Advancing the analysis of gut microbiome in critically ill patients using Hi-C technology

Alexander Tyakht

Institute of Gene Biology
Group of Bioinformatics
Moscow, Russia

Human gut microbiome in intensive care patients

In critically ill subjects, gut dysbiosis leads to inflammation, tissue injury – and back in a ”vicious circle”.

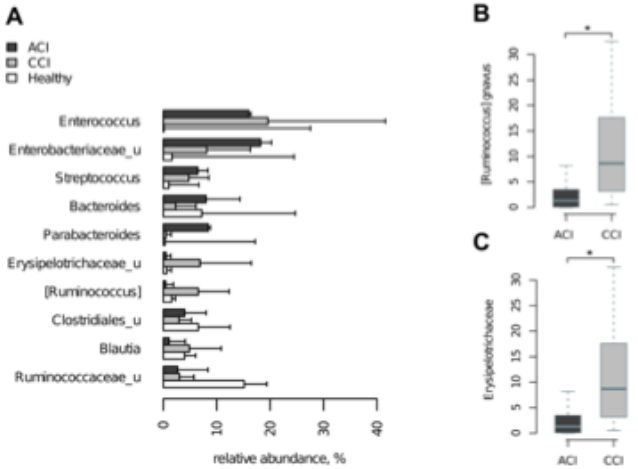
Microbiome - a reservoir for virulent multidrug-resistant nosocomial pathogens.

Chronically critically ill (CCI) patients:

- long stay in ICU
- regular treatment with multiple classes of antibiotics
- inflammation
- myopathy
- increased susceptibility to infection
- metabolic changes (hypercatabolism)
- hormonal changes
- cognitive impairment

Relevance to the COVID-19:

- ~7% of patients get bacterial co-infection.
- Some agents might come from the gut.
- Virulence potential of HGM should be explored.



Chernevskaya et al, Critical Care, 2020

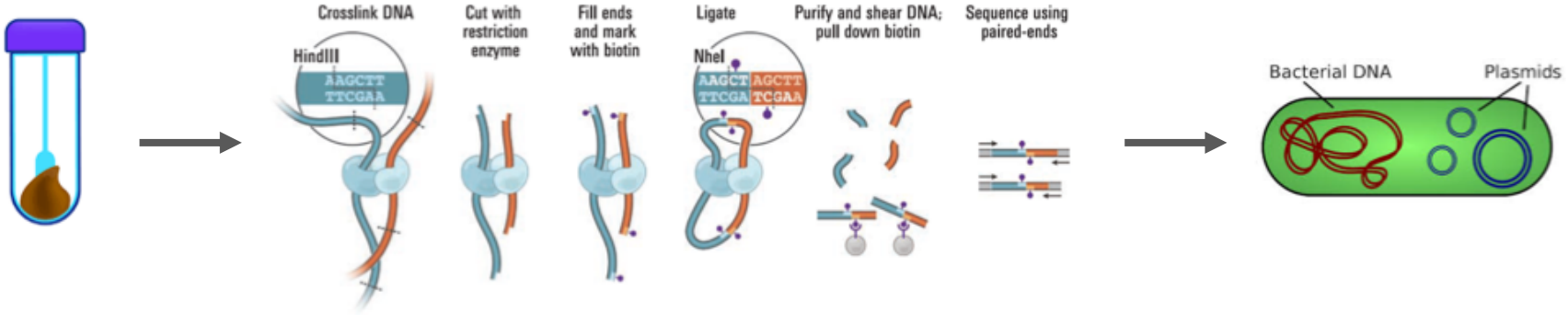
Hi-C metagenomics

The chromosome interaction signal is stronger:

- locally within a bacterial chromosome
- within a cell than between the cells.

