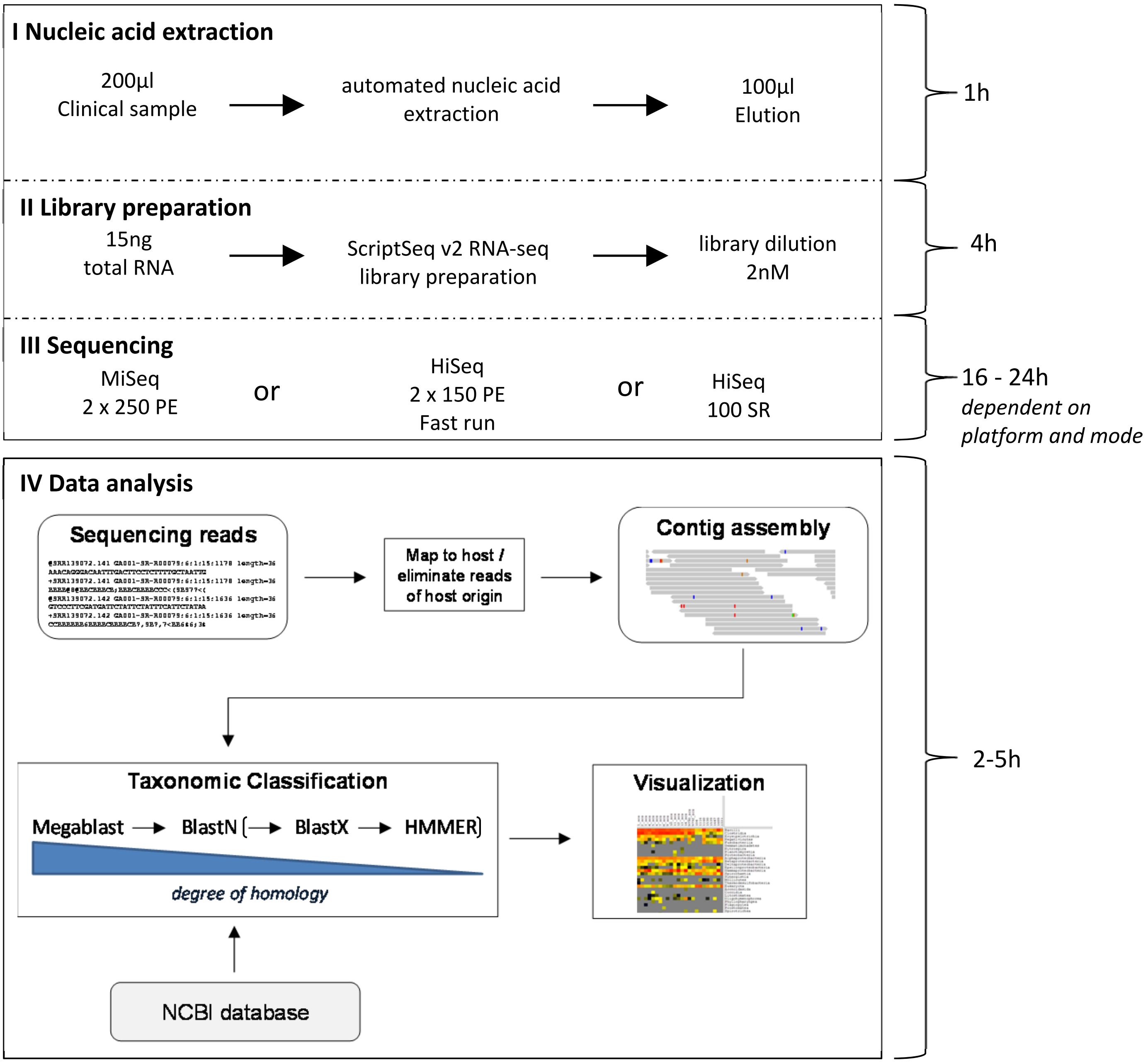
