

Bioinformatics approaches of Metagenomes

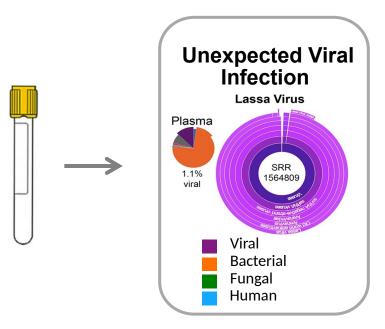
Jiabin Huang

Bioinformatics Meeting 09.09.2019

Metagenomics

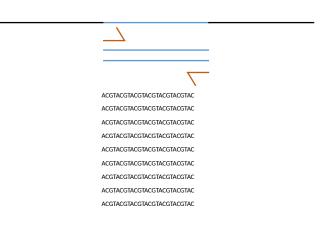


- Unbiased shotgun metagenomics
 - Fragment DNA and sequence randomly



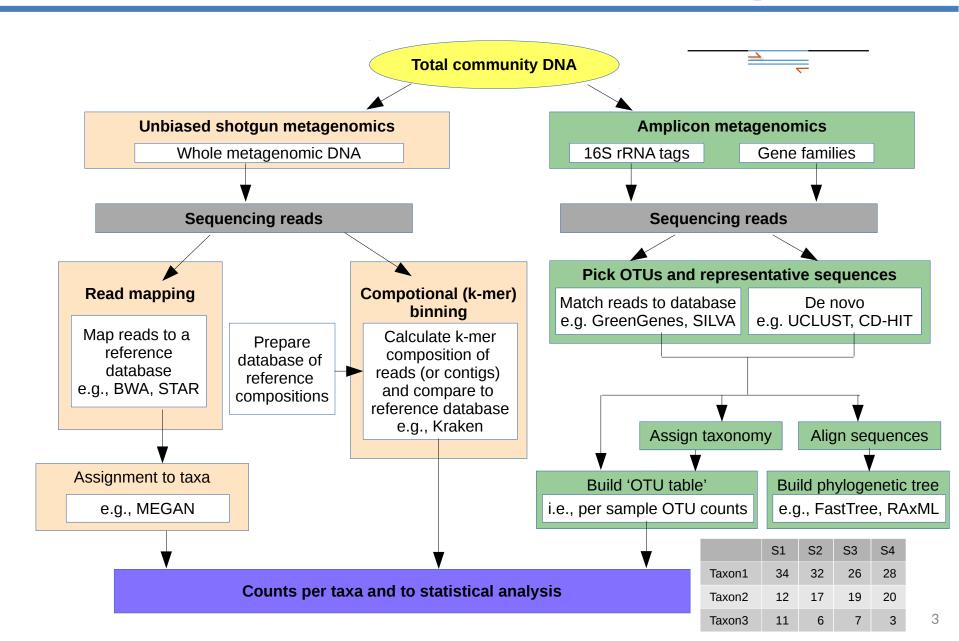
Data from Science, 345, 6202:1369-1372

- Amplicon metagenomics
 - PCR amplify a gene of interest
 - Tells you what types of organisms there are
 - Bacteria/Archaea (16S rRNA),
 Microbial Euks (18S rRNA), Fungi
 (ITS), Virus (no good marker)



