

Riboprofiling - pipeline

23/07/2016

Outline

Stats - trimming, mapping to contaminants and genome

extra stats - genome mapping after rRNA removal

Steps of the pipeline

- ▶ trimming - CutAdapt
- ▶ filter out reads mapping to rRNA
- ▶ filter out reads mapping to small RNA
- ▶ mapping to the genome
 - ▶ minimal match 16nt
 - ▶ minimal match 18nt

Number of reads in the steps of the pipeline

	Initial	Trimming	rRNA	small RNA
labExpld	nb-reads	TR:nb-reads	CM:nb-reads	SR:nb-reads
H001.1RP	147,212,687	127,743,822 (86.8%)	67,252,826	59,169,455
H003.1RP	118,082,384	98,513,634 (83.4%)	70,341,886	62,102,678
H006.1RP	145,860,276	104,443,932 (71.6%)	72,017,672	63,827,154
H009.1RP	138,383,660	113,314,770 (81.9%)	11,732,073	9,185,567
H012.1RP	117,319,391	77,433,945 (66.0%)	12,093,595	10,015,288
H018.1RP	185,026,553	159,066,506 (86.0%)	106,923,516	97,977,438
H024.1RP	145,379,749	132,273,796 (91.0%)	86,596,298	76,087,809
H036.1RP	135,397,014	105,805,374 (78.1%)	44,522,823	33,171,278
H048.1RP	117,264,568	93,580,545 (79.8%)	68,837,224	60,649,426
H072.1RP	149,441,806	94,029,110 (62.9%)	66,967,288	58,948,258
H120.1RP	169,778,727	126,029,247 (74.2%)	73,382,333	63,446,749
H168.1RP	115,386,623	70,220,531 (60.9%)	31,505,037	24,922,235

Stats trimming

labExpld	TR:Total reads processed	TR:Reads with adapters	TR:Reads too short	TR:Reads passing filters
H001.1RP	147,212,687	143,586,273 (97.5%)	15,878,480 (10.8%)	127,743,822 (86.8%)
H003.1RP	118,082,384	114,113,805 (96.6%)	15,691,235 (13.3%)	98,513,634 (83.4%)
H006.1RP	145,860,276	141,021,393 (96.7%)	36,602,250 (25.1%)	104,443,932 (71.6%)
H009.1RP	138,383,660	136,806,541 (98.9%)	23,494,863 (17.0%)	113,314,770 (81.9%)
H012.1RP	117,319,391	115,741,397 (98.7%)	38,311,846 (32.7%)	77,433,945 (66.0%)
H018.1RP	185,026,553	181,188,407 (97.9%)	22,378,837 (12.1%)	159,066,506 (86.0%)
H024.1RP	145,379,749	142,824,488 (98.2%)	10,567,509 (7.3%)	132,273,796 (91.0%)
H036.1RP	135,397,014	128,202,445 (94.7%)	22,577,496 (16.7%)	105,805,374 (78.1%)
H048.1RP	117,264,568	114,883,752 (98.0%)	21,317,227 (18.2%)	93,580,545 (79.8%)
H072.1RP	149,441,806	146,379,483 (98.0%)	52,384,402 (35.1%)	94,029,110 (62.9%)
H120.1RP	169,778,727	166,279,650 (97.9%)	40,431,260 (23.8%)	126,029,247 (74.2%)
H168.1RP	115,386,623	113,752,239 (98.6%)	43,538,731 (37.7%)	70,220,531 (60.9%)

Stats rRNA mapping

labExpld	CM: nb input reads	CM:uniq mapped	CM:prop uniq	CM:multi mapped	CM:prop multi
H001.1RP	127,743,822	59,775,214	46.79%	715,782	0.56%
H003.1RP	98,513,634	27,657,267	28.07%	514,481	0.52%
H006.1RP	104,443,932	31,728,256	30.38%	698,004	0.67%
H009.1RP	113,314,770	101,464,436	89.54%	118,261	0.10%
H012.1RP	77,433,945	65,251,465	84.27%	88,885	0.11%
H018.1RP	159,066,506	51,351,414	32.28%	791,576	0.50%
H024.1RP	132,273,796	44,792,441	33.86%	885,057	0.67%
H036.1RP	105,805,374	60,523,052	57.20%	759,499	0.72%
H048.1RP	93,580,545	24,342,994	26.01%	400,327	0.43%
H072.1RP	94,029,110	26,530,615	28.22%	531,207	0.56%
H120.1RP	126,029,247	52,170,580	41.40%	476,334	0.38%
H168.1RP	70,220,531	38,446,345	54.75%	269,149	0.38%