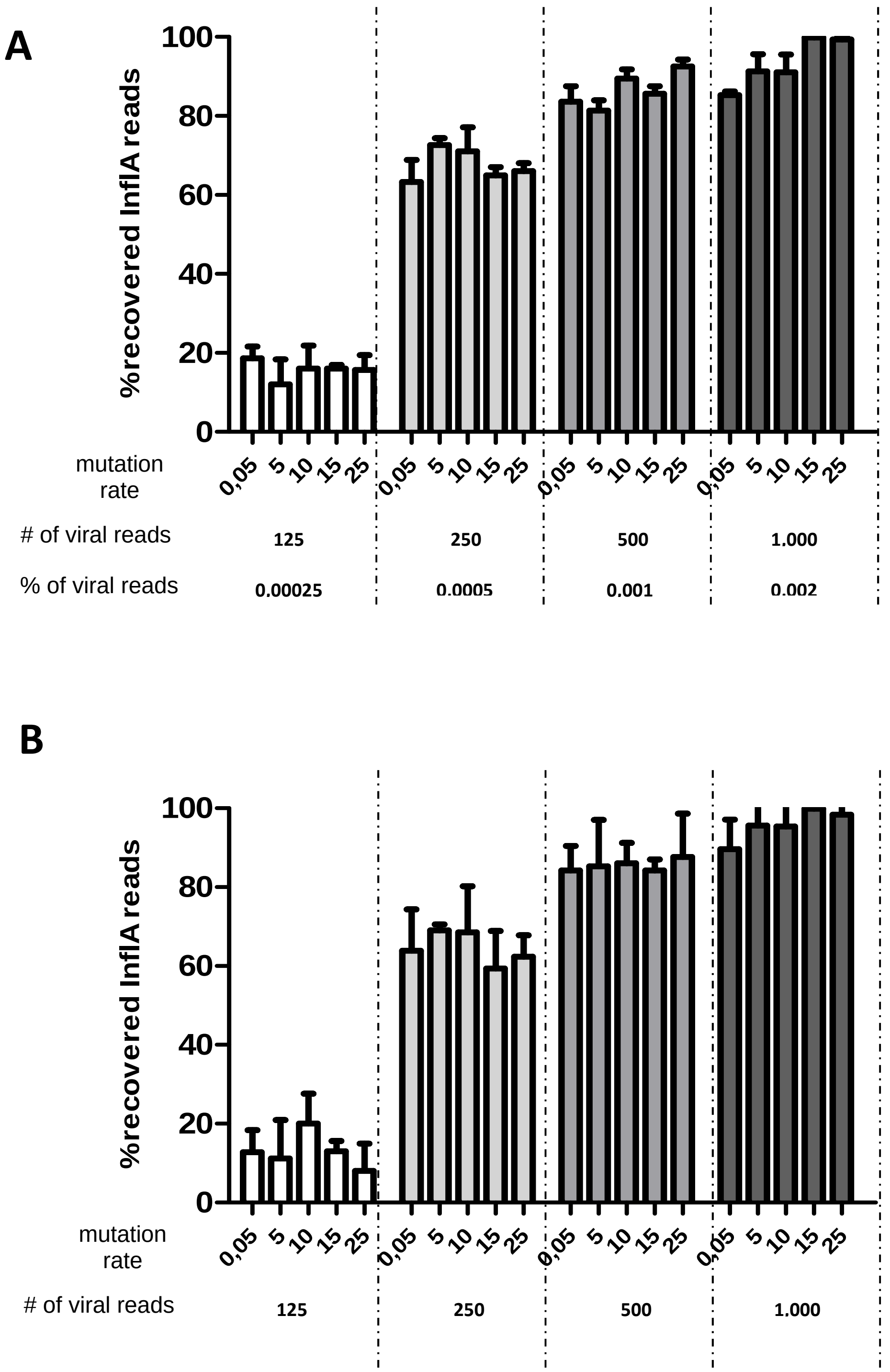
Supplementary Material