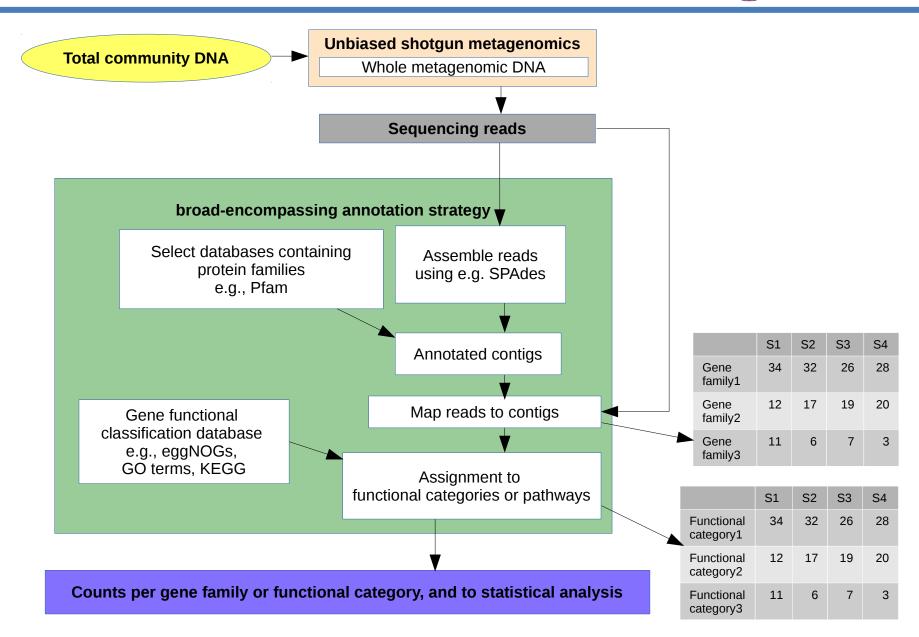
Taxonomic analysis of metagenomes





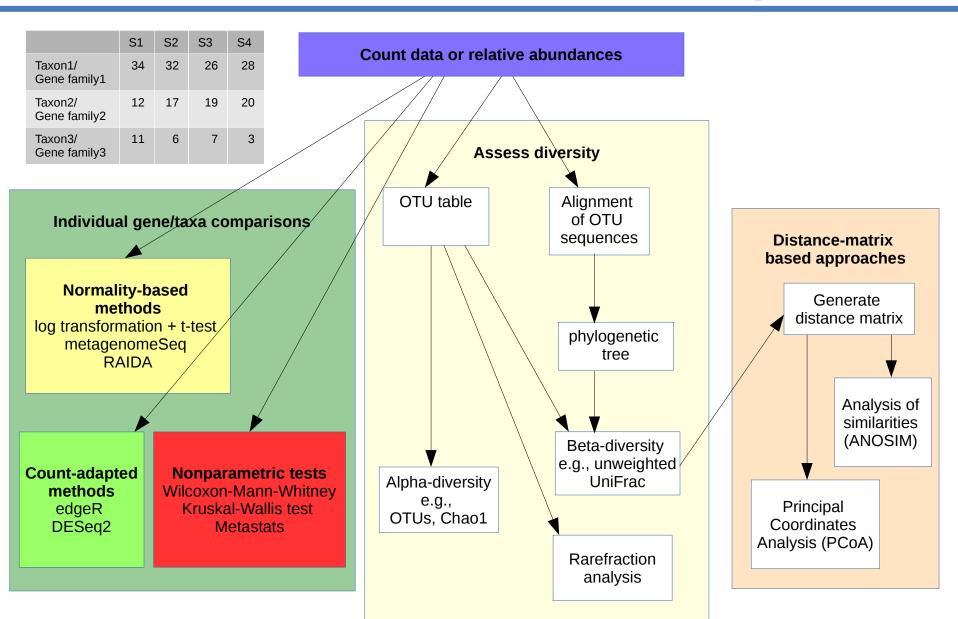
Functional analysis of metagenomes





Statistical analysis of metagenomes





1st case study: 16S amplicon metagenomics

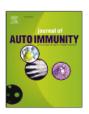




Contents lists available at ScienceDirect

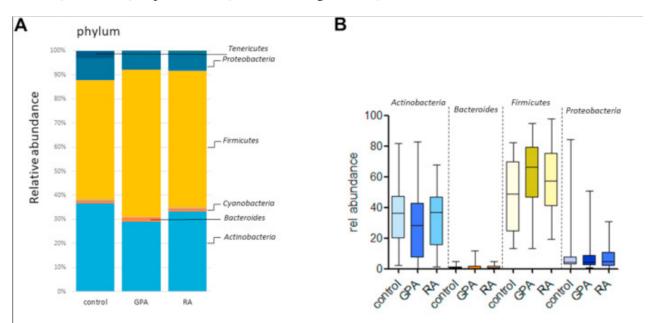
Journal of Autoimmunity

journal homepage: www.elsevier.com/locate/jautimm



Changes in the composition of the upper respiratory tract microbial community in granulomatosis with polyangiitis

