

Contents

1 Basic edgeR results for Project mmuTestRun2-microRNA	1
2 Introduction	2
3 Statistical read count analysis	2
3.1 Read count statistics	4
3.2 microRNA diversity across the samples	7
4 Exploratory data analysis: heatmaps and hierachical clustering	7
4.1 All microRNAs	8
4.2 30 most variant microRNAs	9
5 Exploratory data analysis: principal component analysis (PCA)	10
6 Differential Expression Analysis	10
6.1 young soleus vs. old soleus	11
6.2 young EDL vs. old EDL	15
6.3 young EDL vs. young soleus	19
6.4 old EDL vs. old soleus	23
7 Summary of Differential Expression	27
8 Reproducibility	27
9 Bibliography	29



1 Basic edgeR results for Project mmuTestRun2-microRNA

This report was rendered at 2018-07-06 00:18:33 according to instructions for analysis by the customer. Please read carefully. In case questions arise please contact our responsible project manager at TAmiRNA directly or via office@tamirna.com.

2 Introduction

This report is meant to help explore the results obtained from the analysis of a small RNA sequencing experiment. We have used the publicly available Bioconductor toolbox edgeR (Robinson, McCarthy, and Smyth, 2010; McCarthy, J., Chen, Yunshun, et al., 2012; Chen, Lun, and Smyth, 2014) to prepare these results. While the report is rich, it is meant to just start the exploration of the results.

This report is furthermore complementary to the Quality Control (QC) report, which has been sent separately to this report.

3 Statistical read count analysis

This section includes the following technical replicates:

replicates:

```
ID:  old_sol_1903    sample index:    old_oleus_7
      (input file:    60525_ATGTCA_CBRBBANXX_8_20180128B_20180128 )
```

```
ID:  old_sol_1903_2  sample index:    old_oleus_8
      (input file:    60526_GTCCGC_CBRBBANXX_8_20180128B_20180128_WH )
```

replicates:

```
ID:  y_edl_1915      sample index:    young_EDL_1
      (input file:    60495_ATCACG_CBRBBANXX_7_20180128B_20180128 )
```

```
ID:  y_edl_1915_2    sample index:    young_EDL_6
      (input file:    60510_GTAGAG_CBRBBANXX_7_20180128B_20180128_WH )
```

Note that files `*._WH` are considered throughout this section but excluded from differential expression analysis.

Table 1: input file reference

customerID	indexedSamples	inputFile
old_EDL		
old_edl_1908	old_EDL_1	60500_GCCAAT_CBRBBANXX_7_20180128B_20180128
old_edl_1903	old_EDL_2	60506_CTTGTA_CBRBBANXX_7_20180128B_20180128
old_edl_1905	old_EDL_3	60512_CGATGT_CBRBBANXX_8_20180128B_20180128
old_edl_1907	old_EDL_4	60513_TTAGGC_CBRBBANXX_8_20180128B_20180128
old_edl_1906	old_EDL_5	60518_ACTTGA_CBRBBANXX_8_20180128B_20180128
old_edl_1902	old_EDL_6	60519_GATCAG_CBRBBANXX_8_20180128B_20180128
old_edl_1904	old_EDL_7	60520_TAGCTT_CBRBBANXX_8_20180128B_20180128
old_soleus		
old_sol_1904	old_soleus_1	60498_TGACCA_CBRBBANXX_7_20180128B_20180128
old_sol_1902	old_soleus_2	60504_TAGCTT_CBRBBANXX_7_20180128B_20180128
old_sol_1908	old_soleus_3	60505_GGCTAC_CBRBBANXX_7_20180128B_20180128
old_sol_1906	old_soleus_4	60508_AGTTC_CBRBBANXX_7_20180128B_20180128
old_sol_1905	old_soleus_5	60522_CTTGTA_CBRBBANXX_8_20180128B_20180128
old_sol_1907	old_soleus_6	60523_AGTCAA_CBRBBANXX_8_20180128B_20180128
old_sol_1903	old_soleus_7	60525_ATGTCA_CBRBBANXX_8_20180128B_20180128
old_sol_1903_2	old_soleus_8	60526_GTCCGC_CBRBBANXX_8_20180128B_20180128_WH
young_EDL		
y_edl_1915	young_EDL_1	60495_ATCACG_CBRBBANXX_7_20180128B_20180128
y_edl_1916	young_EDL_2	60496_CGATGT_CBRBBANXX_7_20180128B_20180128
y_edl_1917	young_EDL_3	60497_TTAGGC_CBRBBANXX_7_20180128B_20180128
y_edl_1910	young_EDL_4	60499_ACAGTG_CBRBBANXX_7_20180128B_20180128
y_edl_1914	young_EDL_5	60501_CAGATC_CBRBBANXX_7_20180128B_20180128
y_edl_1915_2	young_EDL_6	60510_GTAGAG_CBRBBANXX_7_20180128B_20180128_WH
y_edl_1912	young_EDL_7	60511_ATCACG_CBRBBANXX_8_20180128B_20180128
y_edl_1913	young_EDL_8	60515_ACAGTG_CBRBBANXX_8_20180128B_20180128
y_edl_1911	young_EDL_9	60521_GGCTAC_CBRBBANXX_8_20180128B_20180128
young_soleus		
y_sol_1915	young_soleus_1	60502_ACTTGA_CBRBBANXX_7_20180128B_20180128
y_sol_1917	young_soleus_2	60503_GATCAG_CBRBBANXX_7_20180128B_20180128
y_sol_1913	young_soleus_3	60507_AGTCAA_CBRBBANXX_7_20180128B_20180128
y_sol_1916	young_soleus_4	60509_ATGTCA_CBRBBANXX_7_20180128B_20180128
y_sol_1911	young_soleus_5	60514_TGACCA_CBRBBANXX_8_20180128B_20180128
y_sol_1910	young_soleus_6	60516_GCCAAT_CBRBBANXX_8_20180128B_20180128
y_sol_1912	young_soleus_7	60517_CAGATC_CBRBBANXX_8_20180128B_20180128
y_sol_1914	young_soleus_8	60524_AGTTC_CBRBBANXX_8_20180128B_20180128