Improved:

- genome assembly
- localization of mobile genetic elements



Lieberman-Aiden et al, 2009

Study design

Pilot set:
2 chronically critically ill patients x
2 time points

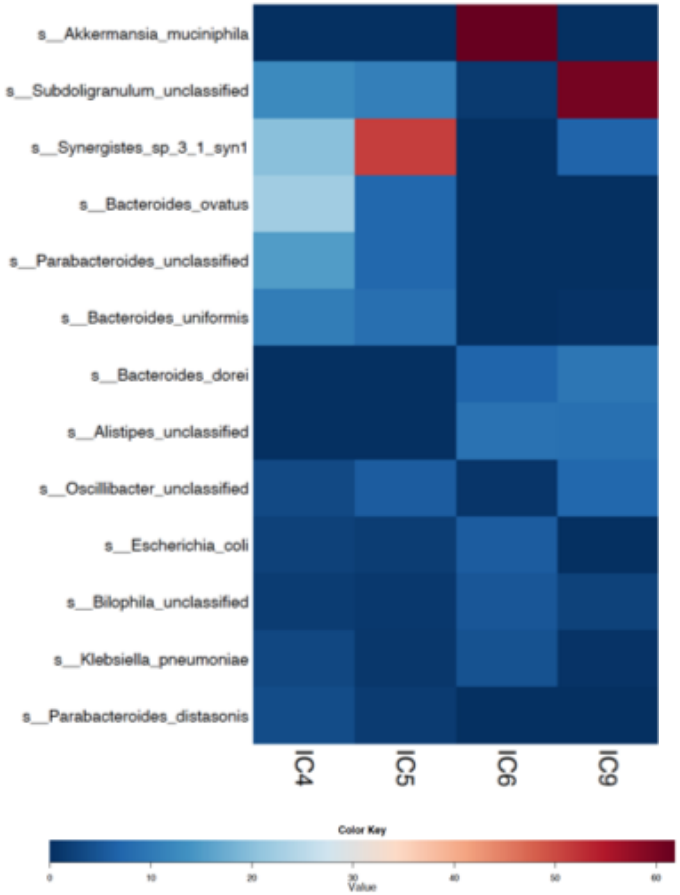
Sample type: stool.

Method: WGS + Hi-C sequencing (2 x 150 bp, Illumina).

	Patient A (samples 4 - 5)	**Patient B (samples 6 - 9)**
Age/Gender	75 / F	74 / M
Diagnosis	intracerebral hemorrhage	ischemic stroke
Length of stay in ICU, w	8	6
Support	Mechanical ventilation, enteral tube feeding (high calorie, low-residue)	
Timepoint 1	suspected bacterial infection pathogens in trachea (cultivation) pneumonia signs from CT Scan	
Timepoint 2	+1 week, negative clinical dynamics	+2 weeks, positive clinical dynamics

Basic analysis of gut community structure: pronounced dysbiosis

Timepoints: 4-5, 6-9.



WGS: decreased diversity driven by *Bacteroidaceae* and various opportunists (*Klebsiella*, *Escherichia*, *Proteus*, *Bilophila*).

MetaPhlan2

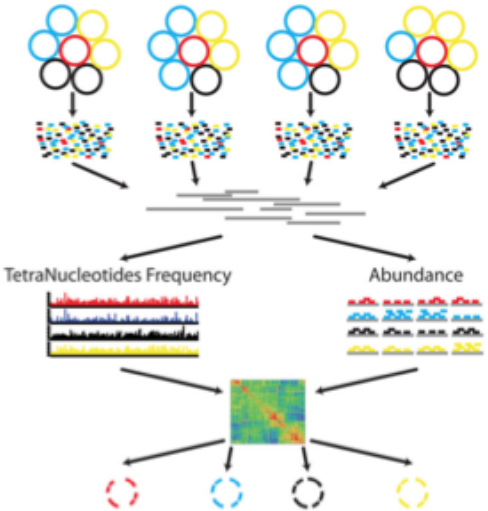
Metagenome-assembled genomes (MAGs)

From short NGS reads to microbes: contigs are binned into MAGs, basing on various principles.

Major MAG quality measures: # of MAGs/sample, completeness (%) and contamination (%).

WGS-MAGs

(MetaBat2 [Kang et al, 2015])

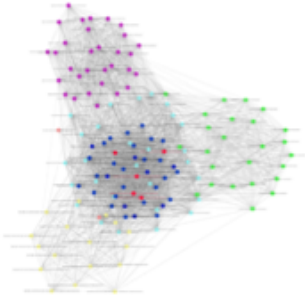


Hi-C-MAGs

(bin3c [DeMaere et al, 2019])

Graph: Hi-C contacts (edges) between WGS contigs (vertices).

