

Nanopore is preferable over Illumina for 16S amplicon sequencing of the gut microbiota when species-level taxonomic classification, accurate estimation of richness, or focus on rare taxa is required

Márton Szoboszlay<sup>1</sup>, Laetitia Schramm<sup>1</sup>, David Pinzauti<sup>1</sup>, Jeanesse Scerri<sup>1</sup>, Anna Sandionigi<sup>2</sup>, Manuele Biazzo<sup>1</sup>

<sup>1</sup>The BioArte Ltd, San Gwann, Malta

<sup>2</sup>University of Milan Bicocca, DISCO, Milan, Italy

## Supplementary tables

**Table S1: Sequencing yield.**

Sample	Illumina raw read pairs	Nanopore raw reads	Nanopore trimmed reads	Nanopore quality filtered reads
Mock	104889	178513	161733	100506
Mock	126117	181006	164518	101523
Mock	136677	162365	147664	91350
A1	126280	199489	180961	111640
A1	138564	198499	179941	111761
A1	108634	187606	169693	105160
B1	110781	161598	146228	90066
B1	130274	177145	159998	99062
B1	118184	165345	149378	92191
AN1	120367	190866	173056	106591
AN1	125680	181985	164611	100666
AN1	112429	182352	164744	101844
MD01	119083	250784	227059	139562
MD01	109324	230994	209913	127798
MD01	106072	224344	204196	125022
S1X	113130	217891	197947	121585
S1X	136979	236344	214118	131186
S1X	130658	240542	218454	133814
S2P	108602	245526	223887	127347
S2P	121263	236612	214531	133116
S2P	105438	239594	217607	135519

**Table S2: The number of Nanopore reads at each stage of the EzBioCloud MTP pipeline.**

Sample	Uploaded	Low quality or not 16S	Not bacterial	Chimeric	Passed	Classified to species, phylotypes, or species groups
Mock	100000	1273	0	13431	85296	84551

<b>Mock</b>	100001	1293	0	14903	83805	83085
<b>Mock</b>	91350	1119	0	12583	77648	76985
<b>A1</b>	100000	1615	0	23154	75231	74834
<b>A1</b>	100000	1493	0	22369	76138	75776
<b>A1</b>	100000	1554	0	23054	75392	75027
<b>B1</b>	90066	1799	0	22104	66163	65172
<b>B1</b>	99062	2023	0	25646	71393	70284
<b>B1</b>	92191	1786	0	22563	67842	66779
<b>AN1</b>	100000	1691	0	20953	77356	76633
<b>AN1</b>	100000	1717	0	20406	77877	77160
<b>AN1</b>	100000	1684	10	20451	77855	77082
<b>MD01</b>	100000	1769	5	15234	82992	82217
<b>MD01</b>	100000	1952	10	18327	79711	79027
<b>MD01</b>	100000	1883	2	16308	81807	81082
<b>S1X</b>	100000	2205	0	25069	72726	69749
<b>S1X</b>	100000	2176	6	24719	73099	70287
<b>S1X</b>	100000	2185	0	24810	73005	70036
<b>S2P</b>	100000	1377	0	14416	84207	83562
<b>S2P</b>	100000	1473	0	16577	81950	81130
<b>S2P</b>	100000	1308	0	13977	84715	83943

**Table S3: The number of Illumina reads at each stage of the EzBioCloud MTP pipeline.**

<b>Sample</b>	<b>Uploaded</b>	<b>Low quality or not 16S</b>	<b>Not bacterial</b>	<b>Chimeric</b>	<b>Passed</b>	<b>Classified to species, phylotypes, or species groups</b>
<b>Mock</b>	100000	825	119	8007	91049	87586
<b>Mock</b>	100000	934	165	6860	92041	87598
<b>Mock</b>	100000	1094	253	6999	91654	86840
<b>A1</b>	100000	527	18	7136	92319	89941
<b>A1</b>	100000	486	21	6550	92943	90640
<b>A1</b>	100000	507	22	6943	92528	90383
<b>B1</b>	100000	526	91	7088	92295	87673
<b>B1</b>	100000	686	17	7636	91661	86507
<b>B1</b>	100000	505	43	7143	92309	88140
<b>AN1</b>	100000	737	55	8143	91065	86083
<b>AN1</b>	100000	506	44	8150	91300	86315
<b>AN1</b>	100000	690	57	8189	91064	85718
<b>MD01</b>	100000	537	2448	8964	88051	82813
<b>MD01</b>	100000	545	2054	8823	88578	83297
<b>MD01</b>	100000	669	1895	9134	88302	83785
<b>S1X</b>	100000	702	180	8651	90467	85917
<b>S1X</b>	100000	659	136	8431	90774	85948
<b>S1X</b>	100000	671	301	8448	90580	86298
<b>S2P</b>	100000	809	108	10299	88784	79804
<b>S2P</b>	100000	855	51	8848	90246	80417