3.1 Read count statistics

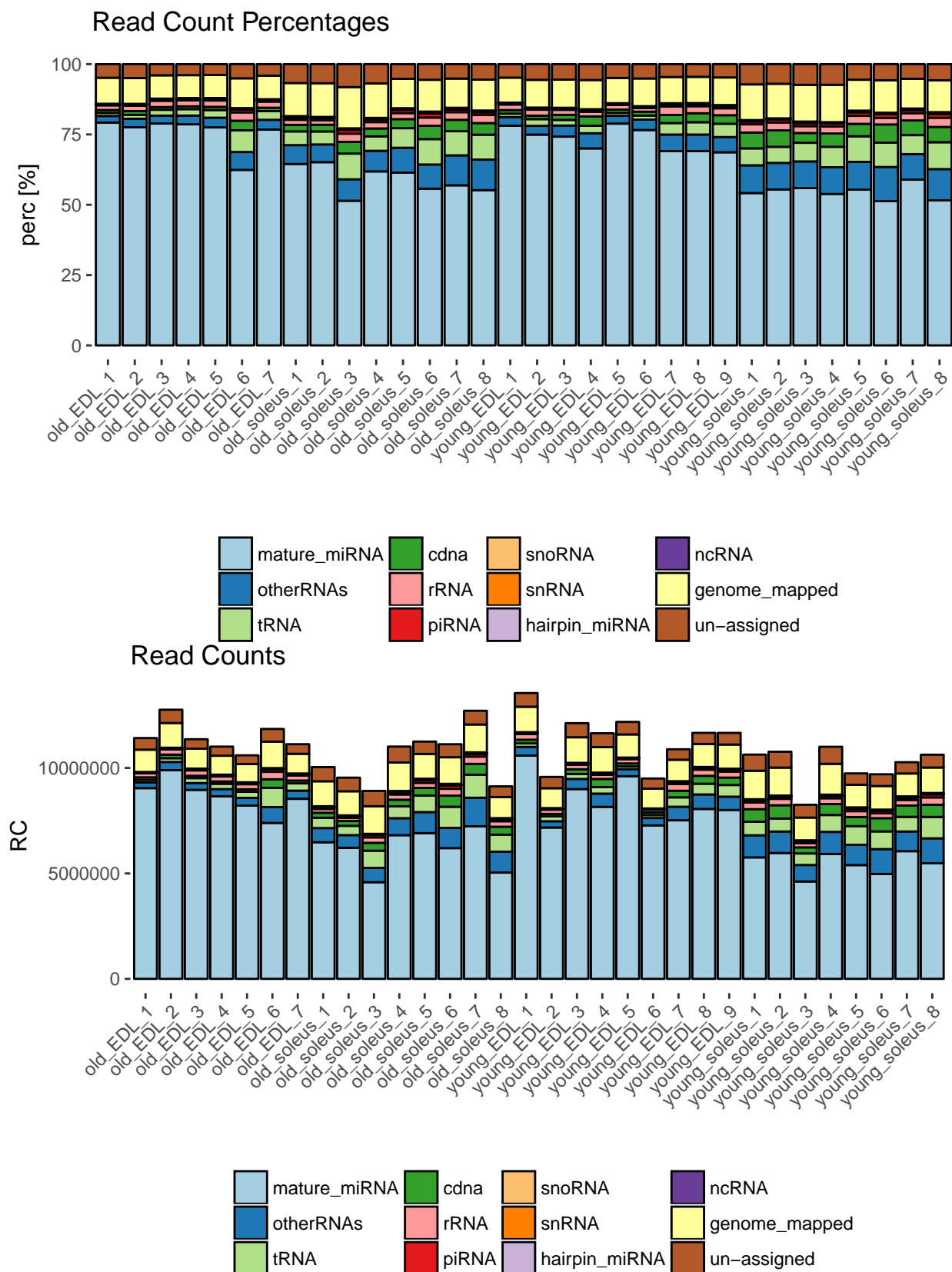


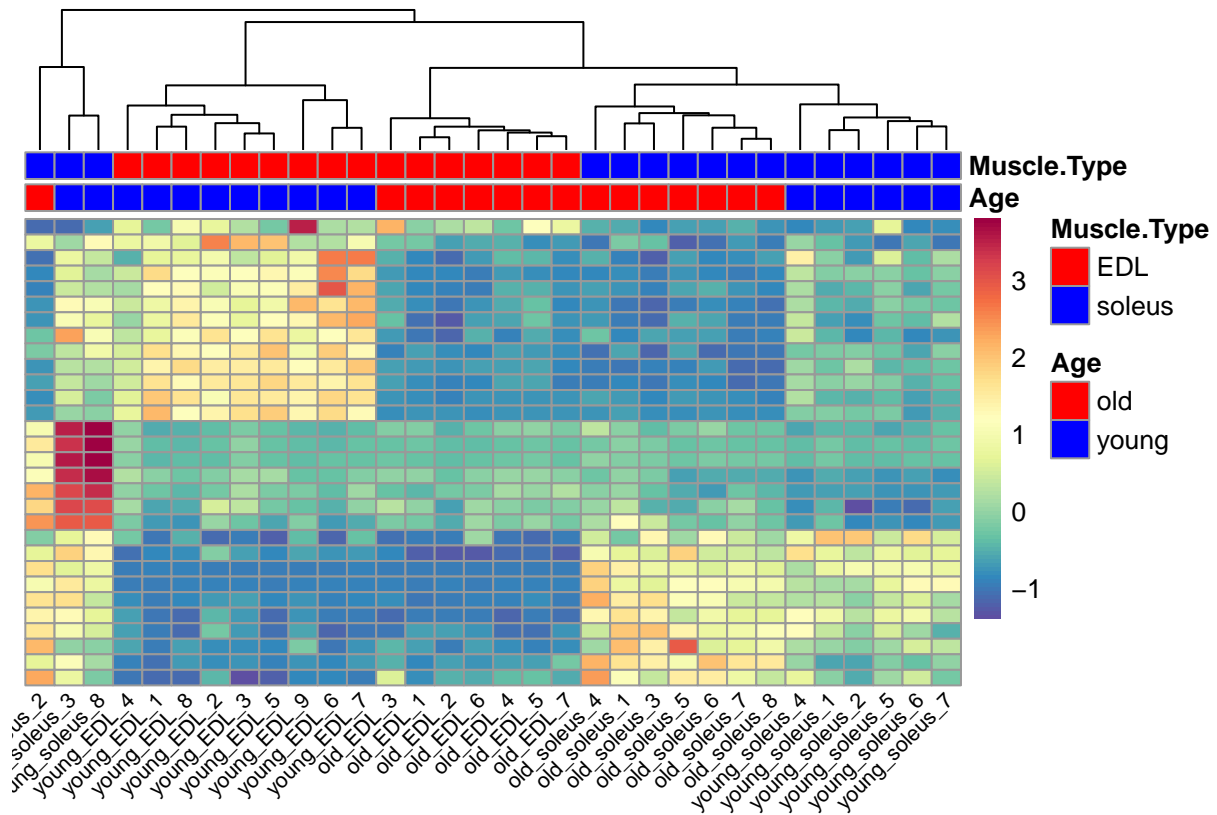
Table 2: Distribution of Reads pt 1

sample	mature_miRNA		otherRNAs		tRNA		cdna		rRNA		piRNA	
	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]
old_EDL												
old_EDL_1	9043776	79.22	268473	2.35	123924	1.09	118603	1.04	188862	1.65	28631	0.25
old_EDL_2	9894661	77.54	389381	3.05	177339	1.39	169185	1.33	244680	1.92	38057	0.30
old_EDL_3	8961635	78.88	321660	2.83	233973	2.06	111871	0.98	252850	2.23	40224	0.35
old_EDL_4	8653880	78.60	343861	3.12	251327	2.28	110520	1.00	235488	2.14	37936	0.34
old_EDL_5	8214390	77.51	365977	3.45	287986	2.72	120306	1.14	238943	2.25	44162	0.42
old_EDL_6	7393959	62.41	752298	6.35	913100	7.71	403828	3.41	351831	2.97	120612	1.02
old_EDL_7	8536474	76.73	386872	3.48	345640	3.11	132769	1.19	242080	2.18	46946	0.42
old_soleus												
old_soleus_1	6477686	64.51	672811	6.70	489278	4.87	235305	2.34	179732	1.79	73640	0.73
old_soleus_2	6214674	65.14	601761	6.31	439589	4.61	228398	2.39	153555	1.61	60913	0.64
old_soleus_3	4582990	51.43	679432	7.62	815467	9.15	371171	4.17	259767	2.91	119997	1.35
old_soleus_4	6812086	61.86	805853	7.32	563853	5.12	311459	2.83	249000	2.26	94502	0.86
old_soleus_5	6910482	61.42	995845	8.85	783507	6.96	362338	3.22	242202	2.15	127404	1.13
old_soleus_6	6199780	55.73	955733	8.59	1004729	9.03	526240	4.73	326032	2.93	157614	1.42
old_soleus_7	7234569	56.92	1353766	10.65	1090620	8.58	514977	4.05	344193	2.71	121581	0.96
old_soleus_8	5038110	55.20	991278	10.86	807373	8.85	376838	4.13	260623	2.86	87853	0.96