Supplementary Figure S1: Summary of workflow and turnaround time of the individual experimental steps including nucleic acid extraction, strand specific RNA library preparation, sequencing on Illumina MiSeq (2 x 250bp paired end, PE), HiSeq2500 (2 x 150bp paired end, fast run mode) or HiSeq2500 (2 x 100bp single read, SR) and bioinformatic data analysis. HiSeq fast run mode was not used in this current study. To allow detection of distantly related infectious agents with a low degree of nucleic acid sequence homology, our data analysis pipeline features optional BlastX and HMMER modules. These modules, however, were not used in the current study, allowing data analysis to complete within 2-5h. Processing time is increased by an additional 5 to 12h if the optional modules are used.



Supplementary Figure S2: Summary of in silico modeling (N=3, shown are mean and standard deviation) to assess sensitivity of bioinformatic analysis. (A) Percentage of InflA reads recovered from datasets with the indicated number of influenza virus reads in a background of 50,000,000 reads generated from the transcriptome dataset of the five influenza virus negative BAL samples. The relative percentage of influenza virus reads is given under the absolute numbers. The mutation rate of simulated viral reads is shown under each column. (B) Percentage of InflA reads recovered from datasets simulated as described in (A), but without the transcriptome background reads.

Supplementary Table S1: Summary on next generation sequencing data.

						filtered contigs[a]			
sample ID	**diagnostic entity**	**instrument**	**total reads**	**host-subtracted reads**	**assembled contigs**	**count**	**mean length**	**mapped**[b] **reads**	**% of host-subtracted reads**
1	BAL	HiSeq2500	10,311,953	4,326,057	510,402	526	459	477,937	11.05%
2	BAL	HiSeq2500	31,998,422	10,158,889	911,694	892	435	2,189,332	21.55%
3	BAL	HiSeq2500	31,953,379	13,744,104	1,582,838	2,288	422	1,111,019	8.08%
4	BAL	HiSeq2500	28,914,182	25,253,498	1,416,064	369	418	1,942,973	7.69%
5	BAL	HiSeq2500	34,164,253	15,246,183	1,316,248	2,459	424	2,405,066	15.77%
104	BAL	HiSeq2500	25,448,729	5,737,099	2,409,216	454	472	121,433	2.12%
755	BAL	HiSeq2500	44,689,504	11,604,195	5,181,101	1,603	448	292,452	2.52%
1116	BAL	HiSeq2500	163,412,609	39,234,172	17,953,404	591	397	78,788	0.20%
1721	BAL	HiSeq2500	27,376,970	5,937,090	2,822,768	444	464	69,980	1.18%
2535	BAL	HiSeq2500	27,472,868	13,945,264	2,338,182	2,589	464	5,920,737	42.46%
2292	BAL	MiSeq	2,226,054	359,262	328,567	16	1,456	5,657	1.57%
3157	BAL	MiSeq	3,265,314	2,050,238	1,175,340	2,799	511	520,512	25.39%
1773	Sputum	HiSeq2500	41,415,879	24,508,834	3,504,725	7,410	538	3,025,434	12.34%
1168	Sputum	HiSeq2500	42,593,701	42,100,823	1,919,424	10,778	659	32,065,730	76.16%
853	Sputum	HiSeq2500	37,054,250	35,908,502	2,216,129	17,233	761	29,404,598	81.89%
677	Sputum	HiSeq2500	32,506,338	21,491,354	2,138,304	6,851	532	16,857,606	78.44%
208	Sputum	HiSeq2500	34,784,935	33,981,137	2,361,555	16,196	679	20,163,035	59.34%
2373	Swap	HiSeq2500	46,324,162	43,280,574	2,261,922	10,700	567	32,703,053	75.56%
2098	Swap	HiSeq2500	46,201,878	23,212,954	2,060,429	4,210	470	2,462,994	10.61%
1689	Swap	HiSeq2500	35,621,107	35,299,029	2,010,101	15,287	718	24,644,607	69.82%
1647	Swap	HiSeq2500	48,790,100	20,794,370	2,435,447	2,191	444	1,250,890	6.02%
1539	Swap	HiSeq2500	29,303,419	22,675,081	3,496,423	8,378	607	16,325,608	72.00%
1538	Swap	HiSeq2500	40,436,634	16,932,677	6,839,715	2,222	545	1,925,912	11.37%
768	Swap	HiSeq2500	34,117,546	32,397,096	2,752,299	18,227	695	23,913,343	73.81%
83	Swap	MiSeq	1,742,025	390,535	268,103	24	979	2,822	0.72%
848	Swap	MiSeq	1,695,921	350,382	129,072	26	908	101,128	28.86%
2295	Swap	HiSeq2500	51,133,693	49,932,909	6,032,741	17,889	515	32,416,062	64.92%
2544	Swap	HiSeq2500	38,427,200	35,076,719	2,533,128	29,688	613	24,287,397	69.24%
14087	secretion	MiSeq	3,538,581	1,673,576	583,224	818	574	745,420	44.54%