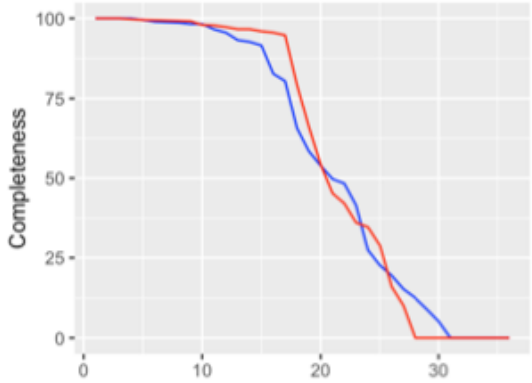
Identify the clusters.

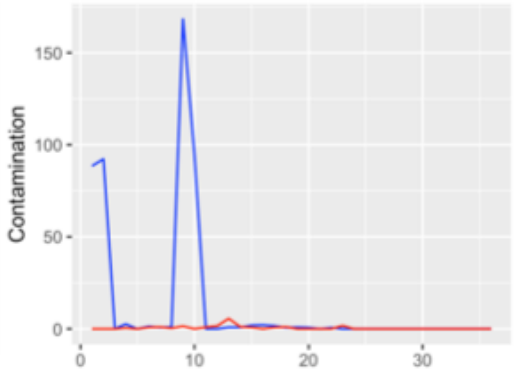


Hi-C allows to obtain higher quality MAGs compared to WGS

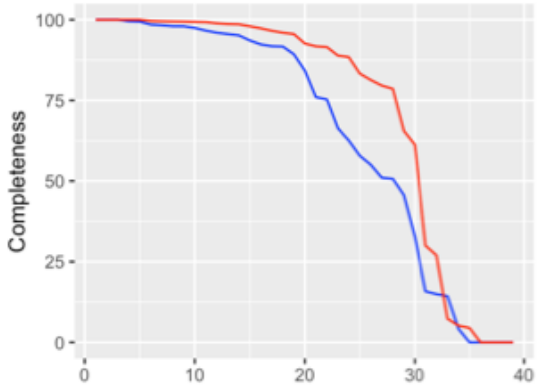
Hi-C
WGS

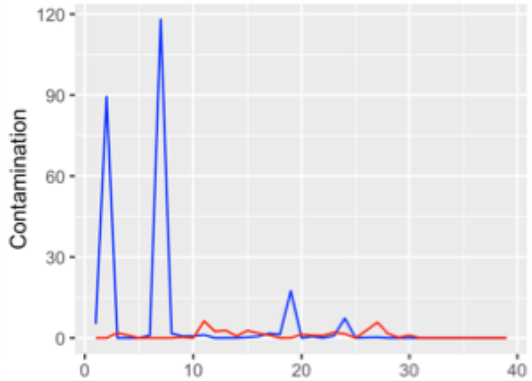
Sample 6





Sample 9





CheckM

~20 high-quality MAGs per sample

Number of high-quality MAGs (completeness > 80%) is similar.

Completeness is higher for Hi-C.