young_EDL												
young_EDL_1	10586029	78.13	412272	3.04	165732	1.22	173741	1.28	268101	1.98	38199	0.28
young_EDL_2	7174982	74.94	298512	3.12	229116	2.39	119631	1.25	200655	2.10	35531	0.37
young_EDL_3	8997942	74.23	480090	3.96	239098	1.97	220681	1.82	215392	1.78	41367	0.34
young_EDL_4	8155773	70.04	627916	5.39	311185	2.67	382767	3.29	196166	1.68	58727	0.50
young_EDL_5	9606885	78.86	337780	2.77	128293	1.05	140680	1.15	203282	1.67	30237	0.25
young_EDL_6	7275270	76.58	351837	3.70	138116	1.45	129533	1.36	126740	1.33	24625	0.26
young_EDL_7	7517794	69.07	642748	5.90	444201	4.08	314138	2.89	323657	2.97	73349	0.67
young_EDL_8	8057814	69.09	685815	5.88	509141	4.37	373018	3.20	292464	2.51	77302	0.66
young_EDL_9	8006637	68.66	631745	5.42	551991	4.73	352842	3.03	286792	2.46	79321	0.68
young_soleus												
young_soleus_1	5758487	54.16	1049243	9.87	639804	6.02	598751	5.63	310810	2.92	104562	0.98
young_soleus_2	5968368	55.43	1018806	9.46	615535	5.72	628585	5.84	303116	2.82	108299	1.01
young_soleus_3	4617133	55.93	782258	9.48	547359	6.63	287306	3.48	200214	2.43	88630	1.07
young_soleus_4	5923612	53.84	1045153	9.50	800047	7.27	520698	4.73	273970	2.49	114546	1.04
young_soleus_5	5394313	55.38	960838	9.86	887392	9.11	430607	4.42	277862	2.85	118562	1.22
young_soleus_6	4974543	51.28	1178167	12.15	837927	8.64	628225	6.48	228165	2.35	122411	1.26
young_soleus_7	6053553	58.93	933164	9.08	694491	6.76	535201	5.21	264642	2.58	106477	1.04
young_soleus_8	5478574	51.55	1183075	11.13	1015384	9.55	577679	5.44	347159	3.27	142788	1.34