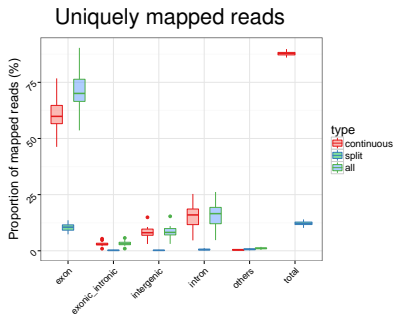
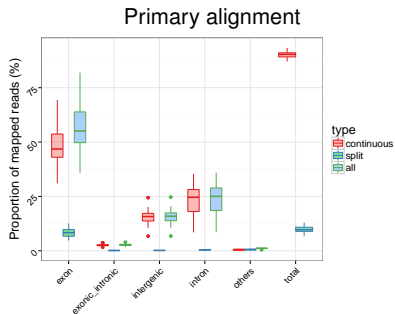
Stats small RNA mapping

labExpld	SR: nb input reads	SR:uniq mapped	SR:prop uniq	SR:multi mapped	SR:prop multi
H001.1RP	67,252,826	4,488,050	6.67%	3,595,321	5.35%
H003.1RP	70,341,886	4,913,827	6.99%	3,325,381	4.73%
H006.1RP	72,017,672	3,526,374	4.90%	4,664,144	6.48%
H009.1RP	11,732,073	1,620,574	13.81%	925,932	7.89%
H012.1RP	12,093,595	1,514,484	12.52%	563,823	4.66%
H018.1RP	106,923,516	4,297,118	4.02%	4,648,960	4.35%
H024.1RP	86,596,298	6,235,908	7.20%	4,272,581	4.93%
H036.1RP	44,522,823	7,901,804	17.75%	3,449,741	7.75%
H048.1RP	68,837,224	4,795,012	6.97%	3,392,786	4.93%
H072.1RP	66,967,288	4,091,084	6.11%	3,927,946	5.87%
H120.1RP	73,382,333	5,011,861	6.83%	4,923,723	6.71%
H168.1RP	31,505,037	3,186,674	10.11%	3,396,128	10.78%

Stats genome mapping (16nt minimal match) after rRNA and small RNA removal

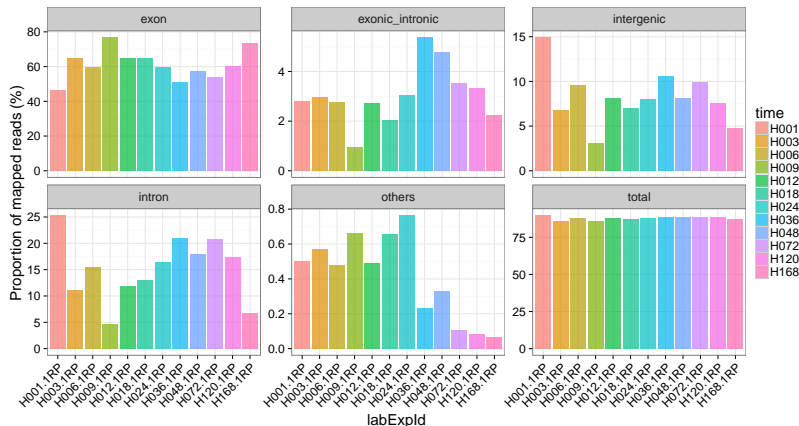
labExpld	GM16: nb input reads	GM:uniq mapped	GM16:prop uniq	GM16:multi mapped	GM16:prop multi
H001.1RP	59,169,455	7,008,442	11.84%	7,513,490	12.70%
H003.1RP	62,102,678	11,657,068	18.77%	7,483,764	12.05%
H006.1RP	63,827,154	7,023,495	11.00%	4,657,957	7.30%
H009.1RP	9,185,567	2,283,386	24.86%	1,007,345	10.97%
H012.1RP	10,015,288	1,050,070	10.48%	563,821	5.63%
H018.1RP	97,977,438	10,899,463	11.12%	7,416,521	7.57%
H024.1RP	76,087,809	11,586,748	15.23%	7,606,255	10.00%
H036.1RP	33,171,278	12,033,726	36.28%	10,650,276	32.11%
H048.1RP	60,649,426	9,238,677	15.23%	5,921,051	9.76%
H072.1RP	58,948,258	9,563,735	16.22%	6,685,270	11.34%
H120.1RP	63,446,749	15,149,743	23.88%	10,464,940	16.49%
H168.1RP	24,922,235	6,768,218	27.16%	2,635,071	10.57%