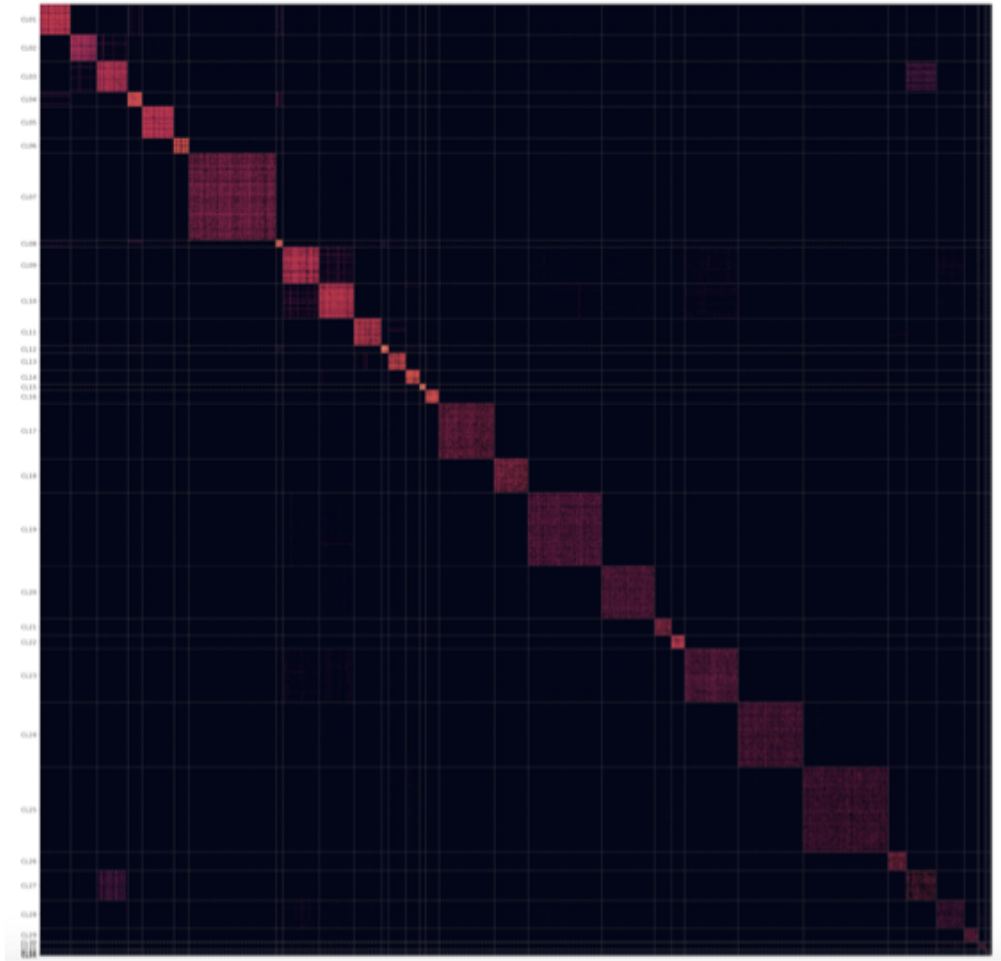
Contamination across high-quality MAGs is lower for Hi-C.

Microbiome-wide Hi-C chromosome contact maps

The Hi-C reads are mapped to WGS contigs.

For each MAG, there is a draft chromosome contact map.

How to finalize it?



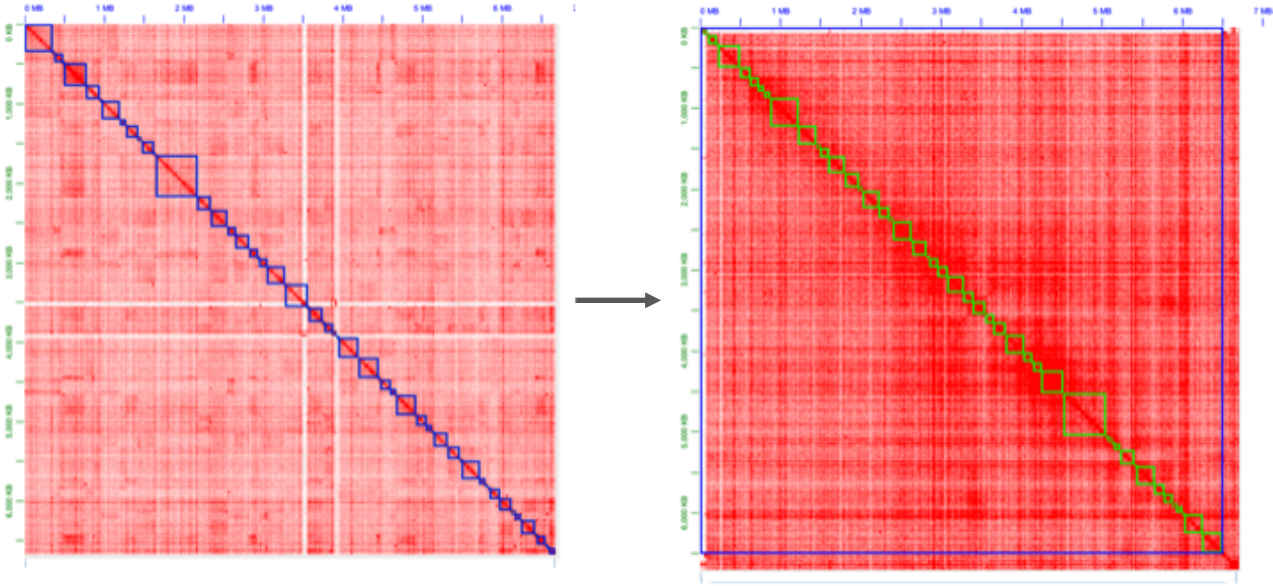
Hi-C map for all MAGs (sample 6)
(1 pixel = 1 contig)

Hi-C allows to obtain chromosome-scale scaffolds for abundant MAGs

Automatic re-ordering of the contigs based on maximization of diagonal signal.