Table 3: Distribution of Reads pt 2

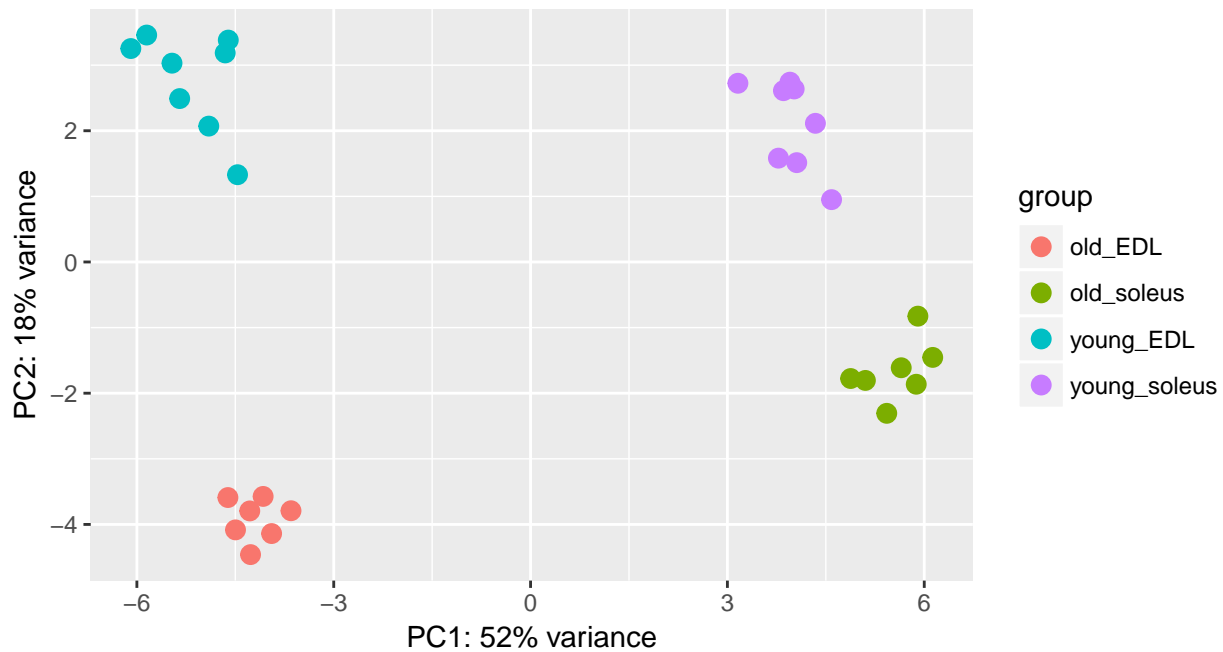
sample	snoRNA		snRNA		hairpin_miRNA		ncRNA		genome_mapped		un-assigned	
	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]	[#]	[%]
old_EDL												
old_EDL_1	21033	0.18	2127	0.02	12830	0.11	1200	0.01	1056760	9.26	550416	4.82
old_EDL_2	25861	0.20	3298	0.03	15911	0.12	1750	0.01	1171189	9.18	628999	4.93
old_EDL_3	31010	0.27	1552	0.01	14525	0.13	1495	0.01	939957	8.27	450400	3.96
old_EDL_4	31474	0.29	1702	0.02	15533	0.14	1406	0.01	896075	8.14	431517	3.92
old_EDL_5	31629	0.30	1889	0.02	13886	0.13	1530	0.01	868315	8.19	408844	3.86
old_EDL_6	33612	0.28	8046	0.07	14834	0.13	2316	0.02	1253660	10.58	599674	5.06
old_EDL_7	29245	0.26	1885	0.02	14693	0.13	1398	0.01	932427	8.38	454979	4.09
old_soleus												
old_soleus_1	36803	0.37	7308	0.07	11772	0.12	1531	0.02	1182255	11.77	673867	6.71
old_soleus_2	32412	0.34	6682	0.07	11336	0.12	1363	0.01	1143099	11.98	646130	6.77
old_soleus_3	28655	0.32	8991	0.10	9958	0.11	1731	0.02	1305917	14.65	727523	8.16
old_soleus_4	44754	0.41	10542	0.10	13513	0.12	2317	0.02	1348329	12.24	756481	6.87
old_soleus_5	44086	0.39	9068	0.08	13668	0.12	2123	0.02	1173094	10.43	586567	5.21
old_soleus_6	49024	0.44	14663	0.13	12096	0.11	2310	0.02	1257816	11.31	618784	5.56
old_soleus_7	48178	0.38	13214	0.10	15803	0.12	2570	0.02	1319233	10.38	651929	5.13
old_soleus_8	36295	0.40	9508	0.10	12295	0.13	1940	0.02	1001504	10.97	503666	5.52
young_EDL												
young_EDL_1	28470	0.21	3245	0.02	17720	0.13	2331	0.02	1205223	8.90	648257	4.78
young_EDL_2	20738	0.22	2399	0.03	12602	0.13	1923	0.02	942461	9.84	536056	5.60
young_EDL_3	25389	0.21	3768	0.03	15454	0.13	2206	0.02	1211699	10.00	668656	5.52
young_EDL_4	23902	0.21	8315	0.07	11649	0.10	1973	0.02	1209556	10.39	655864	5.63
young_EDL_5	25604	0.21	2785	0.02	14821	0.12	1683	0.01	1095583	8.99	595062	4.88
young_EDL_6	20476	0.22	1522	0.02	10943	0.12	1260	0.01	936670	9.86	483408	5.09
young_EDL_7	32499	0.30	7583	0.07	14149	0.13	2317	0.02	1015883	9.33	496497	4.56
young_EDL_8	30339	0.26	5278	0.05	14305	0.12	2602	0.02	1088379	9.33	526226	4.51
young_EDL_9	32547	0.28	5759	0.05	15382	0.13	2218	0.02	1142904	9.80	553237	4.74
young_soleus												
young_soleus_1	29744	0.28	11076	0.10	12586	0.12	2692	0.03	1349290	12.69	766177	7.21
young_soleus_2	24567	0.23	9697	0.09	12218	0.11	2625	0.02	1323873	12.29	752103	6.98
young_soleus_3	28943	0.35	6354	0.08	9732	0.12	1746	0.02	1077741	13.06	607122	7.36
young_soleus_4	28951	0.26	11039	0.10	13416	0.12	2561	0.02	1461277	13.28	807839	7.34
young_soleus_5	37791	0.39	8989	0.09	12396	0.13	2383	0.02	1070192	10.99	539752	5.54
young_soleus_6	28194	0.29	14128	0.15	10753	0.11	2436	0.03	1119059	11.54	556447	5.74
young_soleus_7	35533	0.35	13516	0.13	13095	0.13	2542	0.02	1084745	10.56	535987	5.22
young_soleus_8	40163	0.38	13058	0.12	12801	0.12	3111	0.03	1205401	11.34	607726	5.72

