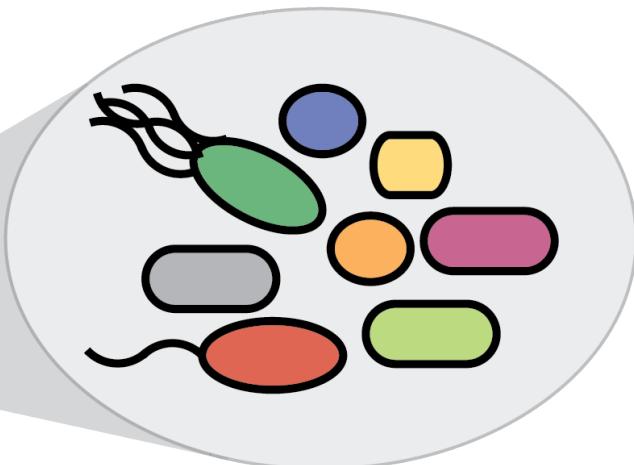
- Introduction to the 2nd semester
- Break
- Discussion of group/project assignment
- Short presentations by SAGE supplement groups
- Start your analysis project!

Gut bacteria of bumble bees!

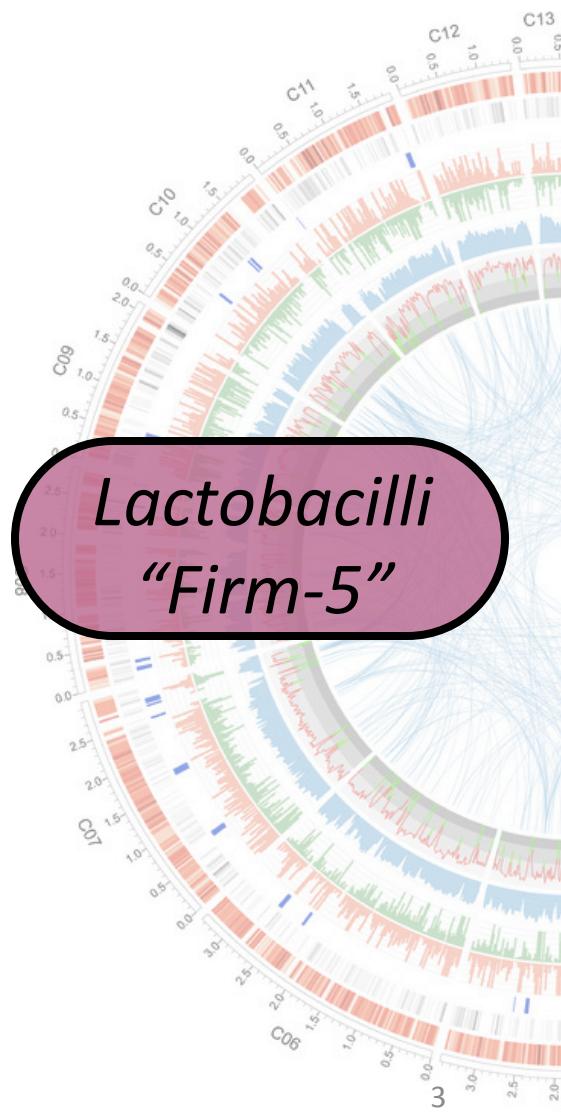
Bumble bee



Gut microbiota

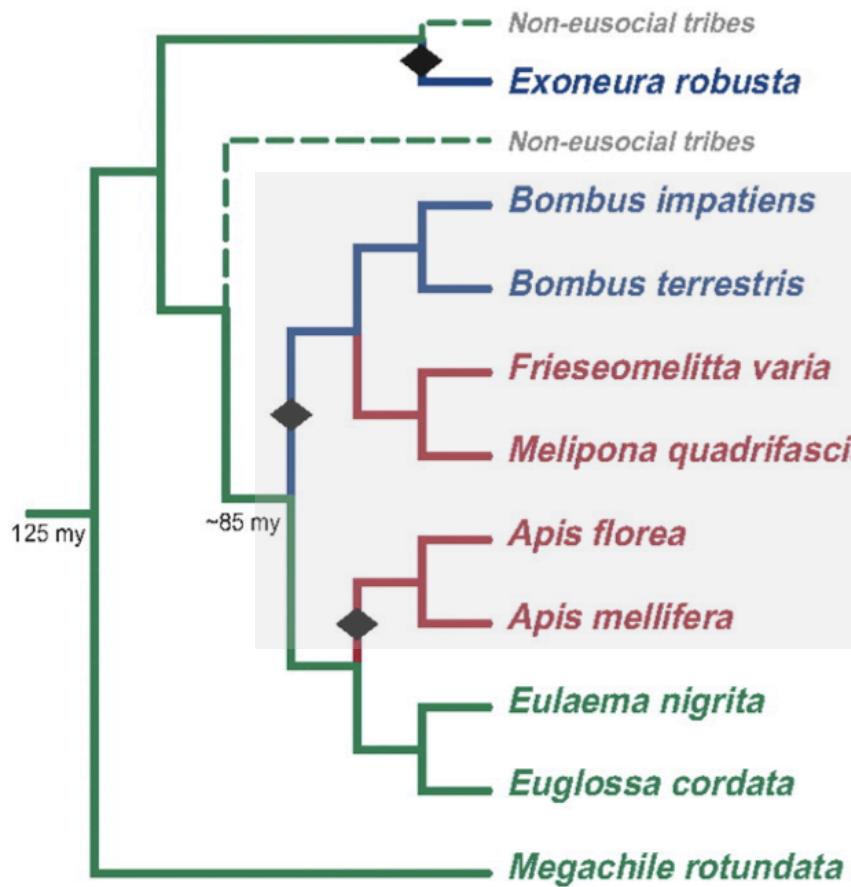


Lactobacilli
“Firm-5”



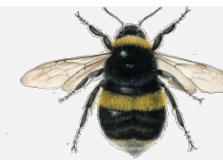
Bumble bees belong to the corbiculated bees

Bee phylogeny



Similar microbiota?

✗	Primitively Eusocial
✗	Highly Eusocial
(✓)	Highly Eusocial
(✓)	Highly Eusocial
✓	Highly Eusocial
✓	Highly Eusocial
✗	Non-Eusocial
✗	Non-Eusocial
✗	Non-Eusocial



Bumble bees
(*Bombus*)



Stingless bees
(*Meliponini*)



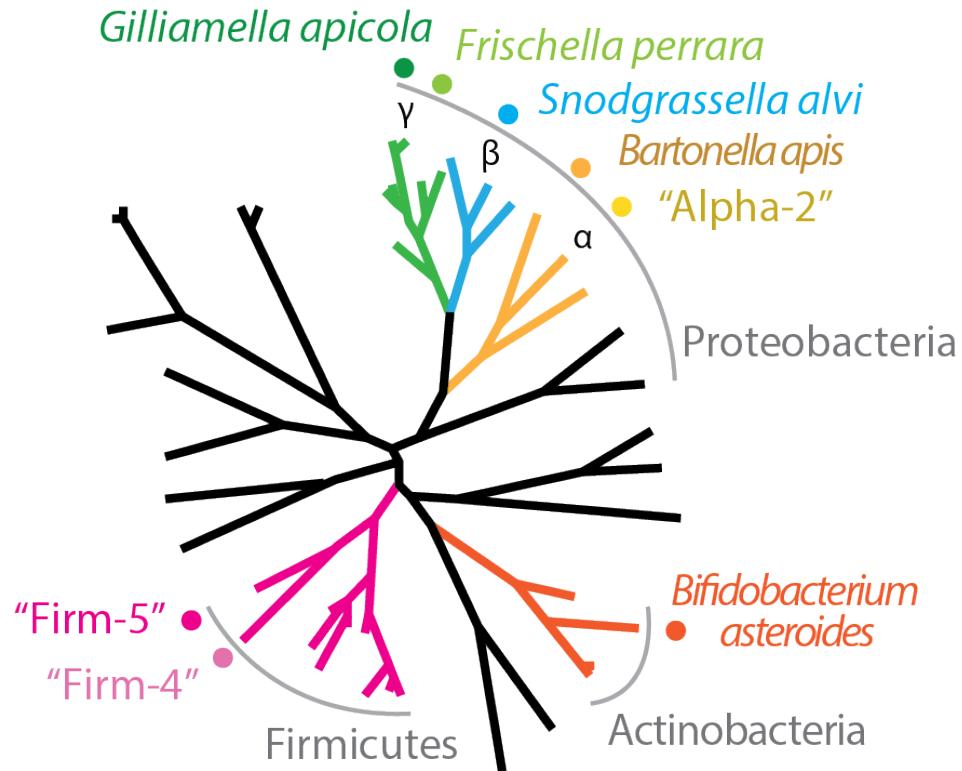
Honey bees
(*Apis*)