Commonly used to improve scaffolding for single genomes of higher eukaryotes.

We applied it to microbiome-recovered maps to scaffold each MAG.



Bacteroides ovatus, sample 4
via 3D-DNA algorithm (Dudchenko et al, 2017).

Comparative genome analysis for major Hi-C MAGs of opportunist taxa

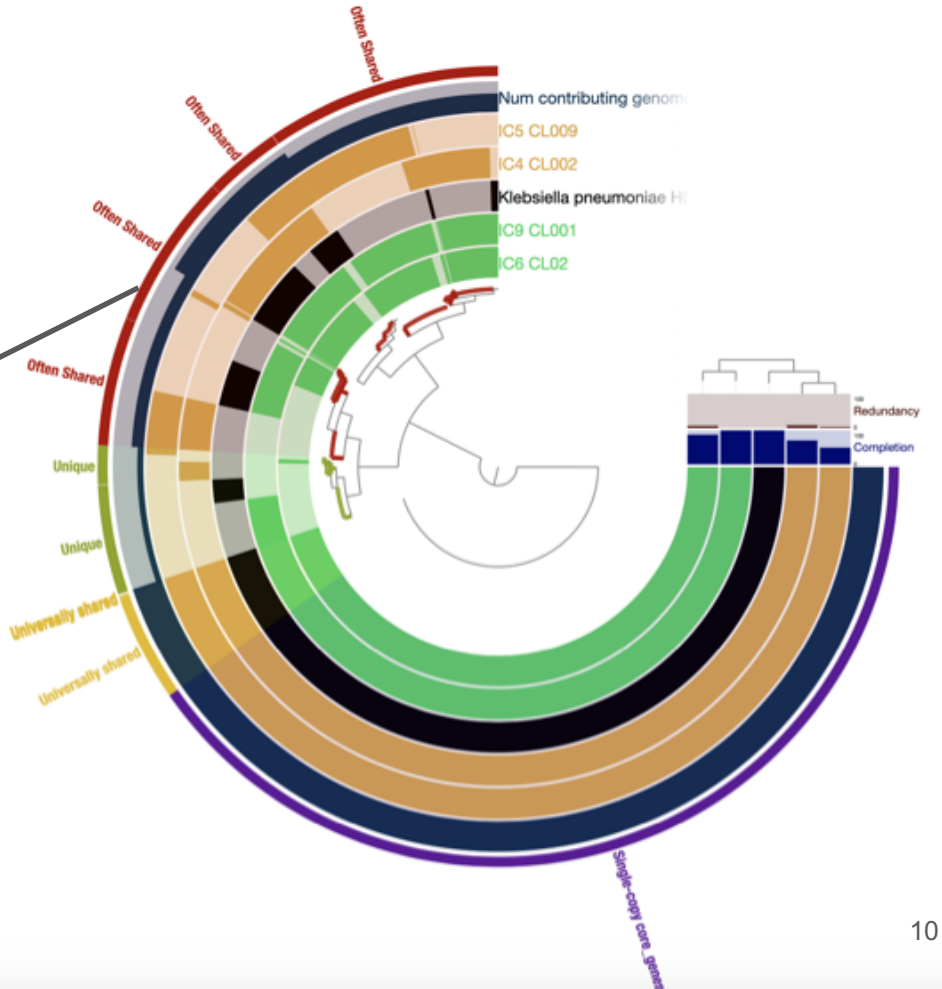
Klebsiella pneumonia (compared to external strain HS11286**)**

Within-patient genome similarity is higher than between the patients.

Patient B samples (6-9) are more similar than for A (4-5).