4.2 30 most variant microRNAs



Values for the top 30 microRNA with the highest variation, measured as the coefficient of variation, are shown (rows). Columns are the samples of the experiment. This figure gives an overview of whether the known factors have an impact on the most variable microRNA in the dataset.

5 Exploratory data analysis: principal component analysis (PCA)



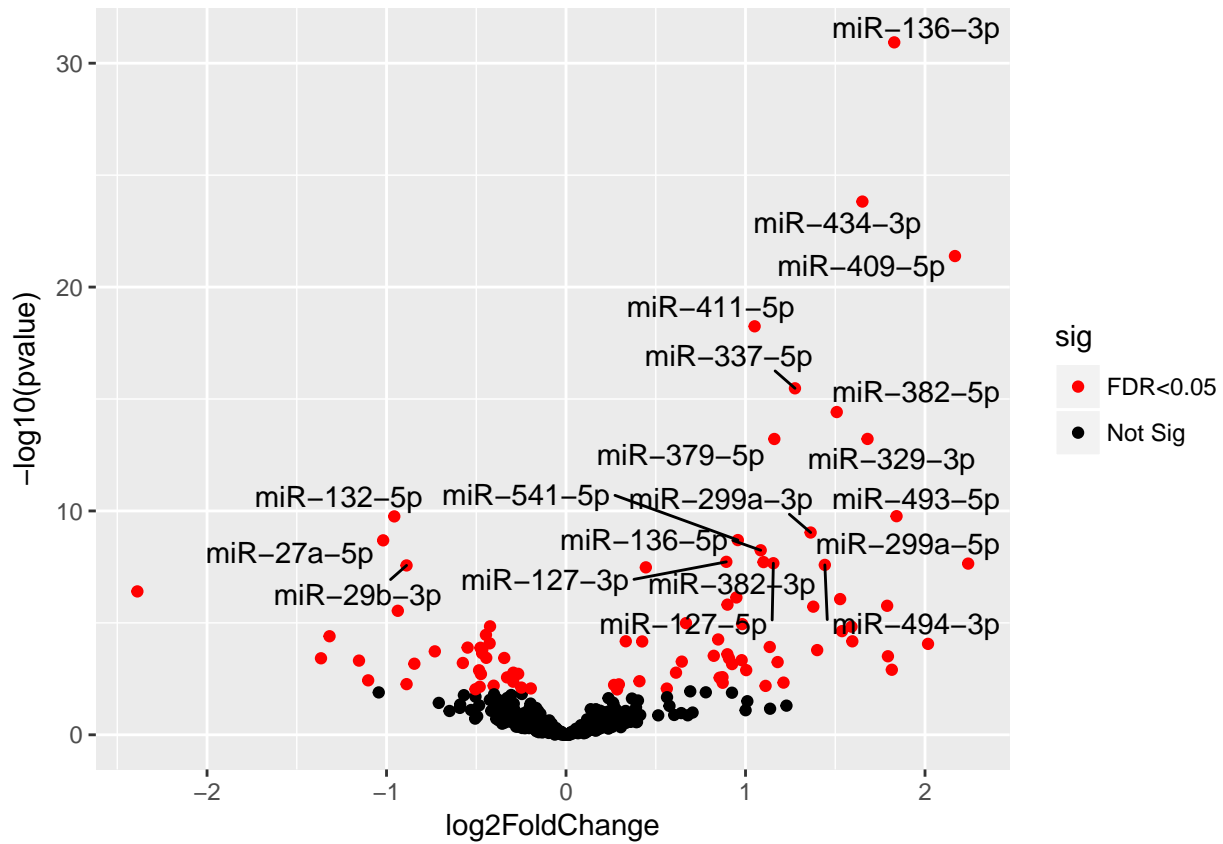
The above plot shows the first two principal components that explain the variability in the data using the regularized log count data. This analysis is based on 415 microRNA. If you are unfamiliar with principal component analysis, you might want to check the Wikipedia entry or this interactive explanation. In this case, the first and second principal component explain 52 and 18 percent of the variance respectively.