[a] contigs with a minimal length of 300 or 400 bp for HiSeq or MiSeq datasets, respectively

[b] total number of host-subtracted reads mapped back to assembled contigs

Supplementary Table S2: Summary on sample composition.

				reads per million mapped reads (RPM):[a]			
sample	**diagnostic entity**	**instrument**	**% host reads**	**bacteria**	**viruses**	**fungi**	**unknown seqs.**
1	BAL	HiSeq2500	58.24%	96,040	5,248	389	329,960
2	BAL	HiSeq2500	68.48%	732,419	1,768	427	95,070
3	BAL	HiSeq2500	57.13%	89,569	80,729	33,574	407,247
4	BAL	HiSeq2500	13.06%	310,900	186,358	0	209,521
5	BAL	HiSeq2500	55.50%	210,577	7,359	179,863	175,017
104	BAL	HiSeq2500	77.78%	7,099	212,348	35,262	175,017
755	BAL	HiSeq2500	74.10%	805,459	2,835	0	63,901
1116	BAL	HiSeq2500	76.00%	19,432	1,193	1,168	594,684
1721	BAL	HiSeq2500	78.47%	80,409	116,190	0	64,404
2535	BAL	HiSeq2500	49.40%	916,464	24	146	59,273
2292	BAL	MiSeq	83.86%	545,519	356,373	29,521	68,588
3157	BAL	MiSeq	37.24%	276,824	5,178	425,454	27,533
1773	Sputum	HiSeq2500	40.91%	903,929	57,562	878	8,217
1168	Sputum	HiSeq2500	1.20%	996,584	210	37	1,378
853	Sputum	HiSeq2500	3.15%	992,253	24	77	900
677	Sputum	HiSeq2500	33.97%	98,296	48	893,643	2,140
208	Sputum	HiSeq2500	2.34%	962,474	29	130	1,635
2373	Swap	HiSeq2500	6.62%	664,920	93	321,710	2,199
2098	Swap	HiSeq2500	50.11%	650,740	9,147	851	167,559
1689	Swap	HiSeq2500	0.93%	990,775	23	1,606	1,254
1647	Swap	HiSeq2500	57.74%	366,297	37,692	2,696	385,947
1539	Swap	HiSeq2500	22.67%	4,829	73	52,070	624
1538	Swap	HiSeq2500	58.26%	19,695	4,495	63,385	4,085
768	Swap	HiSeq2500	5.09%	966,477	37	198	1,460
83	Swap	MiSeq	77.58%	863,926	93,196	0	42,877
848	Swap	MiSeq	79.34%	999,446	208	0	0
2295	Swap	HiSeq2500	2.61%	961,139	550	8	4,504
2544	Swap	HiSeq2500	8.78%	980,301	8	0	1,707
14087	secretion	MiSeq	52.71%	962,951	19	299	36,542

[a] relative abundance of viral and fungal sequences, as well as sequences of unknown origin, is shown as normalized numbers of reads per million mapped reads of non-host origin (RPM).

Supplementary Table S3: Recovery of viral reads from simulated datasets in a background of 50,000,000 reads generated from a transcriptome dataset.