Important pollinators for agriculture and ecology

The bee gut microbiota

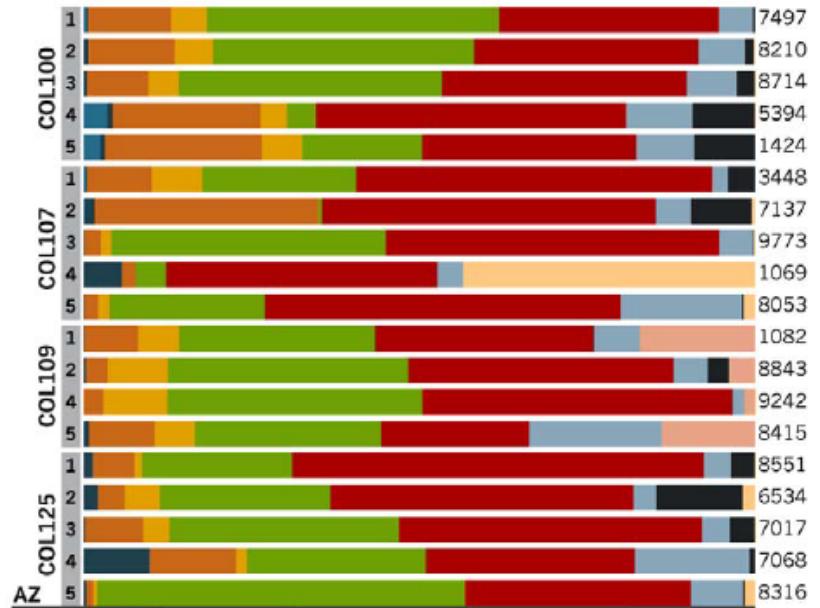


Honey bee (*Apis mellifera*): 8 species (<-> human: 100-500 species)



(adapted from Martinson et al, Mol Ecol 2011)

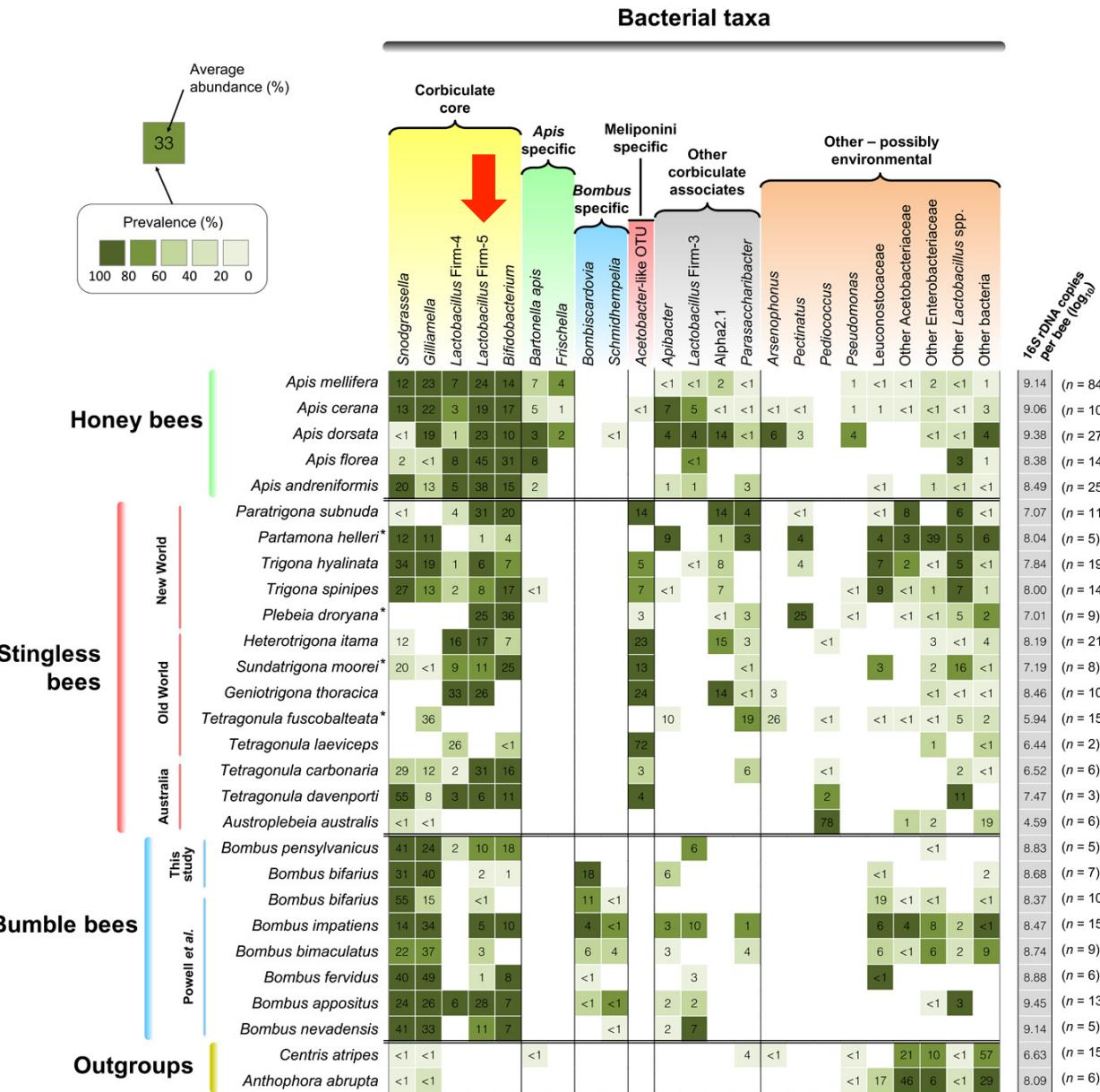
Each honey bee has the same microbiota...



16S rRNA community analysis
(OUT clusters \rightarrow 97% identity)

(Moran et al., PloS One 2012)

Firm-5 belongs to the core microbiota of social bees



Honey bees
(*Apis*)



Stingless bees
(*Meliponini*)



Bumble bees
(*Bombus*)

Insights from genome sequences?

Metabolic functions

Degradation of dietary compounds

Functions for host interaction

Functions for interbacterial interactions

Functions for pathogen interactions

Firm-5 from bumble bee

Phylogenies:
gut bacteria evolution

HGT of genes?

Genome structure

Evolution of gene families

Bacteriophages

Antibiotic resistance genes

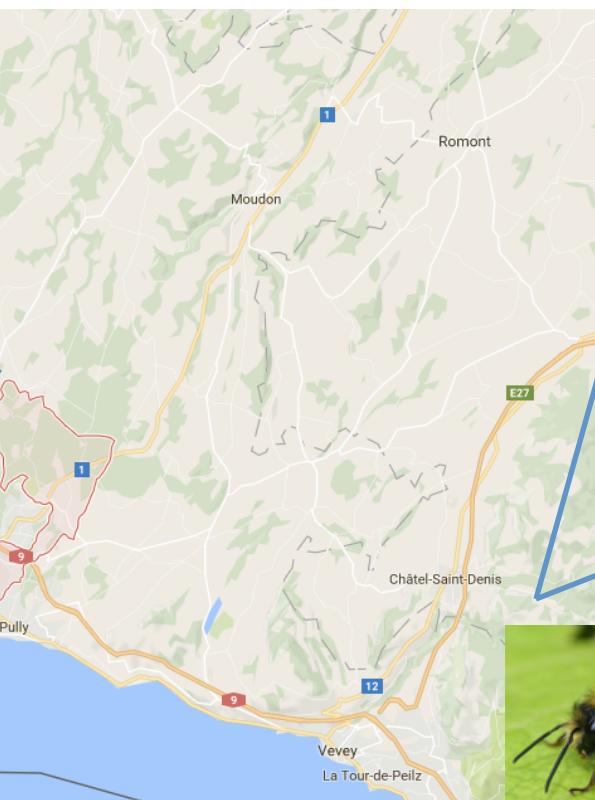
Genomic analysis → functions and evolution of gut bacteria

Sampling of bumble bees in CH

Bumble bee sampling sites:



UNIL-SORGE



B. terrestris



Moléson, FR



B. bohemicus



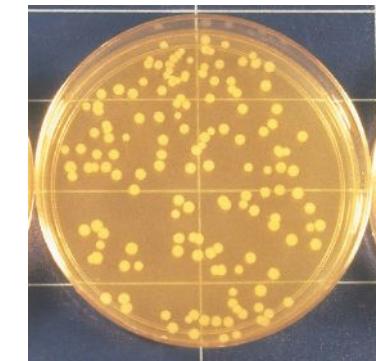
B. pascuorum

Isolation of bacteria from bee guts

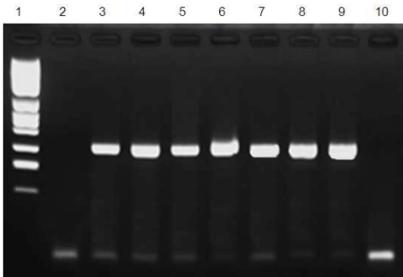
Gut dissection/homogenization



Plating on MRSA an anoxic conditions



DNA isolation & genotyping



15 isolates of Firm-5



two .fastq.gz
files per genome

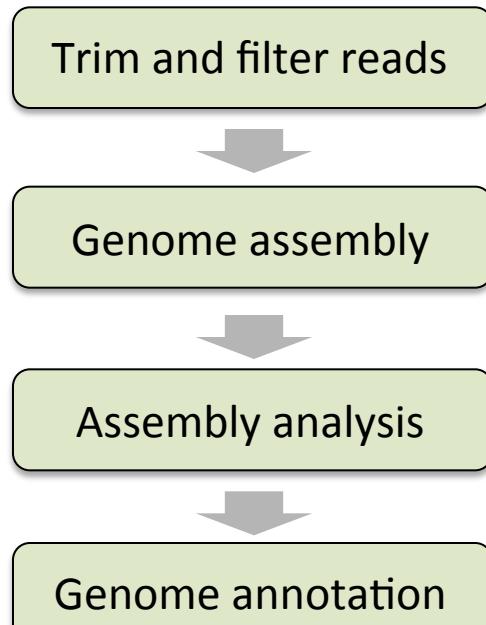
Illumina MiSeq
2x250bp reads

Assignment of genomes to groups

#	Isolate name	Bee species	Bee ID	Sampling	Site	Student group	Student names
1	Lb_233	<i>B. terrestris</i>	Bee#1	UNIL Dorigny	46°31'15.8"N 6°34'32.6"E	Group 1	Shaoline Sheppard Tatiana.Sokoloff
2	Lb_234	<i>B. pascuorum</i>	Bee#4	Moléson, FR	46°32'46.0"N 6°59'27.9"E	Group 2	Anthony Sorel Melvin.Berard
3	Lb_236	<i>B. pascuorum</i>	Bee#4	Moléson, FR	46°32'46.0"N 6°59'27.9"E	Group 3	Domique Jacques-Vuarambon Gaelle Spack
4	Lb_237	<i>B. pascuorum</i>	Bee#4	Moléson, FR	46°32'46.0"N 6°59'27.9"E	Group 4	Nathalie Thoman Titouan Laessle
5	Lb_230	<i>B. pascuorum</i>	Bee#5	Moléson, FR	46°32'20.1"N 6°59'22.8"E	Group 5	Claire Pralong Jennifer Mayor
6	Lb_225	<i>B. pascuorum</i>	Bee#6	Moléson, FR	46°32'15.6"N 6°59'14.2"E	Group 6	Sandrine Pinheiro Charlotte Griessen
7	Lb_245	<i>B. bohemicus</i>	Bee#7	Moléson, FR	46°32'20.1"N 6°59'22.8"E	Group 7	Mirjam Mattei Nathalie Guttmann
8	Lb_246	<i>B. bohemicus</i>	Bee#7	Moléson, FR	46°32'20.1"N 6°59'22.8"E	Group 8	Joaquim Claivaz Virginie Ricci
9	Lb_247	<i>B. bohemicus</i>	Bee#7	Moléson, FR	46°32'20.1"N 6°59'22.8"E	Group 9	Nastassia Gobet Ambrin Farizah Babu
10	Lb_228	<i>B. bohemicus</i>	Bee#8	Moléson, FR	46°32'20.1"N 6°59'22.8"E	Group 10	Line Aubert Yassine.ElChazli
11	Lb_259	<i>A. mellifera</i>	GC_Hbee#1	UNIL Dorigny	Colony "Grand Combin"	Group 11	Karim Hamidi Olivier Gustarini
12	Lb_260	<i>A. mellifera</i>	GC_Hbee#1	UNIL Dorigny	Colony "Grand Combin"	Group 12	Jonas Garessus Sarah Berger
13	Lb_261	<i>A. mellifera</i>	CB_Hbee#1	UNIL Dorigny	Colony "Cornettes de Bise"	Group 13	Laurent Casini Cyri Matthey-Doret
14	Lb_262	<i>A. mellifera</i>	GC_Hbee#1	UNIL Dorigny	Colony "Grand Combin"	Assistant	Kirsten Ellegaard
15	Lb_263	<i>A. mellifera</i>	GC_Hbee#1	UNIL Dorigny	Colony "Grand Combin"	Assistant	Kamil Jaron

Your job in SAGE2016-17

Autumn semester



Spring semester

Comparative genome analysis

Analysis project 1
(Group 1)

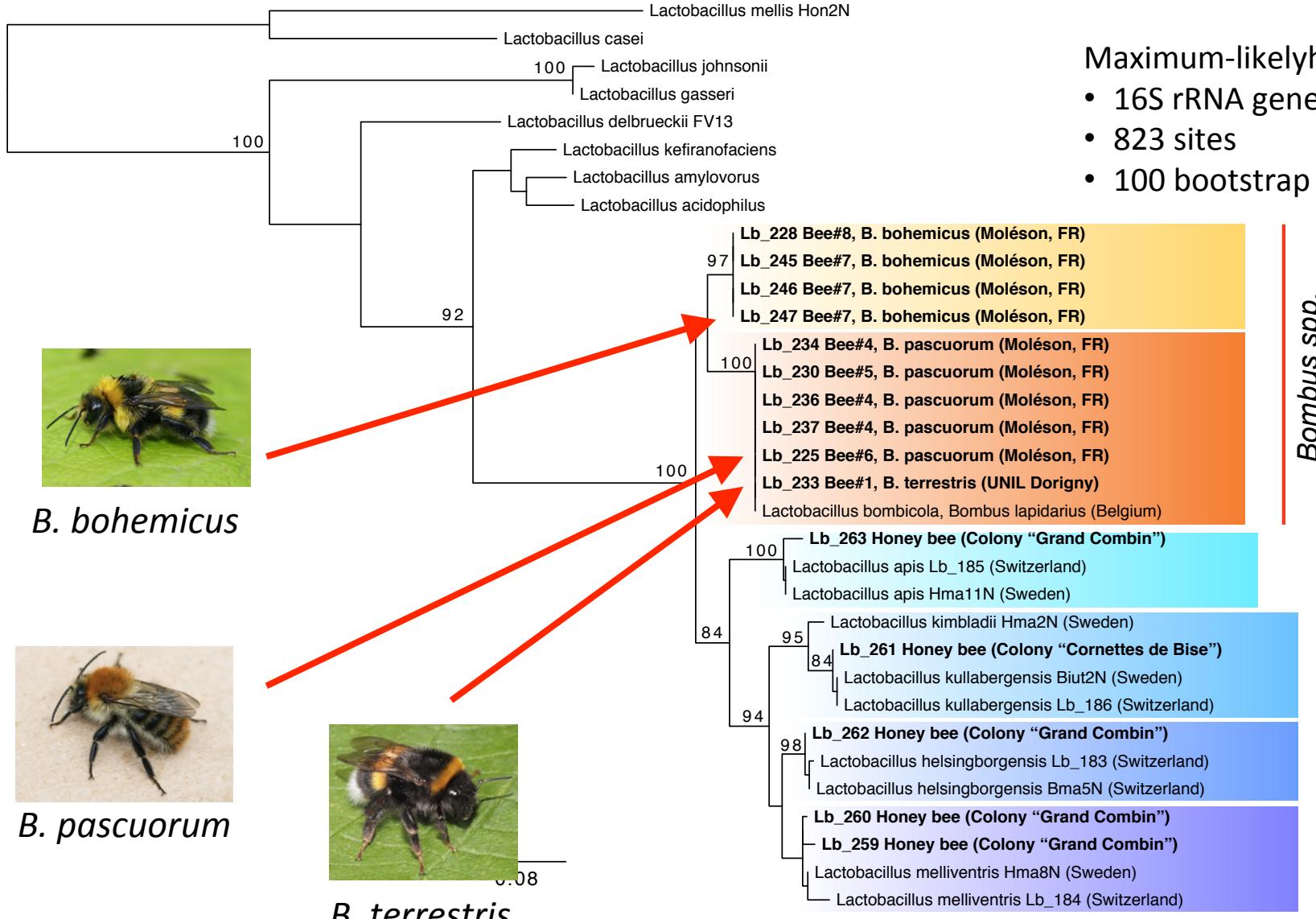
Analysis project 2
(Group 2)

Analysis project 3
(Group 3)

....