6 Differential Expression Analysis

Note that according to the used convention throughout this chapter always the second given group serves as the reference. E.g. for a given comparison of A vs. B a $\log_2\text{FoldChange} > 0$ indicates that the respective microRNA is upregulated in group A .

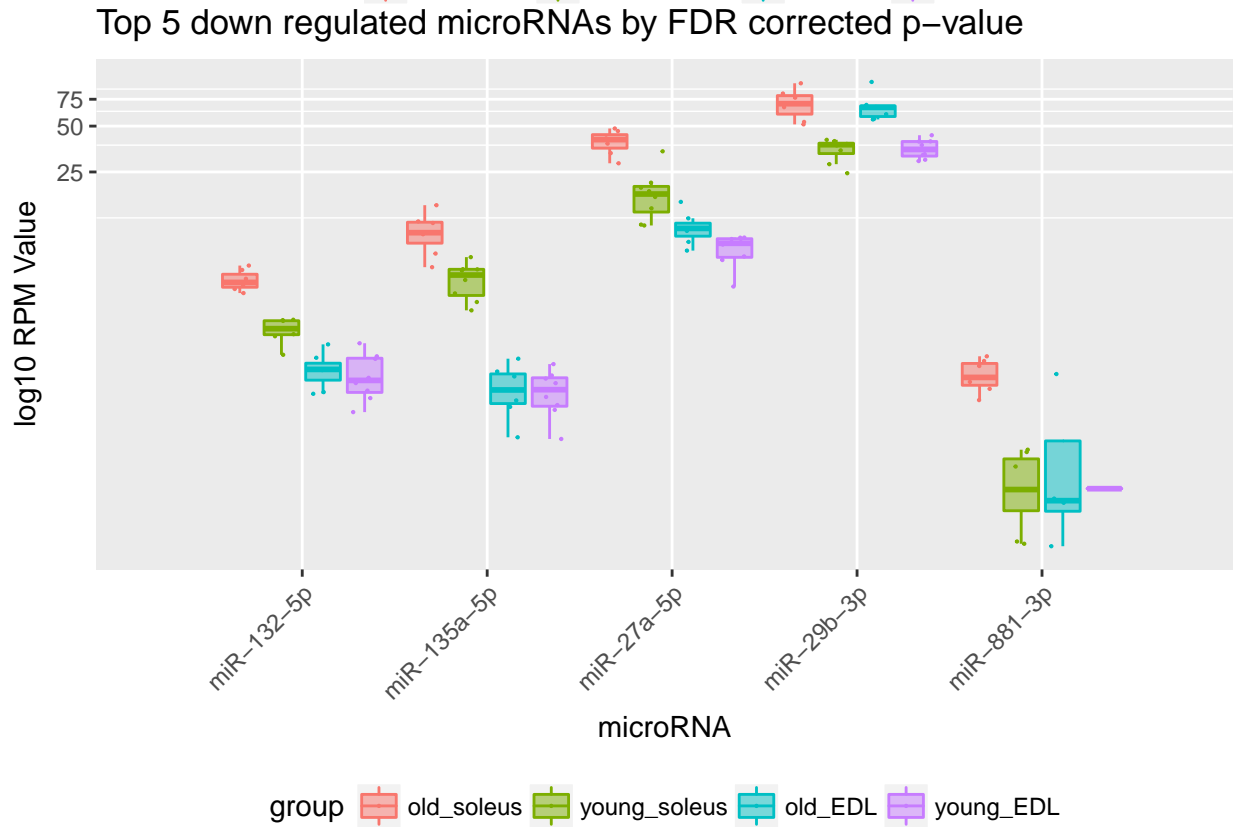
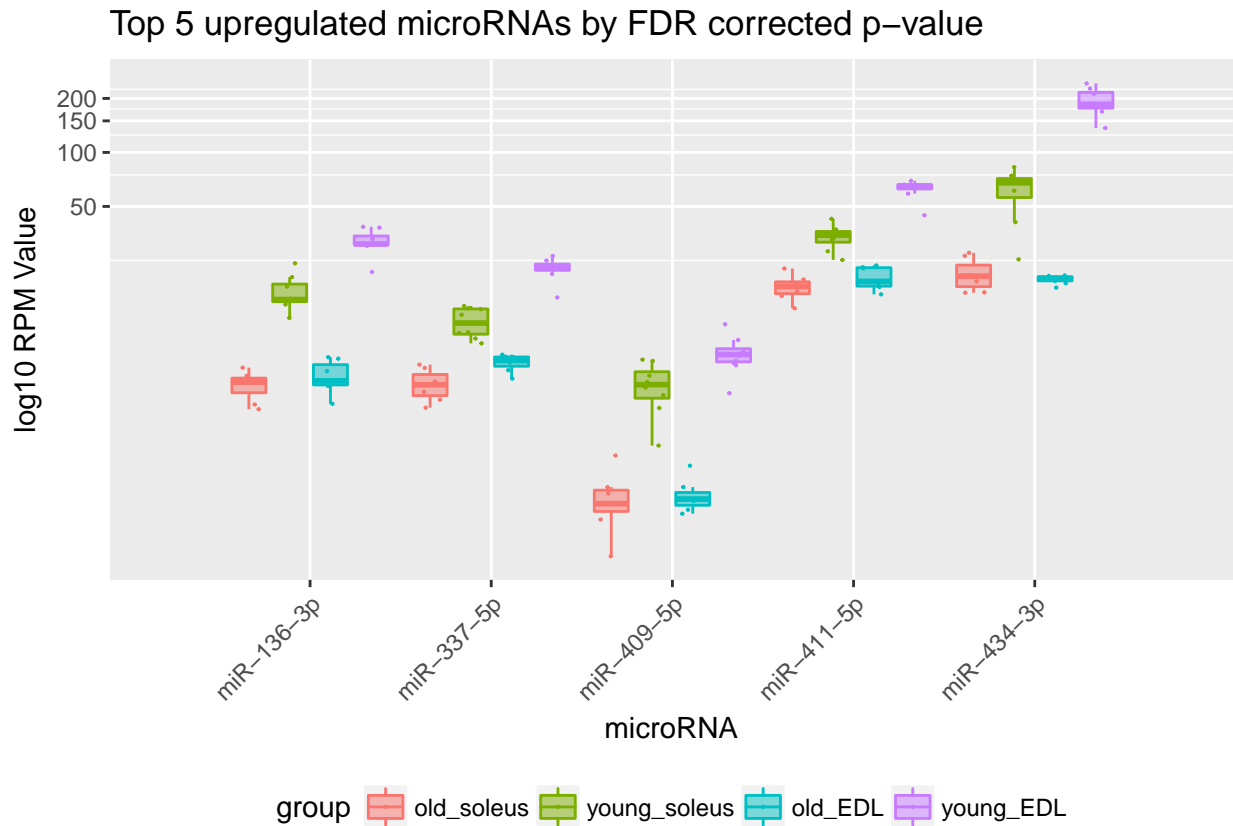
6.1 young soleus vs. old soleus