number /percentage of influenza reads		125 / 0.00025 %					250 / 0.0005 %					500 / 0.001 %					1,000 /0.002 %				
mutation rate :		0.05	5	10	15	25	0.05	5	10	15	25	0.05	5	10	15	25	0,05	5	10	15	25
HA	H1N1	11	6				27		26	9	21		10	57	56	65				130	130
					10		14	27	22	21	14	54		11	65	57				129	130
		8					21	21	10	13	14	6		56	12	65		130	130	130	130
NA	H1N1	10		9	10		20	26	27	20	12	37	54	54	50	43	88	107	107	107	107
		9		7	8		22	14	20	14	16	47	54	54	46	54	107	107	107	107	107
				8	9		12	15	18	19	16	54	43	54	43	54	107	107	107	107	107
PA	H1N1					8	21	40	28	37	33	78	82	81	79	82	164	163	164	164	164
							33		23	22	28	78	60	74	81	81	164	164	164	164	164
			8	9			38	22	27	32	18	82	69	64	75	67	164	164	164	164	164
	H1N2				8			8			13										
	H3N2							29													
PB 1	H1N1	7				7	41	35	31	23	19	86	85	78	64	79	173	173	173	173	173
					8	8	17	29	16		31	43	78	82	53	86	173	173	173	173	
			7	9	7		17	29	29	32	17	70	72	78	82	80	173	173	173	173	173
	H3N2										4				8						
											9										
											6										
	H1N2									16											
					7																
	H3N8								16												
	H11N9																				173
PB 2	H1N1				10	7	21	38	33	28	29	77	73	75	74	81	174	174	174	174	174
				10			36	17	18	30	15	110	87	80	64	61	163	174	164	174	174
			7				10	34	23		33	70	80	74	74	66	174	174	174	174	174
	H3N2									10	5										
	H7N2															21					
	N1								11												
	H1N2							10													
										11											
NP	H1N1					7	23	16	28	19		55	57	50	44	47	114	114	114	112	114
							12	20	16	27	23	56	57	55	50	58	114	114	114	114	114
		9	9	8		5	13	16	27	21	13	57	52	57	57	57	114	114	114	114	70
																2					
	H3N2										10										
																					26
	H1N2								4												
M	H1N1						19	19		16	14	37	37	37	37	37	74	74	74	74	74
							11	19	18	6		37	37	37	37		52	74	74	74	74
							13	16	8	13	19	25	37	37			74	74	74	74	74
	H3N2								16												
			8								19					37					
																37					
	H1N2								6						37						
NS	H1N1						12	15	16		13	32	32	32	31		64	64	64	64	64
		7					7	13		16		32	32	32	32	3	64	64	64	64	64
		9					12	16	16	16	16	32	32	32	32	20	64	64	64	64	64
	H1N2						3									16					
	H3N2									15						32					
											15										
Polymerase Subunit	H1N2																				
sum		28	6	9	20	29	184	189	205	167	155	402	430	464	443	468	851	869	870	998	1000
		16	8	17	18	16	155	178	153	152	170	457	405	425	428	473	837	870	860	999	1000
		26	31	34	22	14	136	177	175	167	170	396	385	452	412	446	870	1000	1000	1000	98.2

Shown are the numbers of 3 independent simulations of recovered influenza reads from a datasets of 50,000,000 transcriptome reads (generated from the RNASeq experiments of five InflA negative BAL samples. Different fractions of viral reads (see numbers/percentages in top row) at the indicated mutation rates were used. The viral genome segments and genotypes to which the viral reads were mapped are shown to the left. For each segment and genotype, the number of recovered viral reads as observed in three independently performed simulations is given on a separate line. See supplementary Fig. S2A for a graphical representation of the table data.

Supplementary Table S4: Recovery of viral reads from simulated datasets without any background reads.