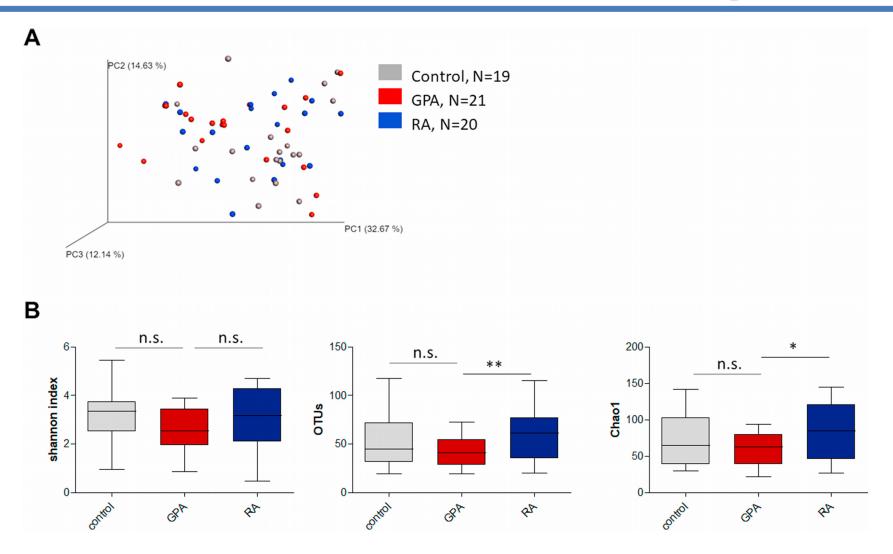
Peter Lamprecht^{a,*,1}, Nicole Fischer^{b,c,1}, Jiabin Huang^b, Lia Burkhardt^d, Marc Lütgehetmann^b, Fabian Arndt^e, Ida Rolfs^f, Anja Kerstein^a, Christof Iking-Konert^g, Martin Laudien^f



Bacterial community composition of the microbial upper respiratory tract in patients with granulomatosis with polyangitis (GPA) and rheumatoid arthritis (RA) and healthy controls.

1st case study: 16S amplicon metagenomics



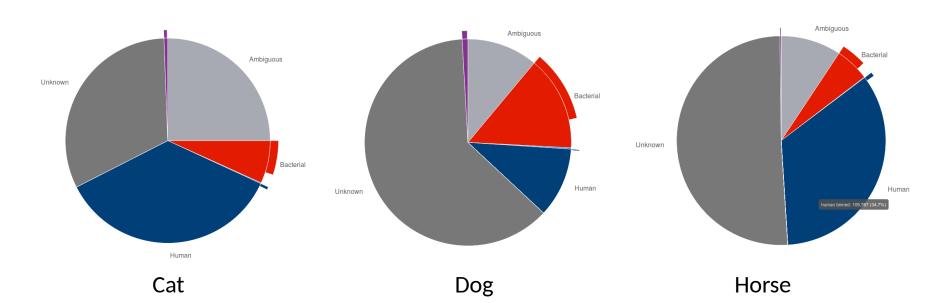


(A) PCoA plot of weighted_unifrac (B) Alpha diversities

2nd case study: unbiased shotgun metagenomics

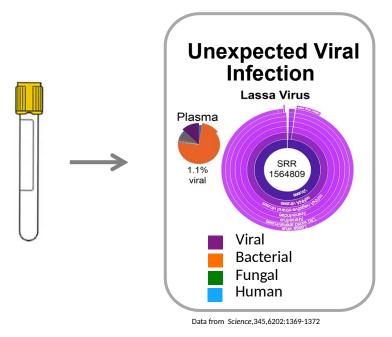


Samples	Human	Bacterial	Viral	Fungal	Ambiguous	Unknown
Cat	35.6%	6.7%	0.5%	0.1%	25%	31.8%
Dog	10.9%	14.7%	0.8%	0.1%	11%	62.1%
Horse	34.7%	5.3%	0.1%	0.0%	9.4%	51.5%

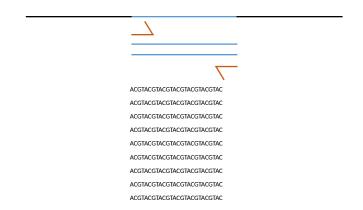




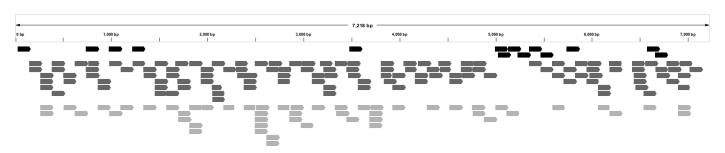
- Unbiased shotgun metagenomics
 - Fragment DNA and sequence randomly