6.1.1 Volcano Plot for young soleus vs. old soleus

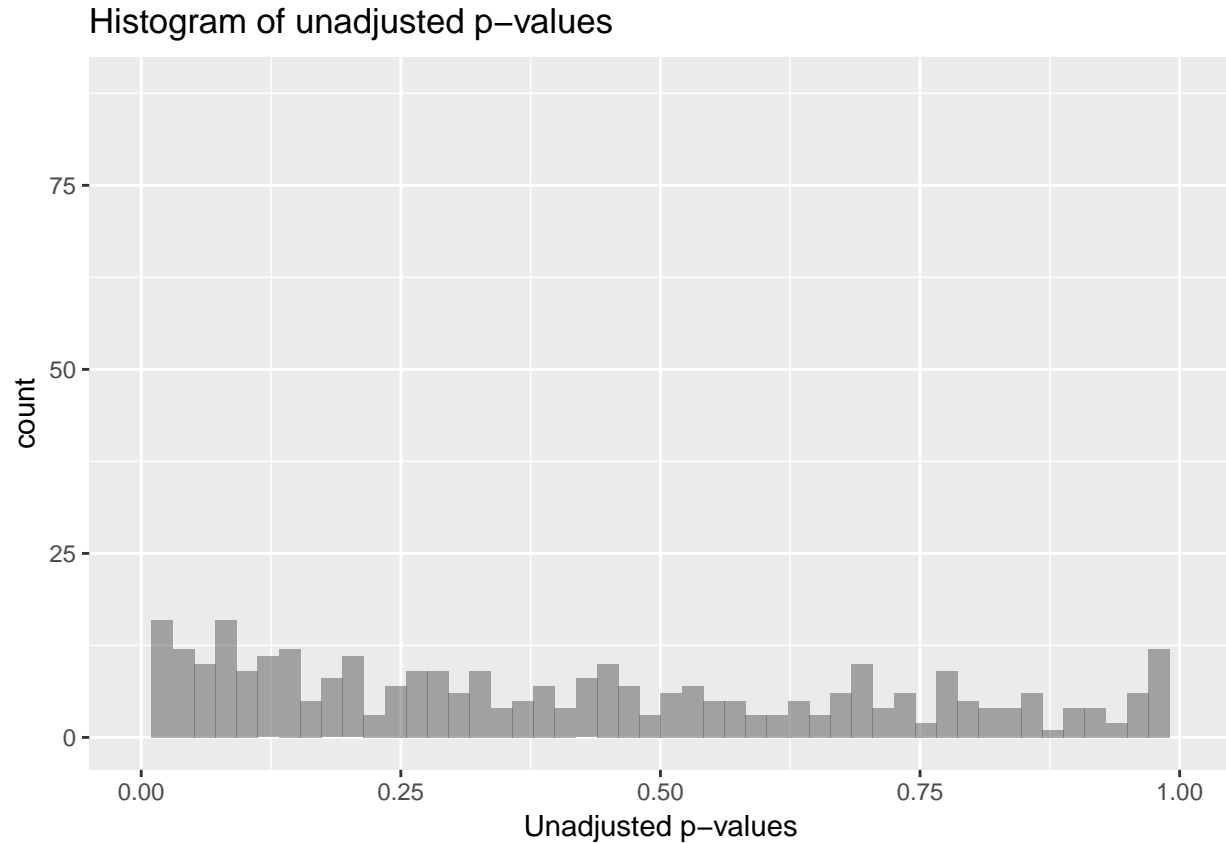


6.1.2 Count Plot for Top 5 microRNAs by FDR corrected p-value for young soleus vs. old soleus

Note that the following plots show RPM values and not the count data used internally in edgeR after library normalization.