<b>S2P</b>	100000	787	93	9977	89143	80501
------------	--------	-----	----	------	-------	-------

**Table S4: Relative abundance (%) of the bacterial genera in the mock community compared to the Nanopore and Illumina results from the EzBioCloud MTP pipeline.**

Genus	Composition	Nanopore			Illumina		
<i>Akkermansia</i>	0.97	1.61	1.69	1.58	0.67	0.58	0.39
<i>Bacteroides</i>	9.94	14.06	14.27	13.95	11.26	8.73	14.96
<i>Bifidobacterium</i>	8.78	6.75	7.28	6.99	8.24	9.66	8.28
<i>Clostridioides</i>	2.62	5.43	5.06	5.43	1.69	2.70	2.15
<i>Clostridium</i>	0.0002	0	0	0	0.270	0.013	0.023
<i>Enterococcus</i>	0.0009	0	0	0	0.0868	0.0054	0.0011
<i>Escherichia</i>	12.12	10.82	10.97	10.63	5.41	8.42	7.53
<i>Faecalibacterium</i>	17.63	15.00	14.58	14.64	13.70	17.37	17.37
<i>Fusobacterium</i>	7.49	7.64	7.38	7.09	9.09	13.61	4.47
<i>Lactobacillus</i>	9.63	5.68	5.67	6.01	2.34	4.13	3.26
<i>Prevotella</i>	4.98	6.58	6.44	6.48	5.62	0.94	4.14
<i>Roseburia</i>	9.89	7.15	7.11	7.12	5.79	6.78	6.11
<i>Salmonella</i>	0.009	0.018	0.024	0.012	0	0	0
<i>Veillonella</i>	15.87	19.24	19.52	20.06	5.67	3.33	3.99
<b>Total</b>	99.93	99.97	99.98	99.99	69.83	76.27	72.67

**Table S5: Relative abundance (%) of the bacterial genera in the mock community compared to the Illumina results from the QIIME2 and mothur pipelines.**

Genus	Composition	Illumina QIIME2			Illumina mothur		
<i>Akkermansia</i>	0.97	0.71	0.63	0.38	0.68	0.58	0.36
<i>Bacteroides</i>	9.94	11.30	8.93	15.36	9.51	7.85	13.95
<i>Bifidobacterium</i>	8.78	8.22	9.66	8.31	8.23	9.57	8.13
<i>Clostridioides</i>	2.62	1.75	2.75	2.19	1.69	2.60	2.10
<i>Clostridium</i>	0.0002	0.24	0.0060	0.015	0.23	0.012	0.019
<i>Enterococcus</i>	0.0009	0.11	0	0	0.081	0.0037	0.00086
<i>Escherichia</i>	12.12	5.56	8.62	7.74	5.66	8.43	7.58
<i>Faecalibacterium</i>	17.63	13.73	17.86	17.58	14.20	17.74	17.59
<i>Fusobacterium</i>	7.49	9.13	13.80	4.47	9.22	13.60	4.39
<i>Lactobacillus</i>	9.63	2.43	4.25	3.43	2.28	3.10	2.25
<i>Prevotella</i>	4.98	5.67	0.95	4.27	5.66	0.91	4.13
<i>Roseburia</i>	9.89	4.44	5.68	5.01	4.38	5.58	4.91
<i>Salmonella</i>	0.009	0	0	0	0	0	0
<i>Veillonella</i>	15.87	5.64	3.34	4.16	5.68	3.28	4.03
<b>Total</b>	99.93	69.0	76.5	72.9	67.6	73.3	69.5

**Table S6: The percentage of reads classified into species and genera in the faecal samples.**

	Reads classified as species or phylotypes (%)		Reads classified as genera or phylotype genera (%)	
Sample	Nanopore	Illumina	Nanopore	Illumina