Analysis project 15
(Group 15)

Phylogenetic position of bumble bee isolates



Overview of Analysis projects

1. General genomic features and whole genome-alignments
2. Core Genome phylogeny
3. Firm5-specific gene content
4. Bumble bee and honey bee-Firm-5 specific gene content strains
5. Horizontal gene transfer between Firm5 and other bacteria
6. Horizontal gene transfer within the Firm5 group
7. Prophage islands
8. CRISPR mediated immunity
9. Carbon metabolism
10. Genome Scale Model

Analysis project “General genomic features”

Genome feature table

Strain	Size ^a (Mb)	Contigs ^a	GC % ^a	CDS ^a	rRNA loci ^a	tRNAs genes ^a	ANI ^b	dS ^{b,c}
<i>B. apis</i> PEB0122 ^T	2.60	16	45.5	2,174	2	46	-	-
<i>B. apis</i> PEB0149	2.53	7	45.5	2,113	2	46	98.3	0.03 (589)
<i>B. apis</i> PEB0150	2.58	12	45.6	2,196	1	43	97.5	0.05 (589)
<i>B. apis</i> BBC0178	2.60	1	45.3	2,228	2	47	85.8	0.77 (584)
<i>B. apis</i> BBC0244	2.64	1	45.2	2,264	2	46	85.8	0.77 (586)
<i>B. apis</i> BBC0122	2.91	1	45.7	2,396	2	47	85.1	0.82 (585)
<i>B. tamiae</i> Th239	2.26	10	38.0	1,983	2	44	69.8	1.61 (408)
<i>B. tamiae</i> Th307	2.20	10	38.0	1,977	2	44	69.8	1.62 (407)
<i>B. australis</i> NH1	1.58	1	41.8	1,265	2	42	68.4	1.24 (284)
<i>B. bacilliformis</i> KC583	1.43	1	38.2	1,322	2	44	68.6	1.48 (359)
<i>B. bovis</i> 91-4	1.62	1	37.3	1,379	2	43	68.5	1.58 (379)
<i>B. clarridgeiae</i> 73	1.52	1	35.7	1,326	2	59	68.6	1.48 (368)
<i>B. henselae</i> Houston-1	1.93	1	38.2	1,631	2	43	68.6	1.56 (374)

Overall similarity between analyzed genomes

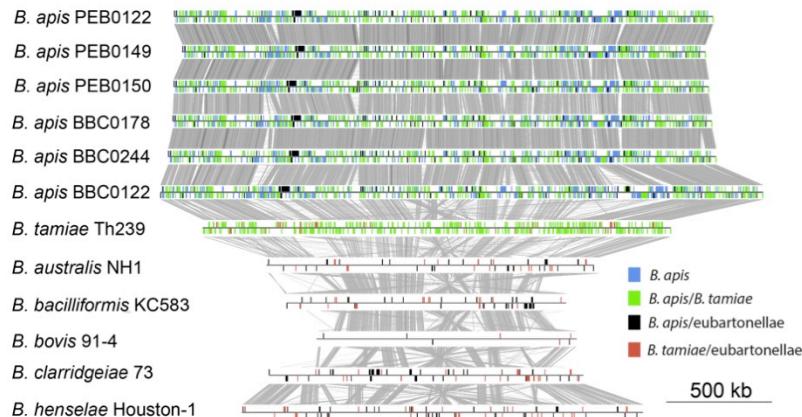
Quality check: all tRNAs and rRNAs present?

Genome rearrangements or static genome structure?

Large genomic islands → interesting functions!

Gene subsets can be mapped onto the aligned genome

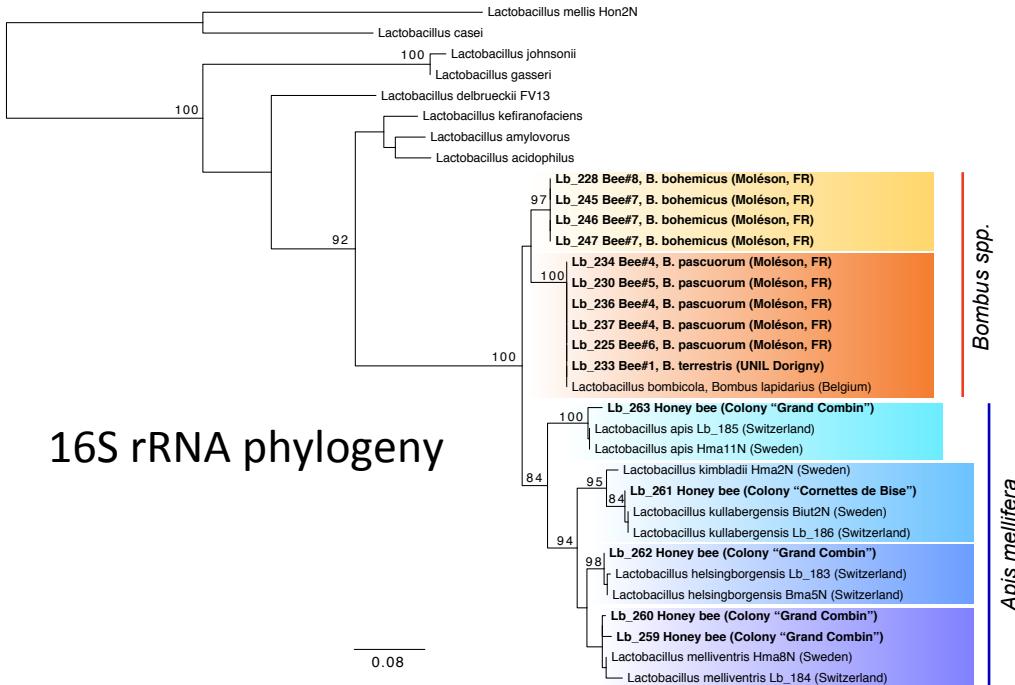
Whole genome alignments



Yassine El Chazli
Anthony Sonrel



Analysis project “Core genome phylogeny”



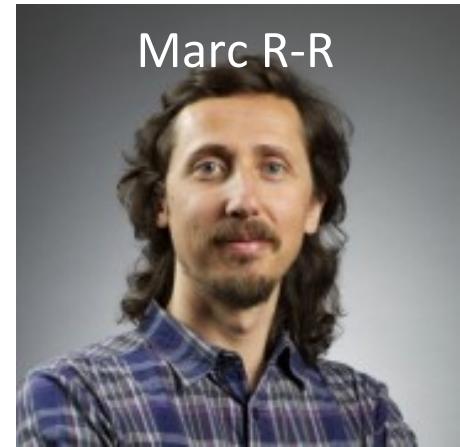
Approach:

1. Identify single-copy core genes (Homolog table)
2. Align protein and/or DNA sequences
3. Infer phylogenies on single genes and concatenate