number of influenza reads		125					250					500					1,000				
mutation rate :		0.05	5	10	15	25	0.05	5	10	15	25	0.05	5	10	15	25	0,05	5	10	15	25
HA	H1N1	11					25		26	9	21		65	50	57	65	130	130	130	130	130
				7	10		14	28	23	21	14	54		11	57	57				129	130
		5					21	21	10		14	19		11	12	55		130	130	130	130
	N1									2											
NA	H1N1	10	6	9	10		20	26	26	20	12	46	54	54	50	43	88	107	107	107	107
				7	8		22	14	20	14	16	47	54	54	46	54	107	107	107	107	107
				9			12	15	18	7	16	54	43	54	43	54	107	107	107	107	107
PA	H1N1			7			27	38	28	34	33	70	82	81	66	82	164	163	164	164	164
							33		23	22	28	78	60	74	81	81	164	164	164	164	164
			12	9			38	22	27	26	4	82	69	60	75	67	164	164	164	164	164
	H1N2																				
								21													
								8													
	H3N2																				
								29													
PB 1	H1N1					7	41	35	31	28	19	86	85	78	64	79	173	173	173	173	173
					8	8	17	29	19		17	71	78	85	62	86	173	173	173	155	7
				9	7		17	29	29	32	17	70	72	81	75	71	173	173	173	173	173
	H3N2														8						
											6										
	H7N2																				
										16	10									18	166
	H1N2																				
					7																
	H3N8										4										
									16		13										
	H11N9																				
																					173
PB 2	H1N1				10		21	20	33	26	29	77	79	75	74	81	174	174	174	174	174
				10			36	17	18	30	15	119	87	80	64	61	163	174	164	174	174
		9	7				10	34	23		33	64	78	74	74	72	174	174	174	174	149
	H3N2																				
										10	5										
	H7N2																				
																21					
	N1																				
									11				5								
	H1N2																				
								10													
										11											
NP	H1N1					7	23	16	28	19		55	57	50	44	47	114	114	114	112	114
							12	20	16	27	23	56	49	55	50		114	114	114	114	114
			9	8			13	16	16	21	13	57	52	57	57		114	114	114	114	70
	H1															2					
	H3N2										10										
																57					
																					20
	H1N2																				
									4												
M	H1N1						19	19		16	14	37	37	37	37	37	74	74	74	74	74
							11	19	18	6			37	37	37	37	52	74	74	74	59
							13	16	6		19	25	37	37			74	74	74	74	74
	H5N1																				
																					15
	H3N2								16												
			8								19										
																37					
	H1N2																				
															37						
NS	H1N1						12	15	16		13	32	32	32	26		64	64	64	64	64
		8					7	13		16		32	32	32	32	19	64	64	64	64	64
		5					12	16	12	16	16	32	32	32	32	20	64	64	64	64	64
	H1N2																				
							3														
	H3N2									15						32					
											15										
Polymerase Subunit	H1N2																				
sum		21	6	16	20	14	188	169	204	169	155	403	491	457	426	468	981	999	1000	998	1000
		8	8	24	18	16	155	171	157	152	170	457	397	428	432	473	837	870	860	999	1000

19	28	35	14	0	136	177	152	123	143	403	388	406	405	376	870	1000	1000	1000	951

Shown are the numbers of 3 independent simulations of recovered influenza reads. Different fractions of viral reads (see numbers in top row) were used. The viral genome segments and genotypes to which the viral reads were mapped are shown to the left. For each segment and genotype, the number of recovered viral reads as observed in three independently performed simulations is given on a separate line. See supplementary Fig. S2B for a graphical representation of the table data.

Supplementary Table S5: Neuraminidase inhibitor resistance mutations.

	104	**208**	**1539**	**1538**	**83**	**2292**
aa position						
H1N1						
Q136K	-	-	-	-	-	-
H275Y	-	-	-	-	-	-
N70S	+	+	+	+	+	+
Y155H	-	n.a	n.a	n.a	n.a	-
I122V/M	-	n.a	-	n.a	-	-
D199N	-	n.a.	-	n.a.	-	-
I223V	-	-	-	-	-	-

	1721	**1773**	**1647**	**1168**	**3157**	**2098**
H3N2						
N294S	-	-	-	-	-	-
R292K	-	-	-	-	-	-
E119A/D	-	-	-	-	-	-
E119G	-	-	-	-	-	-
E119I	-	-	-	-	-	-
E119V	-	-	-	-	-	-
Q136K	-	-	-	-	-	-
R371K	-	-	-	-	-	-
D151A	-	-	-	-	-	-

All H3N2 strains are predicted to be sensitive to Neuraminidase inhibitors (Oseltamivir, Zanamivir, Peramivir). H1N1 strains are likely sensitive to Oseltamivir and Peramivir (H275Y mutation), but showed the N70S mutation which typically mediates medium resistance against Zanamivir (1, 2).
n.a.: not applicable

Supplementary Table S6: Influenza negative samples; NGS data of pathogens known to cause respiratory infections and results of diagnostic culture.