<b>A1</b>	86.2	53.0	99.98	99.40
<b>A1</b>	86.2	52.8	99.99	99.44
<b>A1</b>	85.6	51.5	99.98	99.51
<b>AN1</b>	92.4	28.4	99.97	99.10
<b>AN1</b>	92.4	30.8	99.93	99.16
<b>AN1</b>	92.0	31.1	99.98	99.27
<b>B1</b>	94.8	40.2	99.96	99.35
<b>B1</b>	94.5	39.7	99.97	99.02
<b>B1</b>	94.7	38.7	99.95	99.33
<b>MD01</b>	83.2	53.8	99.77	98.46
<b>MD01</b>	82.2	53.0	99.89	99.18
<b>MD01</b>	82.1	53.5	99.83	99.06
<b>S1X</b>	91.8	53.8	99.79	99.27
<b>S1X</b>	92.0	53.1	99.90	99.25
<b>S1X</b>	91.9	53.6	99.75	99.34
<b>S2P</b>	97.4	52.3	100.00	99.24
<b>S2P</b>	96.8	52.5	99.98	98.95
<b>S2P</b>	97.3	51.1	100.00	99.12

**Table S7: The number of species-level taxa detected in the faecal samples.**

<b>Sample</b>	<b>Species</b>		<b>Species groups</b>		<b>Phylotypes</b>		<b>Phylotype groups</b>	
	Nanopore	Illumina	Nanopore	Illumina	Nanopore	Illumina	Nanopore	Illumina
<b>A1</b>	82	214	18	309	46	933	1	115
<b>A1</b>	82	198	15	286	51	791	1	101
<b>A1</b>	88	193	18	315	49	916	1	107
<b>AN1</b>	127	241	32	352	130	967	1	105
<b>AN1</b>	117	187	31	298	123	816	1	92
<b>AN1</b>	122	229	30	316	122	965	1	116
<b>B1</b>	99	213	15	297	164	900	1	102
<b>B1</b>	96	245	14	343	158	1065	2	117
<b>B1</b>	93	236	11	355	158	864	1	92
<b>MD01</b>	95	197	25	258	179	801	1	88
<b>MD01</b>	101	195	25	293	189	946	1	109
<b>MD01</b>	102	218	24	307	192	1139	2	123
<b>S1X</b>	99	240	18	316	234	1158	1	129
<b>S1X</b>	104	231	20	324	249	1089	1	110
<b>S1X</b>	99	219	20	328	241	1113	2	122
<b>S2P</b>	92	242	16	351	104	1124	1	135
<b>S2P</b>	87	292	11	394	102	1236	1	128
<b>S2P</b>	84	271	9	411	103	1146	1	120

**Table S8: The number of genera detected in the faecal samples.**

<b>Sample</b>	<b>Genera</b>		<b>Phylotype genera</b>	
	Nanopore	Illumina	Nanopore	Illumina

<b>A1</b>	63	392	5	354
<b>A1</b>	64	376	6	318
<b>A1</b>	68	388	6	378
<b>AN1</b>	94	452	31	418
<b>AN1</b>	91	360	28	337
<b>AN1</b>	89	398	29	391
<b>B1</b>	83	395	40	363
<b>B1</b>	76	439	45	442
<b>B1</b>	77	431	42	374
<b>MD01</b>	89	353	56	315
<b>MD01</b>	92	392	59	349
<b>MD01</b>	90	406	65	435
<b>S1X</b>	91	415	65	428
<b>S1X</b>	94	411	76	410
<b>S1X</b>	93	409	72	439
<b>S2P</b>	77	435	26	424
<b>S2P</b>	72	477	27	494
<b>S2P</b>	71	483	26	469