Interindividual differences in gene content are linked to virulence

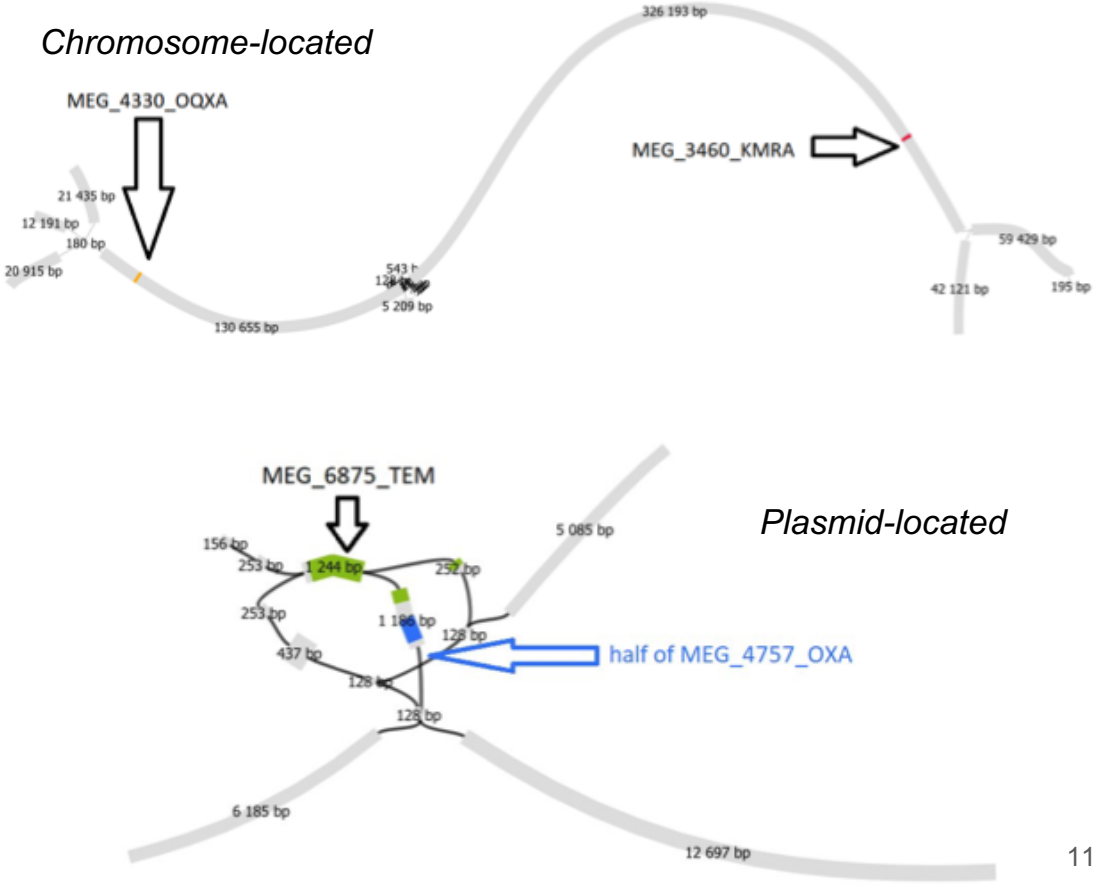
The genes unique to the patient B include the ones related to pili, fimbriae, phages and HGT.



Elucidating genomic context location of antibiotic resistance genes

Identify the environment of selected AR genes in the assembly graph.

Hi-C augmented version: in progress.



ESBL and other genes of K. pneumonia, sample 6 (using metaSPADES)

WGS shows extensive presence of plasmids

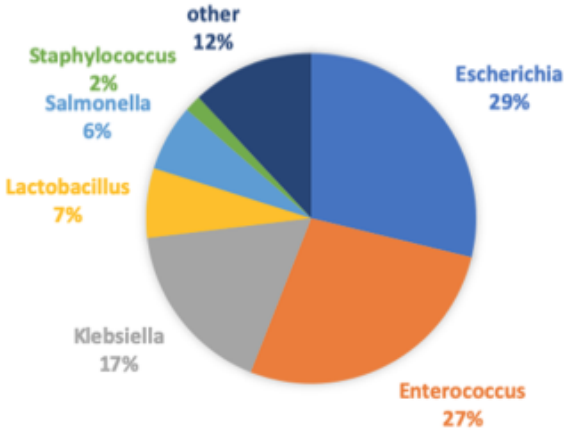
Plasmids are important HGT channel responsible for AR genes exchange.

Proteobacteria and *Enterococci* disproportionately contribute to the plasmid pool compared to the commensal taxa.