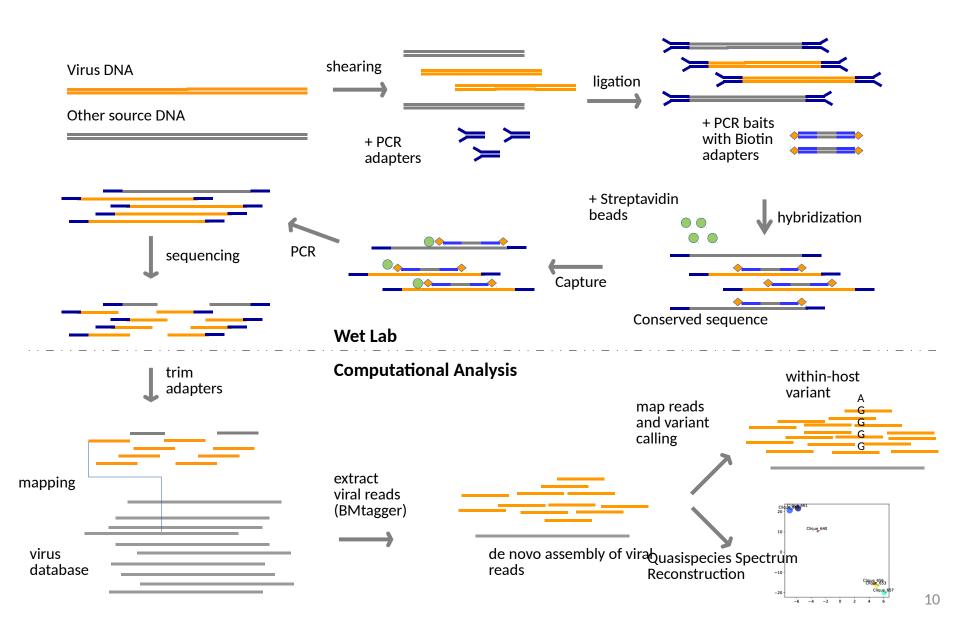
- Amplicon metagenomics
 - PCR amplify a gene of interest
 - Tells you what types of organisms there are
 - Bacteria/Archaea (16S rRNA), Microbial Euks (18S rRNA), Fungi (ITS), Virus (no good marker)



- Targeted-capture enrichment
 - Require probes for a hybridization reaction









Universitätsklinikum Hamburg-Eppendorf

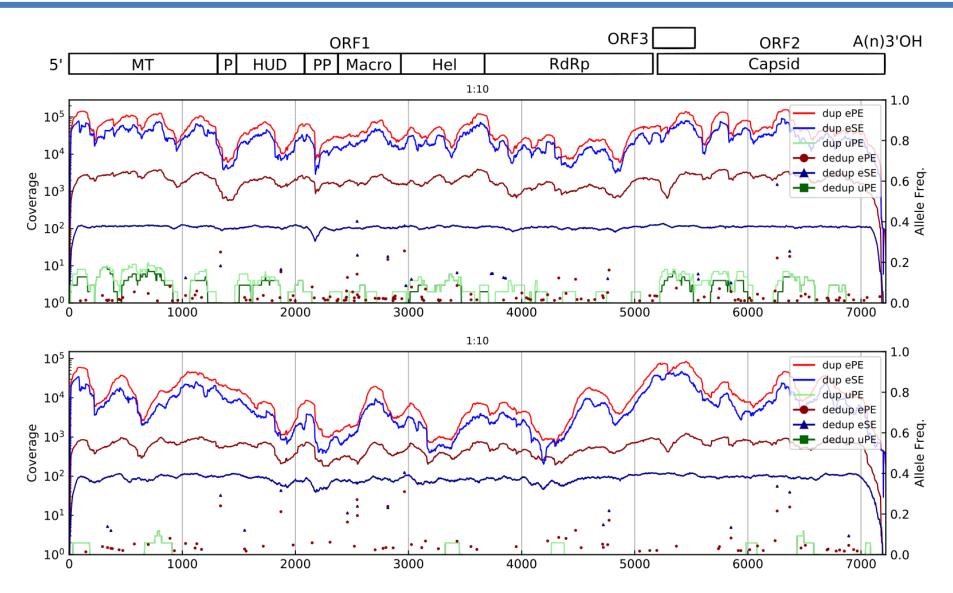
Sample ID	Copy numbe r/ml	enric hed	PE/S E*	Total reads**	HEV reads	Enrich ment rate (%)	De- duplicated HEV reads	Assembly completen ess (%)	Average coverage	SNV numb er***
p2 2x10e 6	10e yes	PE	1,500,000*2	1342724*2	89.51	53542*2	100.0	2024.5	143	
			SE	3,000,000	2675949	89.20	10231	100.0	107.1	16
		no	PE	1,500,000*2	80*2	0.01	51*2	69.8	2.9	0
p2 1:10	1.85x1	.85x1 yes e5	PE	1,500,000*2	396899*2	26.46	14466*2	100.0	576.1	74
1:10 065	ues		SE	3,000,000	768283	25.61	8260	100.0	86.9	14
	no	PE	1,500,000*2	9*2	0.00	8*2	16.3	1.8	0	
p2 2.55x1 1:100 0e4	•	PE	1,500,000*2	12643*2	0.84	846*2	100.0	34.6	2	
		SE	3,000,000	24057	0.80	1267	100.0	13.4	0	
		no	PE	1,500,000*2	0	0.00	0			
p2 2.5x 1:1000 e3	2.5x10	,	PE	1,500,000*2	4987*2	0.33	25*2	41.1	2.2	0
	CO		SE	3,000,000	9309	0.31	17	15.8	1.1	0
		no	PE	1,500,000*2	0	0.00	0			