## Supplementary figures

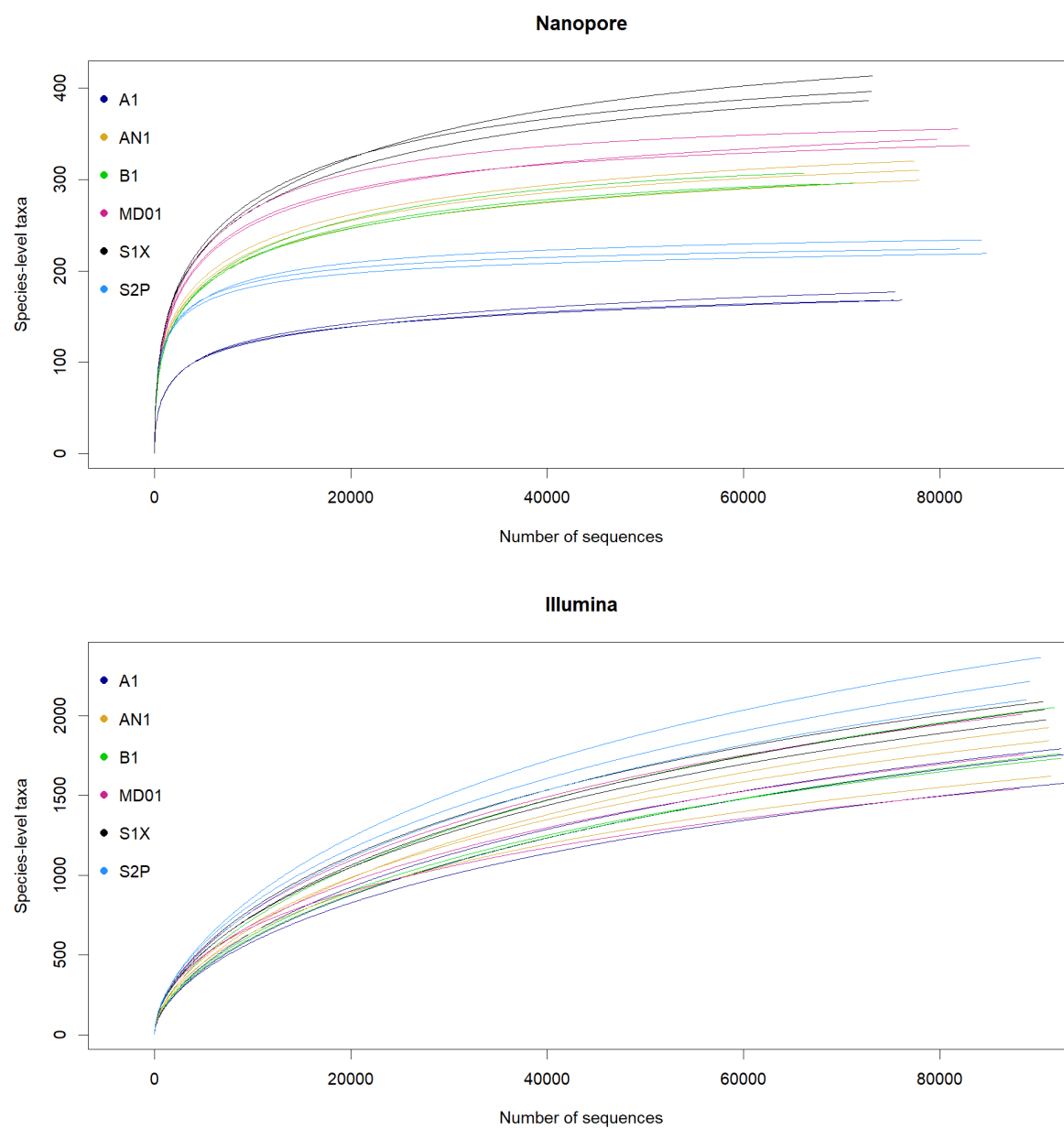
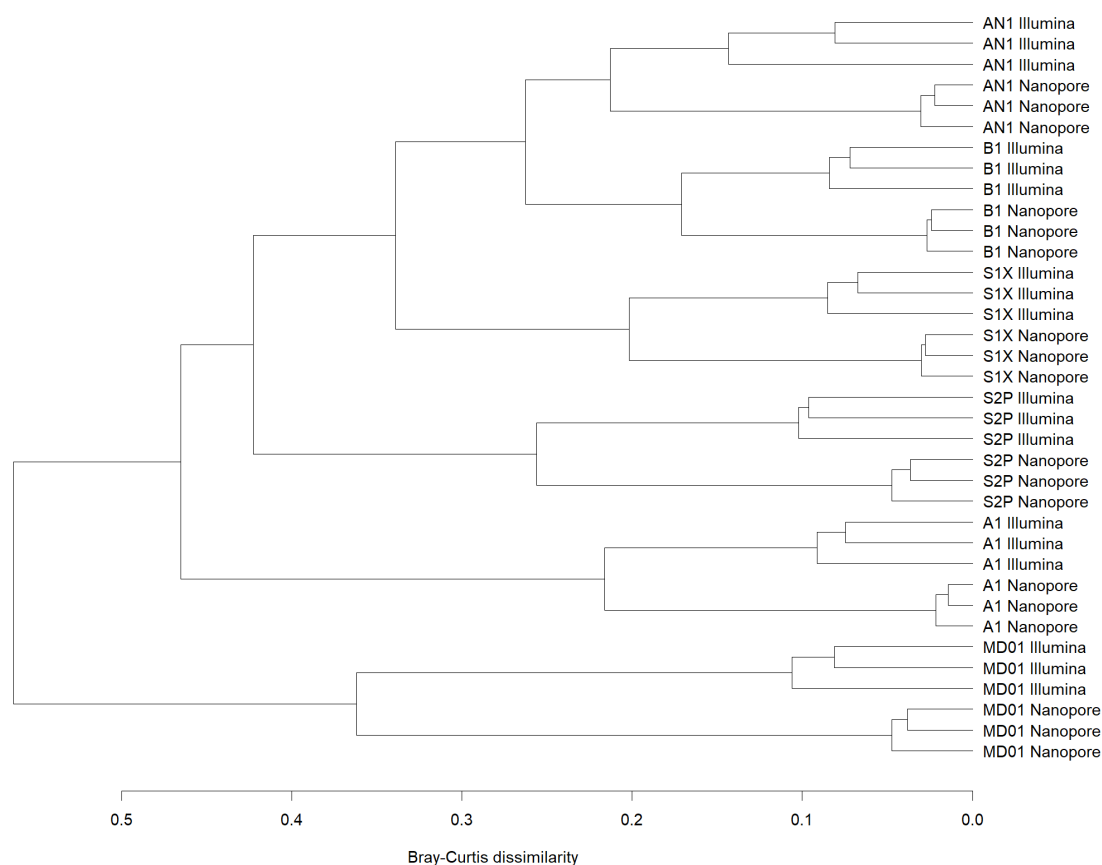
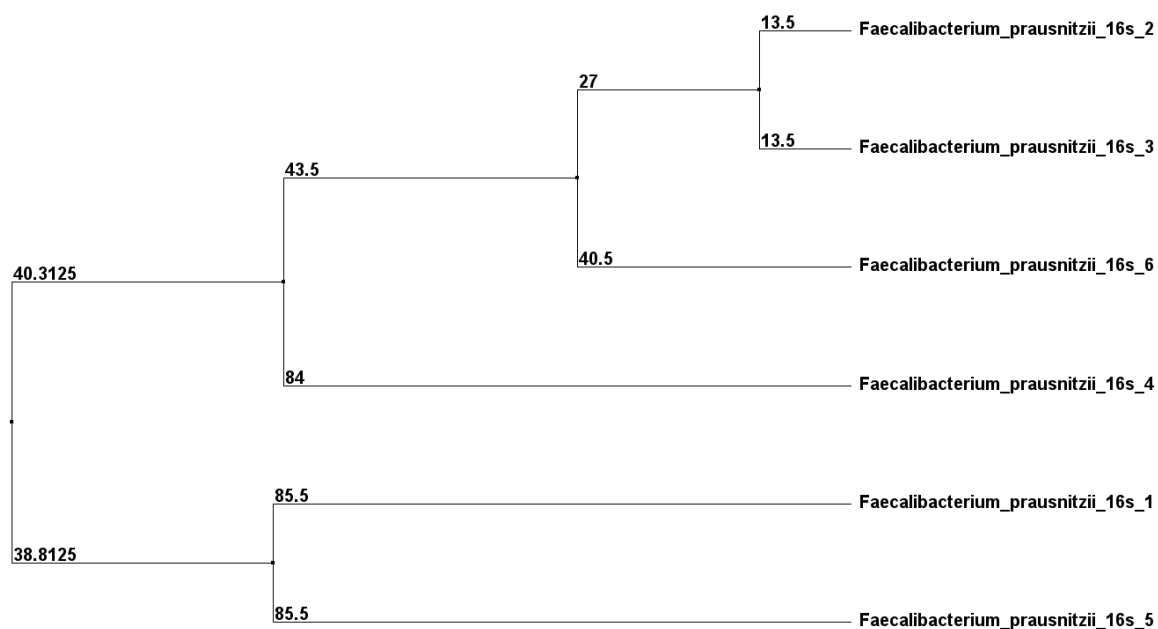


Figure S1: Rarefaction curves.



**Figure S2: Clustering of the faecal samples based on their genus-level bacterial community structure**



**Figure S3: Average distance clustering of the 16S rRNA gene sequences of the *Faecalibacterium prausnitzii* AP34BHI genome after Clustal WS alignment**