Stats genome mapping (16nt minimal match) after rRNA and small RNA removal



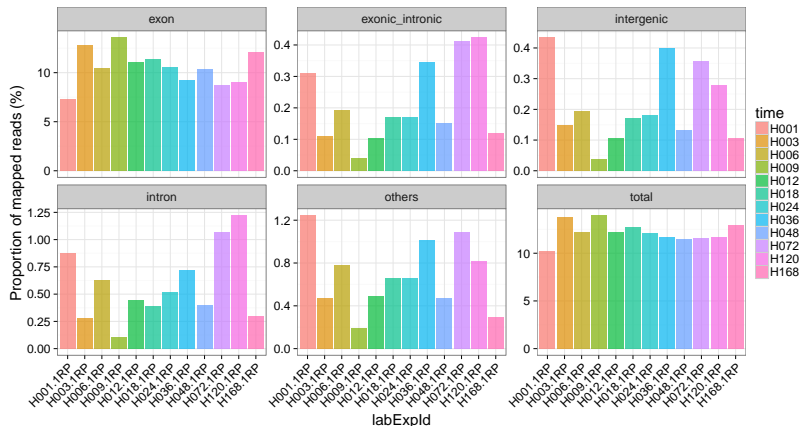
Stats genome mapping (16nt minimal match) after rRNA and small RNA removal

Uniquely mapped reads



Stats genome mapping (16nt minimal match) after rRNA and small RNA removal

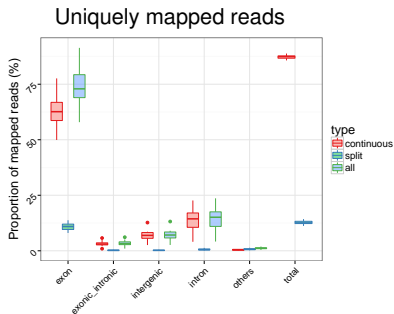
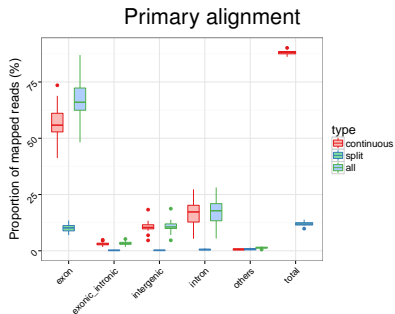
Uniquely mapped reads



Stats genome mapping (18nt minimal match) after rRNA and small RNA removal

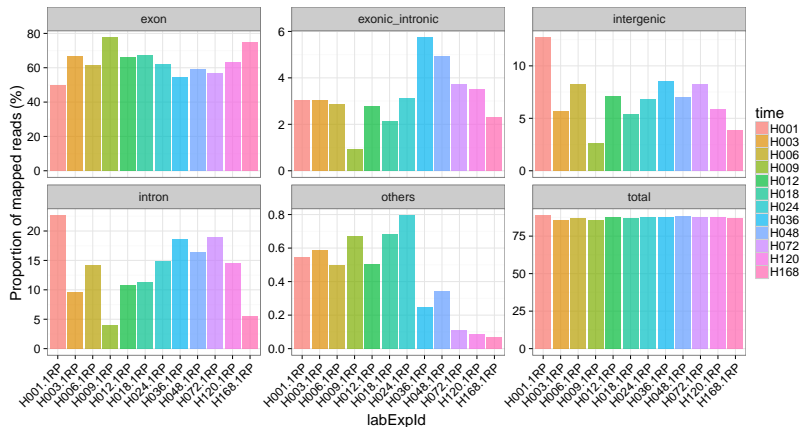
labExpId	GM18: nb input reads	GM18:uniq mapped	GM18:prop uniq	GM18:multi mapped	GM18:prop multi
H001.1RP	59,169,455	6,354,662	10.74%	3,404,735	5.75%
H003.1RP	62,102,678	11,274,362	18.15%	4,929,473	7.94%
H006.1RP	63,827,154	6,733,771	10.55%	2,803,484	4.39%
H009.1RP	9,185,567	2,254,908	24.55%	824,828	8.98%
H012.1RP	10,015,288	1,022,701	10.21%	395,726	3.95%
H018.1RP	97,977,438	10,401,340	10.62%	4,374,842	4.47%
H024.1RP	76,087,809	11,116,270	14.61%	4,426,776	5.82%
H036.1RP	33,171,278	11,124,893	33.54%	5,019,882	15.13%
H048.1RP	60,649,426	8,903,103	14.68%	3,220,880	5.31%
H072.1RP	58,948,258	8,984,081	15.24%	3,586,280	6.08%
H120.1RP	63,446,749	14,117,559	22.25%	5,084,212	8.01%
H168.1RP	24,922,235	6,590,715	26.45%	1,692,794	6.79%