^{*:} Single-end (76nt) or paired-end (2*150nt)

^{**:} The reads are subsampled to 150,000 pairs (PE) or 300,000 (SE).

^{***:} Filtering strategies: minimum number of reads on each strand >= 5, maximum allowable ration of Number of reads on the two strands <= 10, allele freq >= 0.01.





Coverages and SNVs



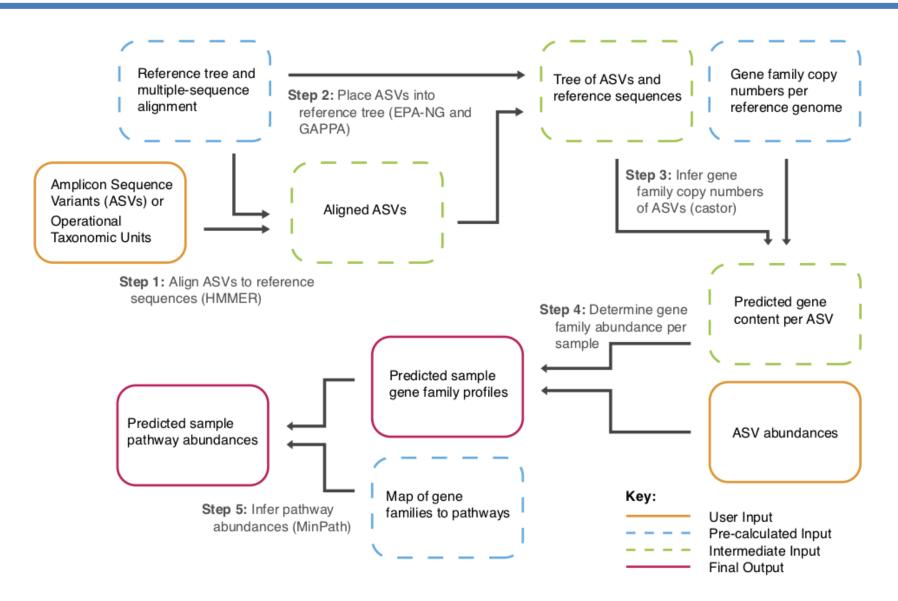
	S1	S2	S3	S4
Taxon1	34	32	26	28
Taxon2	12	17	19	20
Taxon3	11	6	7	3

Precalculated

	Taxon 1	Taxon 2	Taxon3
Gene family 1	1	2	3
Gene family 2	2	1	2
Gene family 3	2	1	1

	S1	S2	S3	S4
Gene family 1	91	84	85	77
Gene family 2	102	93	85	82
Gene family 3	91	87	78	79





Word cloud of Metagenomics