ID	RPM InflA/InflB	RPM S. pneum.	RPM Neisseria meningitidis	RPM H. influenzae	RPM M. catarrhalis.	RPM S. aureus	RPM Candida dubliniensis	RPM Candida albicans	RPM E. faecalis	Culture (bacteria; yeast) and PCR for viruses
BAL										
1	-	-	-	-	-	-	-	-	-	negative
2	-	-	-	699,508	-	-	-	-	-	H. influenzae,
3[a]	-	-	-	-	-	-	-	26,419	-	Candida sp.; RSV
4[b]	-	-	-	-	-	-	-	-	-	Enterococcus sp; Pseudomonas sp., Rhinovirus
5	-	-	-	-	-	-	-	-	-	Yeast, Pseudomonas sp. Lactobacillus,

Values shown represent normalized number of reads per million mapped reads (RPM).
a, b: viral pathogens detected, a: RSV 63,941RPM; b: human Rhinovirus A 8,360RPM.

Supplementary Table S7: Analysis of rRNA sequence contigs

					% of reads in rRNA contigs unambigously assigned on level:[c]					
sample	rRNA contig s[a]	% of all contigs	rRNA reads [b]	% of all mapped reads	super-kingdom	phylum	class	family	genus	species
1	26	4.9%	27,713	5.80%		12.8%	0.7%	37.5%	10.9%	38.2%
2	18	2.0%	762,195	34.81%	0.4%				0.2%	99.5%
3	16	0.7%	44,564	4.01%					85.2%	14.8%
4	4	1.1%	437,843	22.53%					0.1%	99.9%
5	24	1.0%	586,259	24.38%	3.0%				0.6%	96.4%
104	9	2.0%	273	0.22%				9.2%		90.8%
755	32	2.0%	201,514	68.90%	0.3%			0.2%	0.1%	99.4%
1116	6	1.0%	306	0.39%	57.8%			19.6%	9.2%	13.4%
1721	19	4.3%	2,050	2.93%	20.2%				11.0%	68.8%
2535	41	1.6%	2,766,365	46.72%	0.0%		0.0%	0.0%	3.6%	96.4%
2292	3	18.8%	1,502	26.55%						100.0%
3157	13	0.5%	155,896	29.95%		0.1%				99.9%
1773	39	0.5%	1,216,704	40.22%		0.0%	0.1%		3.4%	96.6%
1168	47	0.4%	8,983,773	28.02%				0.1%	28.2%	71.7%
853	43	0.2%	2,231,156	7.59%	0.7%				14.8%	84.5%
677	26	0.4%	5,857,447	34.75%					0.5%	99.5%
208	51	0.3%	2,888,938	14.33%	0.0%		2.5%	2.7%	27.5%	67.3%
2373	56	0.5%	2,010,084	6.15%	9.3%			23.9%	1.2%	65.6%
2098	21	0.5%	1,252,535	50.85%				0.5%	0.8%	98.7%
1689	46	0.3%	4,746,334	19.26%	6.7%			0.1%	9.3%	83.9%
1647	25	1.1%	147,774	11.81%	0.3%				24.8%	74.9%
1539	34	0.4%	117,881	0.72%	1.1%				5.9%	93.0%
1538	18	0.8%	1,679,115	87.19%	0.0%				0.0%	100.0%
768	49	0.3%	3,239,936	13.55%	1.2%			1.0%	8.4%	89.4%
83	6	25.0%	2,292	81.22%	0.4%				1.4%	98.2%
848	6	23.1%	49,350	48.80%					0.2%	99.8%
2295	64	0.4%	8,988,490	27.73%	1.4%		2.6%		22.3%	73.6%
2544	37	0.1%	3,785,736	15.59%	0.0%	8.9%			26.5%	64.6%
14087	15	1.8%	198,810	26.67%						100.0%

[a] number of sequence contigs exhibiting significant homology to ribosomal RNAs (see material and methods for details)

[b] total number of reads mapped to ribosomal sequence contigs

[c] percentage of reads (of all reads mapped to ribosomal contigs) from contigs that can be unambiguously assigned at or below each of the indicated taxonomic levels

Supplementary References

1. **Deyde VM, Sheu TG, Trujillo AA, Okomo-Adhiambo M, Garten R, Klimov AI, Gubareva LV.** 2010. Detection of molecular markers of drug resistance in 2009 pandemic influenza A (H1N1) viruses by pyrosequencing. Antimicrob Agents Chemother **54:**1102-1110.
2. **Kamali A, Holodniy M.** 2013. Influenza treatment and prophylaxis with neuraminidase inhibitors: a review. Infect Drug Resist **6:**187-198.