Core genome phylogeny

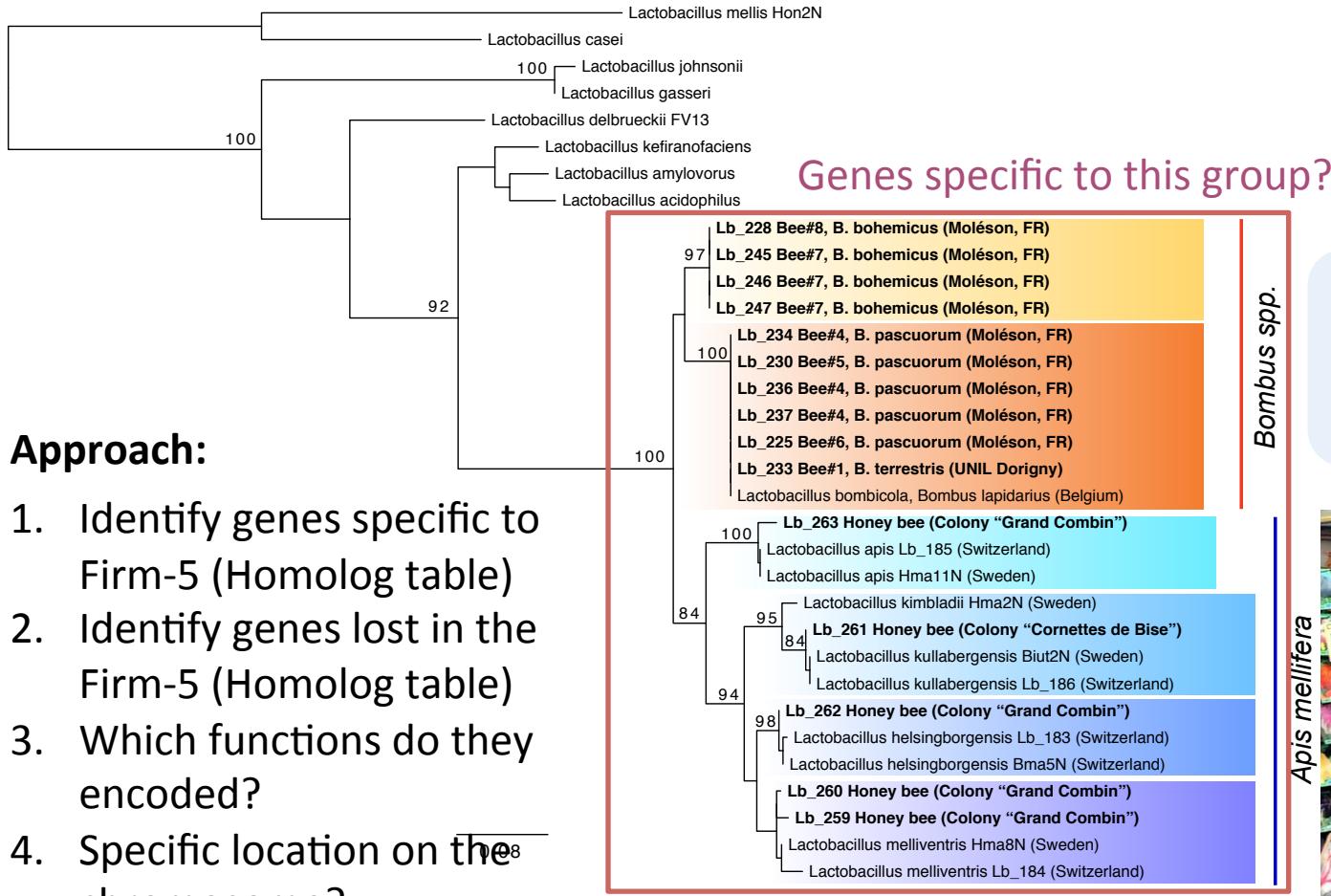
- Based on more than one genes
- Protein-coding genes
- ...
- More robust
- More resolved

Titouan Laessle &
Sandrine Pinheiro



Analysis project “Firm5-specific gene content”

Which gene functions are specific to Firm-5?



Sarah Berger &
Jonas Garessus

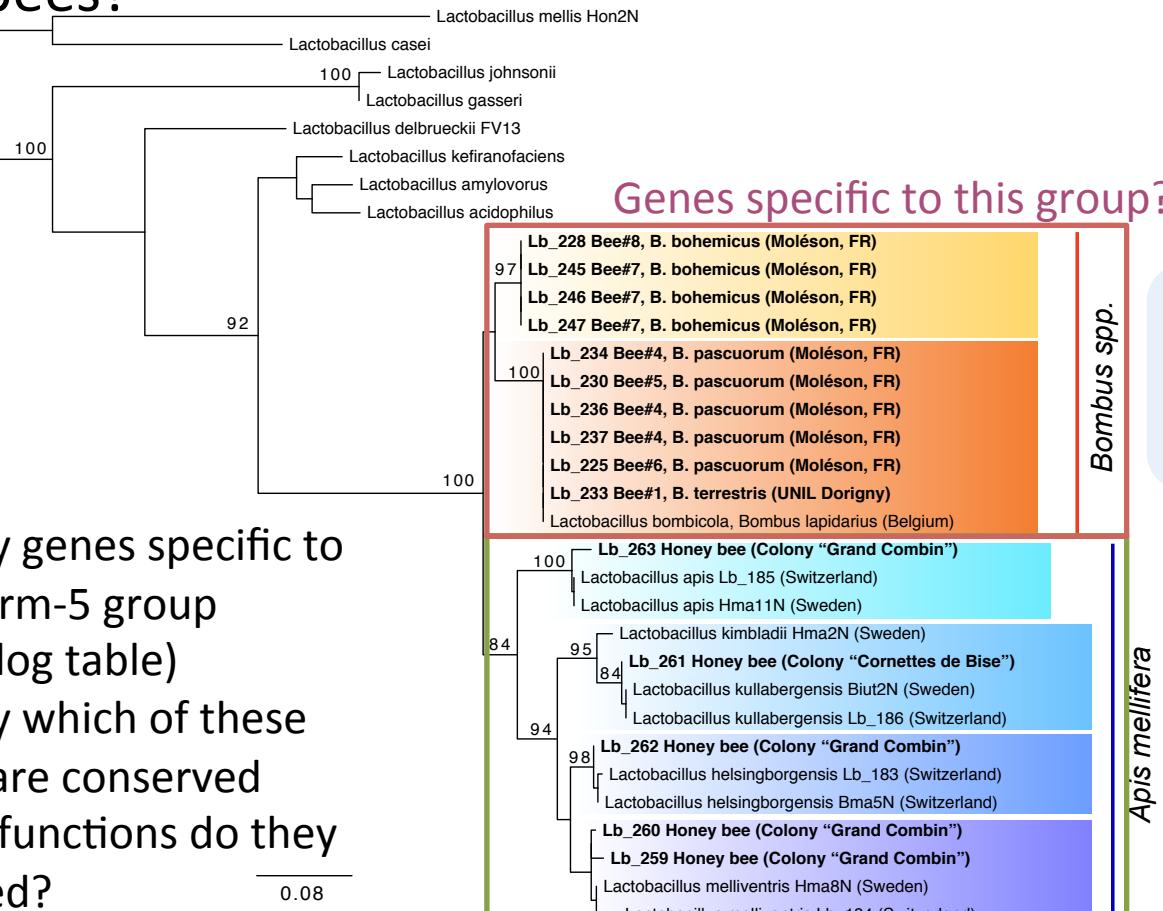
Approach:

1. Identify genes specific to Firm-5 (Homolog table)
2. Identify genes lost in the Firm-5 (Homolog table)
3. Which functions do they encode?
4. Specific location on the chromosome?



Analysis project “Host-specific gene content”

Which gene functions are specific to Firm-5 from bumble bees/
honey bees?



Approach:

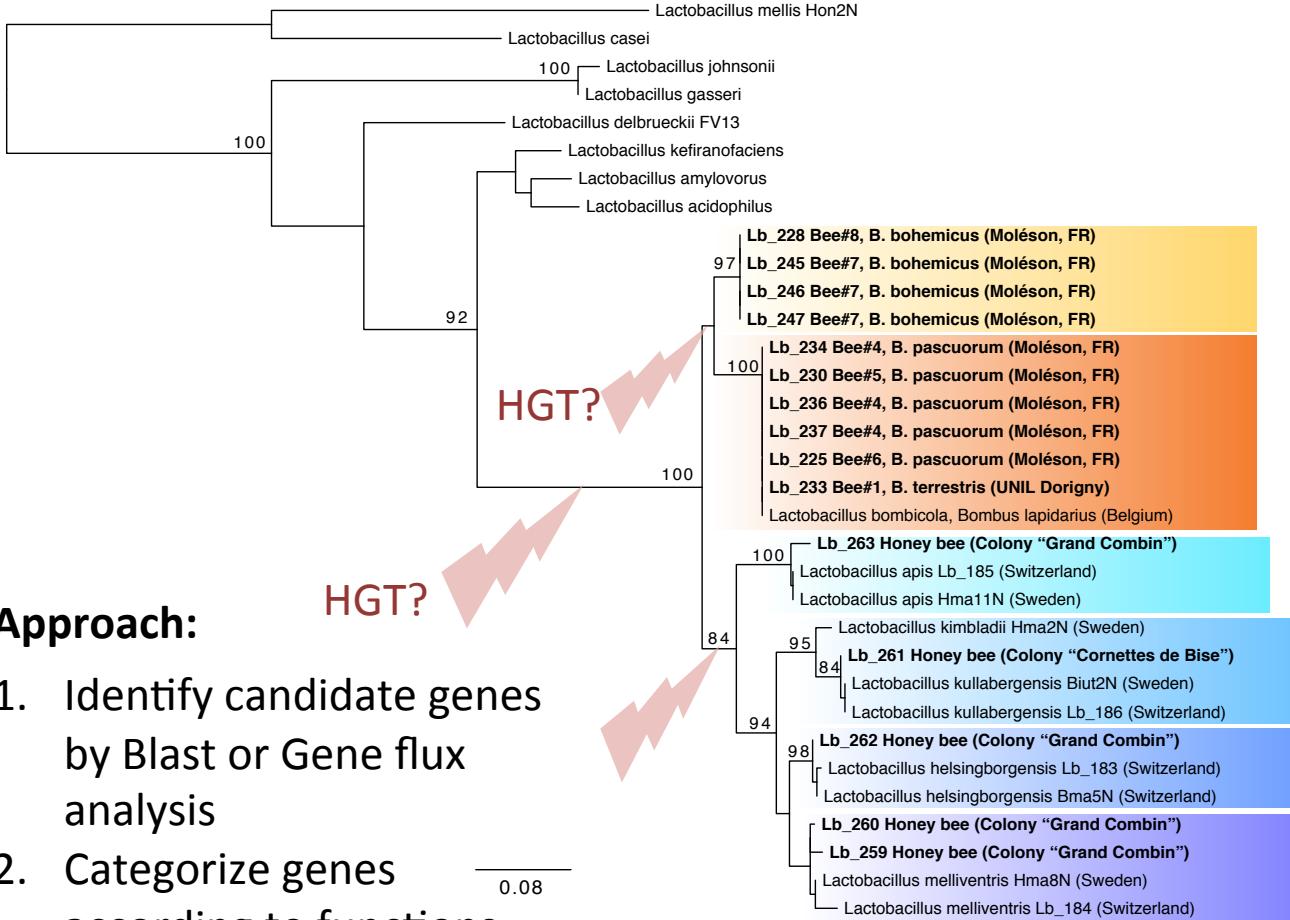
1. Identify genes specific to each Firm-5 group (Homolog table)
2. Identify which of these genes are conserved
3. Which functions do they encode?
4. Specific location on the chromosome?