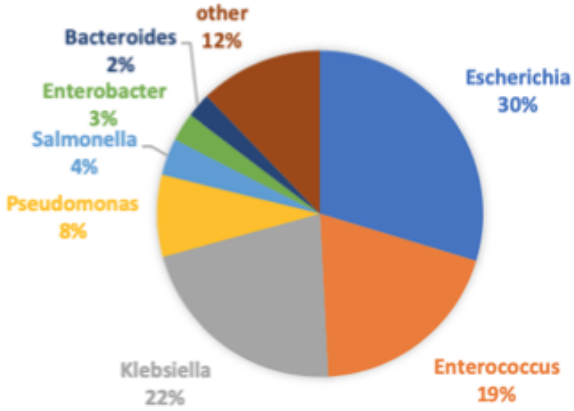
WGS allows to identify plasmid-like contigs, but not to link them to bacterial species.

Plasmids are underrepresented in WGS-MAGs.

#4



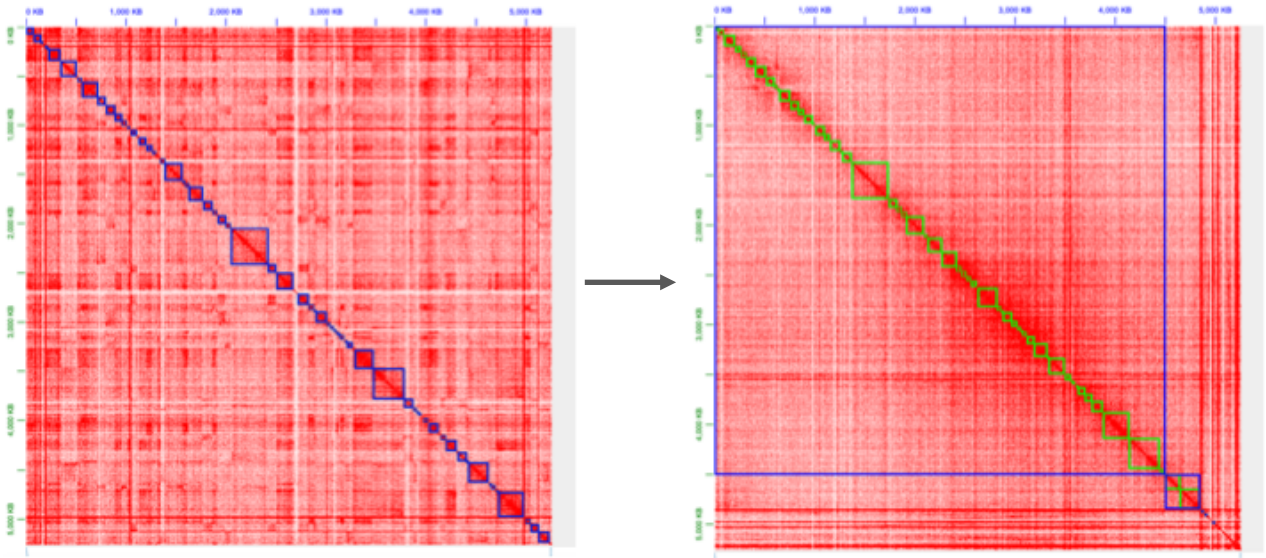
#5



PlasFlow (# of plasmid contigs per taxon)

Hi-C data help to elucidate “host-plasmid” associations

A plasmid is linked stronger to its hosts’ chromosome by the pairs of Hi-C reads than to ones of the other microorganisms.



E. coli, sample 6

The contigs forming solid vertical lines come from plasmids

Conclusions

- Hi-C metagenomics is a promising tool for analyzing clinical microbiome samples.
- Improved Hi-C MAGs: higher completeness, lower contamination, chromosome-scale scaffolds.
- For critical care, the method allows to get insights into virulence potential and antibiotic resistance, as well as tracking mobile genetic elements dynamics.
- The findings can help optimize the treatment schemes and understand mechanisms of pathogenesis.
- The pilot pipeline is established for analyzing involvement of gut microbiome in COVID-19 pathogenesis using Hi-C metagenomics.

Thank you for your attention!

Valeriia Ivanova
Petr Vasiliev
Sergey Ulianov
Institute of Gene Biology, Moscow

Ekaterina Chernevskaya
Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, Moscow

Artem Ivanov
Vladimir Ulyantsev
ITMO University, Saint-Petersburg

Funded by
RSF Russian Science Foundation
19-74-10092

Contacts: a.tyakht@gmail.com