Stats genome mapping (18nt minimal match) after rRNA and small RNA removal



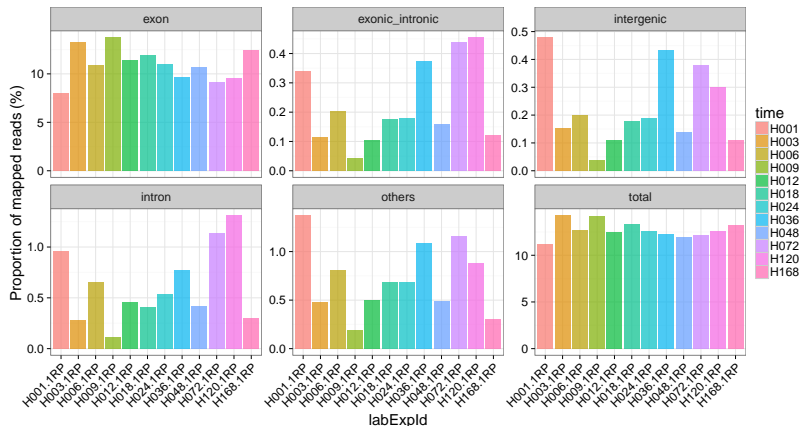
Stats genome mapping (18nt minimal match) after rRNA and small RNA removal

Uniquely mapped reads



Stats genome mapping (18nt minimal match) after rRNA and small RNA removal

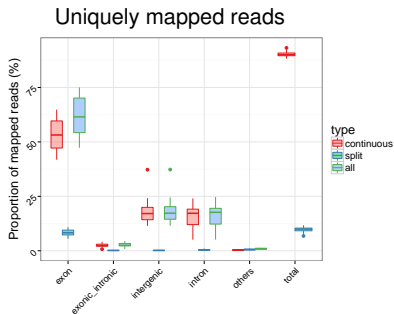
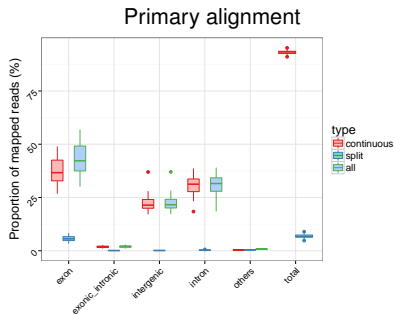
Uniquely mapped reads



Stats genome mapping (16nt minimal match) after rRNA removal

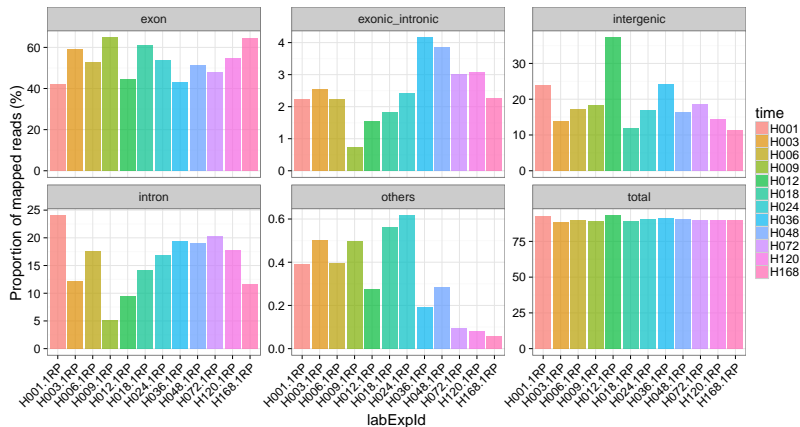
labExpId	GM16: nb input reads	GM16:uniq mapped	GM16:prop uniq	GM16:multi mapped	GM16:prop multi
H001.1RP	67,252,826	10,185,642	15.15%	12,077,202	17.96%
H003.1RP	70,341,886	15,008,085	21.34%	12,107,973	17.21%
H006.1RP	72,017,672	9,456,166	13.13%	10,179,924	14.14%
H009.1RP	11,732,073	3,296,351	28.10%	2,461,399	20.98%
H012.1RP	12,093,595	2,038,743	16.86%	1,585,971	13.11%
H018.1RP	106,923,516	14,250,488	13.33%	12,753,527	11.93%
H024.1RP	86,596,298	16,171,824	18.67%	13,263,058	15.32%
H036.1RP	44,522,823	17,177,484	38.58%	16,367,311	36.76%
H048.1RP	68,837,224	12,534,585	18.21%	10,645,521	15.46%
H072.1RP	66,967,288	12,330,119	18.41%	11,660,935	17.41%
H120.1RP	73,382,333	18,959,585	25.84%	16,166,265	22.03%
H168.1RP	31,505,037	9,247,618	29.35%	6,621,122	21.02%

Stats genome mapping (16nt minimal match) after rRNA removal



Stats genome mapping (16nt minimal match) after rRNA removal

Uniquely mapped reads



Stats genome mapping (16nt minimal match) after rRNA removal

Uniquely mapped reads