Laurent Casini &
Cyril Matthey-Doret



Analysis project “Horizontal gene transfer 1”

Which genes have been acquired by HGT from other lineages?



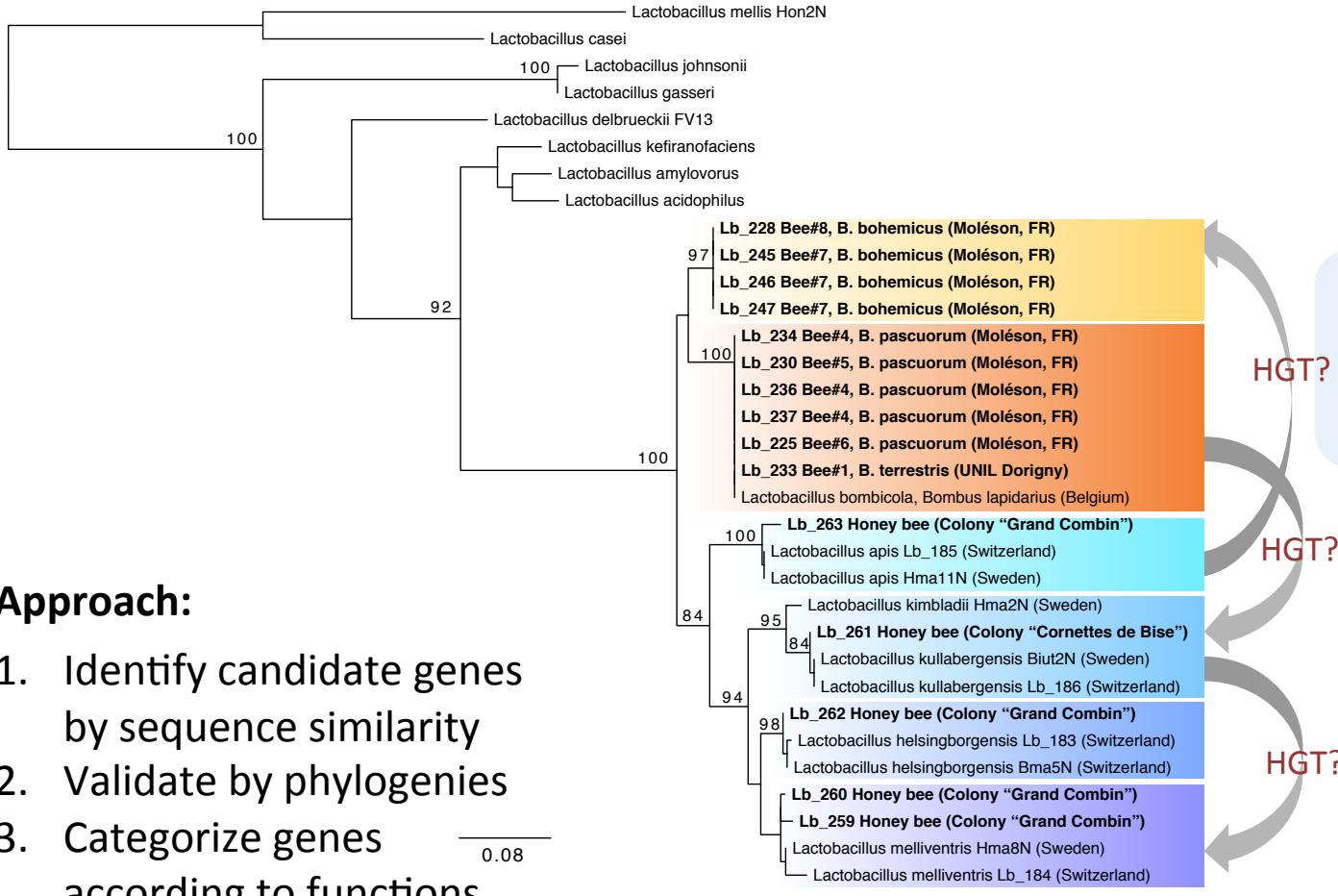
Joaquim Claivaz &
Virginie Ricci



Apis mellifera

Analysis project “Horizontal gene transfer 2”

Which genes have been exchanged among Firm5 strains via HGT?



Melvin Bérard & Charlotte Griessen



German B-R

Approach:

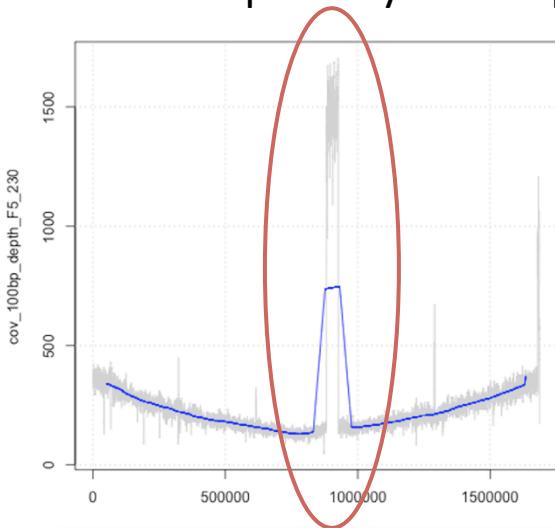
1. Identify candidate genes by sequence similarity
2. Validate by phylogenies
3. Categorize genes according to functions

0.08

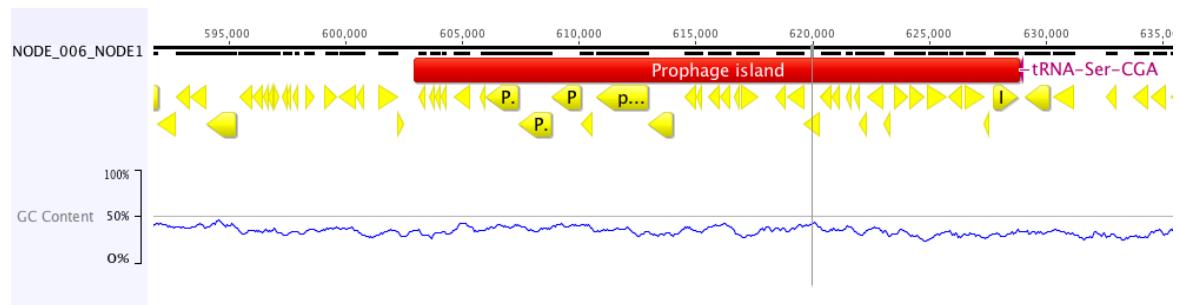
Analysis project “Prophages”

Characterize prophages hosted in the Firm-5 genomes

We have ‘possibly’ active phages!



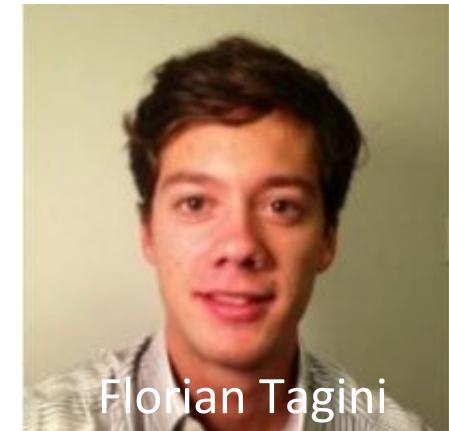
Quick check (using Geneious...)



Aim: Find all phage islands

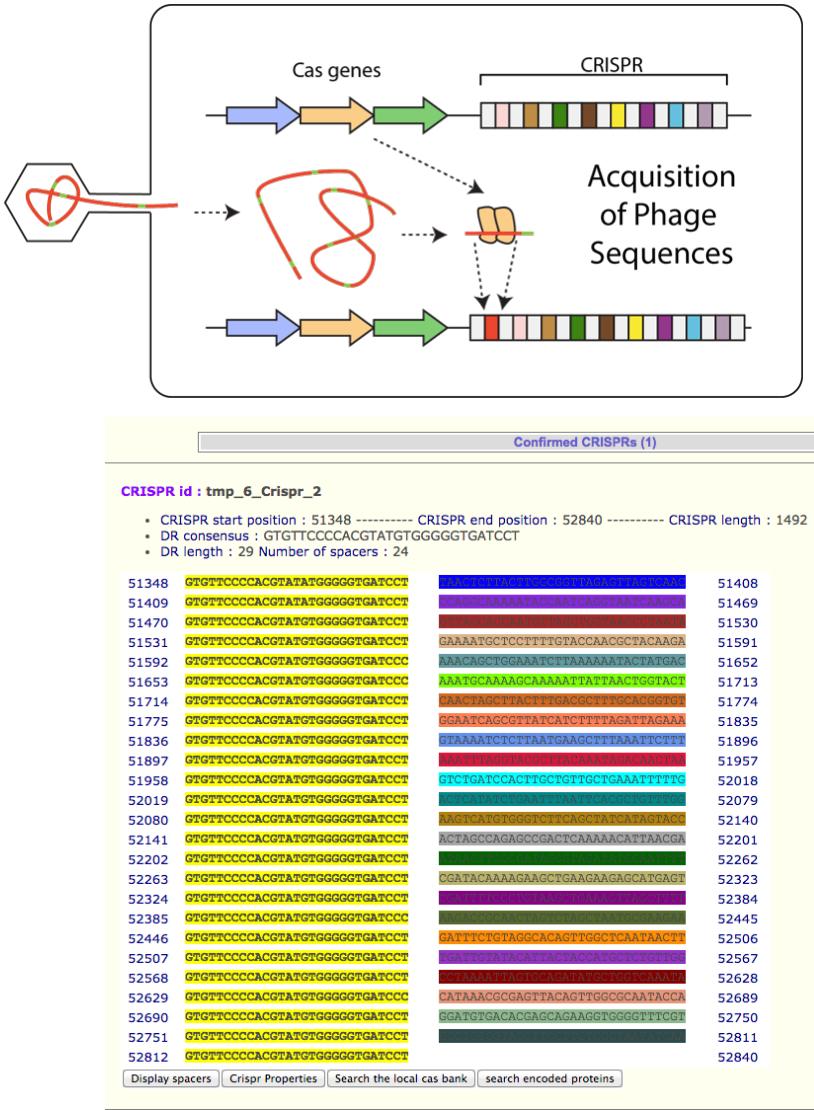
- use dedicated tools
- GC content, integration sites
- Genes encoding host functions?
- Are phages shared between genomes?
- Similar to phages found in other bacteria?

Jennifer Mayor &
Claire Pralong



Florian Tagini

Analysis project “Phage immunity”



Karim Hamidi &
Oliver Gustarini



Kirsten Ellegaard

Blast CRISPR sequences against:

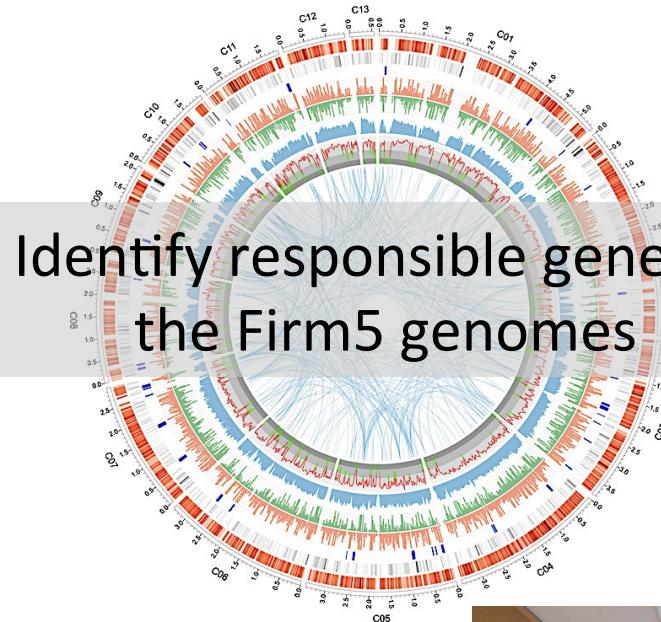
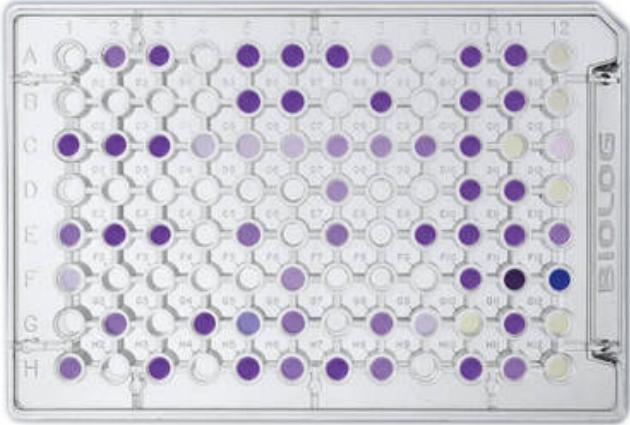
- nr DB of NCBI
- our Firm-5 genomes
- Metagenomes of social bees



Are certain strain
immune against phages/
mobile genes found in
other strains?

Analysis project “Carbon metabolism”

Biolog phenotyping



Polysaccharides



Glycoside hydrolases

Monosaccharides

Tool to identify glycoside hydrolases

Shaoline Sheppard &
Mirjam Mattei
Ambrin Farizah Babu



A web server and database for Carbohydrate-active enzyme ANnotation

dbCAN



Home

Browse

Genome

BLAST

Annotate your proteins

Download

Help

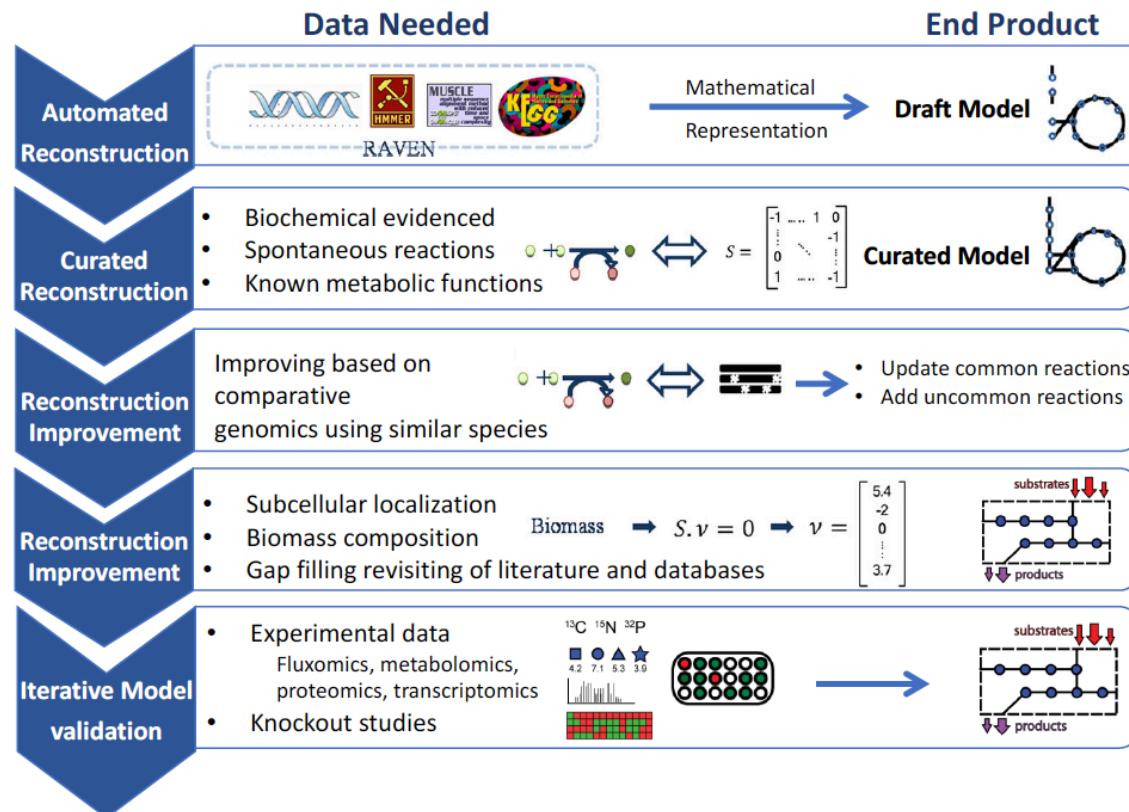
About us

Search: keyword(s), space (and), | (or)

Analysis project “Metabolic modeling”

Establish a model that can predict growth under given conditions

Genome Scale Model Reconstruction



Gaëlle Spack &
Nastassia Gobet &
Dominique Jacques-
Vuarambon



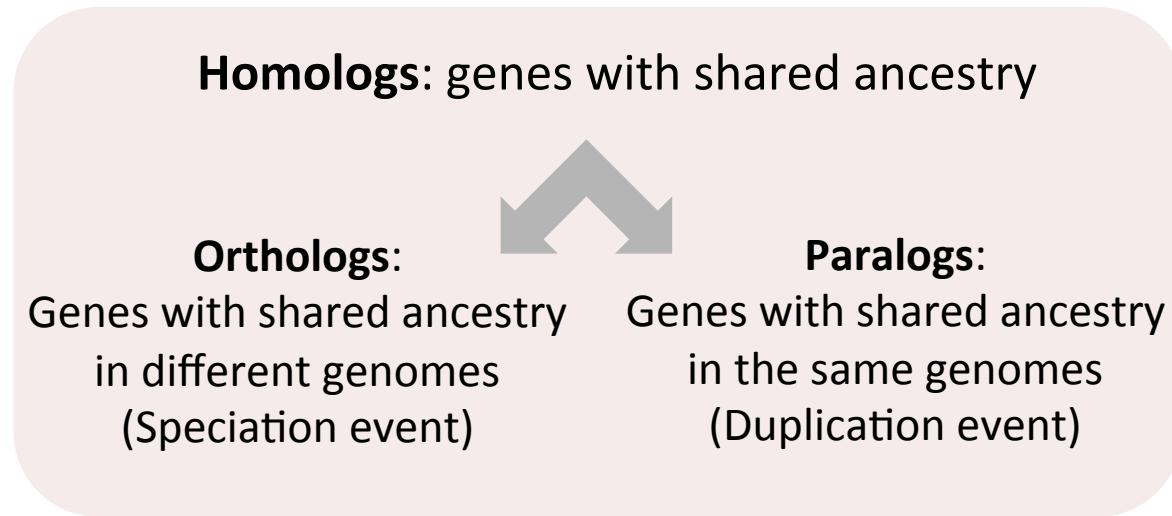
Noushin Hadadi

Available resources

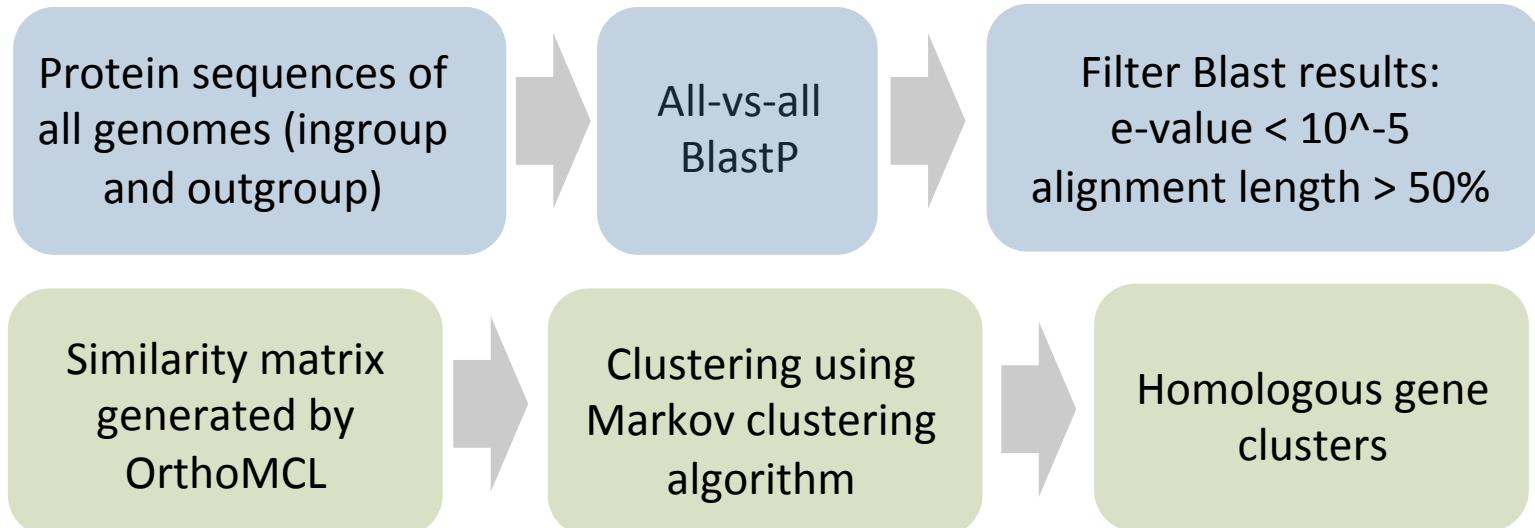
1. Genomes of all Firm-5 strains and outgroup species
 - Contigs in multiple fasta format
 - Genes in multiple fasta format:
 - DNA sequences
 - Protein sequences
2. RAST annotation (on RAST server or downloaded)
3. 16S rRNA gene phylogeny (as reference tree)
4. Homolog table
5. Google doc to share progress with others
6. Your personal assistant
7. WWW

Resources 1-4 are available on Vital-IT!

Homolog table – a list of all homologs



OrthoMCL: a tool to identify all homologs in genomes



Homolog table – a list of all homologs

The format of the homolog table:

- Each line = one gene family (= homolog cluster)
- A gene family can consist of orthologs and paralogs
- Genes which do not have a homolog are not listed
- Each gene is shown in the following format:
genome identifier|gene identifier
- Examples:

	homolog 1	homolog 2	homolog 3	homolog 4
Gene family 1	F225 1578.157.peg.1179	F228 1578.155.peg.1269	F230 1578.175.peg.1263	F233 1578.182.peg.1242
Gene family 2	F225 1578.157.peg.1180	F228 1578.155.peg.1270	F230 1578.175.peg.1264	F233 1578.182.peg.1243
Gene family 3	F225 1578.157.peg.1181	F228 1578.155.peg.1271	F230 1578.175.peg.1265	F233 1578.182.peg.1244
Gene family 4	F225 1578.157.peg.1192	F230 1578.175.peg.1277	F233 1578.182.peg.1259	F234 1578.177.peg.1117

	homolog 1	homolog 2	homolog 3
Gene family 5	WANG RS01060	WANG RS07780	WANG RS08760
Gene family 6	WANG RS01690	WANG RS04805	WANG RS08980
Gene family 7	WANG RS03325	WANG RS07535	JF74 05540
Gene family 8	WANG RS09400	WANG RS09420	LHV RS09840

File name: genefamilies_all.txt

Group 2-6 will probably need to use this homolog table, but others may be too...

Document your analysis and results!

- Take notes of everything!
- Report your progress in the google doc (link will be shared)
 - approaches and ideas
 - methods
 - scripts to filter/extract data
 - links to databases that were used
 - links to relevant literature
 - your results
 - your interpretation
 - troubleshooting
- Get help from your assistant!
- Be creative and flexible --> focus of a project may change

Evaluation of the course

Evaluation

- Final report about the analysis project
- Final presentation
- Participation during the course

Other activities beside group work

- Each class will start in plenum at 1.15pm, Amphipole 189
- Each week at the end of the course, one group will give a quick update about their project (around 4.30pm)

Questions?