6.1.3 P-values distribution for young soleus vs. old soleus



This plot shows a histogram of the unadjusted p-values. It might be skewed right or left, or flat as shown in the Wikipedia examples. The shape depends on the percent of features that are differentially expressed. For further information on how to interpret a histogram of p-values check David Robinson's post on this topic.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.03049 0.25483 0.33458 0.57946 1.00000
```

This is the numerical summary of the distribution of the p-values.

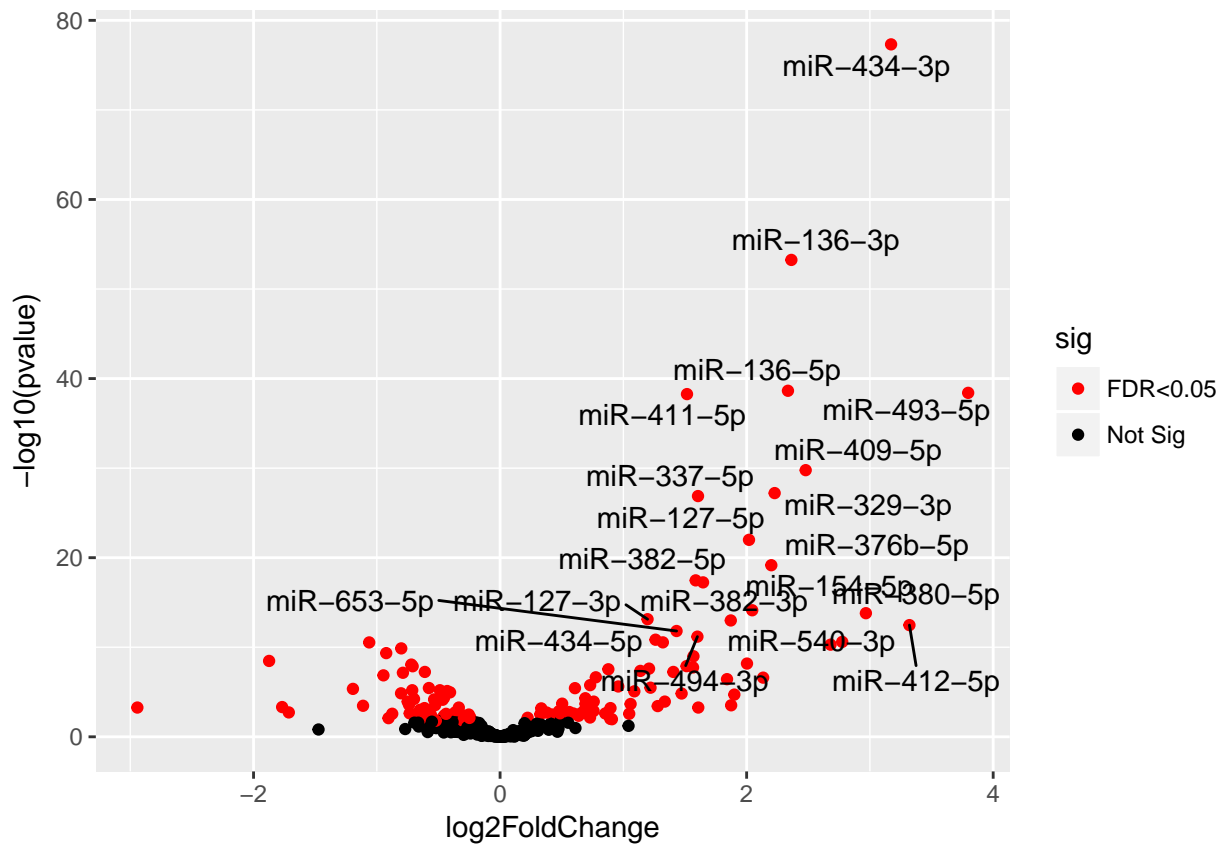
Cut	Count
0.0001	41
0.0010	62
0.0100	88
0.0250	101
0.0500	115
0.1000	146
0.2000	188
0.3000	227
0.4000	258
0.5000	289
0.6000	315
0.7000	339
0.8000	366
0.9000	385
1.0000	415

This table shows the number of features with p-values less or equal than some commonly used cutoff values.

Feature	log2FoldChange	baseMean	pvalue	padj
miR-136-3p	1.83	4.59	1.171342e-31	4.861069e-29
miR-434-3p	1.65	6.91	1.514048e-24	3.141650e-22
miR-409-5p	2.17	2.66	4.110463e-22	5.686141e-20
miR-411-5p	1.05	5.74	5.659017e-19	5.871230e-17
miR-337-5p	1.28	4.22	3.327530e-16	2.761850e-14
miR-382-5p	1.51	3.61	3.822319e-15	2.643770e-13
miR-329-3p	1.68	2.81	6.080008e-14	3.166965e-12
miR-379-5p	1.16	6.75	6.104993e-14	3.166965e-12
miR-493-5p	1.84	2.29	1.716425e-10	7.275605e-09
miR-132-5p	-0.96	1.93	1.753158e-10	7.275605e-09
miR-299a-3p	1.36	3.34	9.208987e-10	3.474300e-08
miR-27a-5p	-1.02	4.82	2.056166e-09	6.563914e-08
miR-541-5p	0.96	5.79	2.002391e-09	6.563914e-08
miR-136-5p	1.08	3.47	5.717112e-09	1.694715e-07
miR-382-3p	1.10	3.64	1.932626e-08	5.012749e-07
miR-127-3p	0.89	8.96	1.878877e-08	5.012749e-07
miR-299a-5p	2.24	0.49	2.250750e-08	5.189230e-07
miR-127-5p	1.16	4.03	2.168445e-08	5.189230e-07
miR-494-3p	1.44	1.71	2.575862e-08	5.626226e-07
miR-29b-3p	-0.89	6.32	2.716394e-08	5.636518e-07

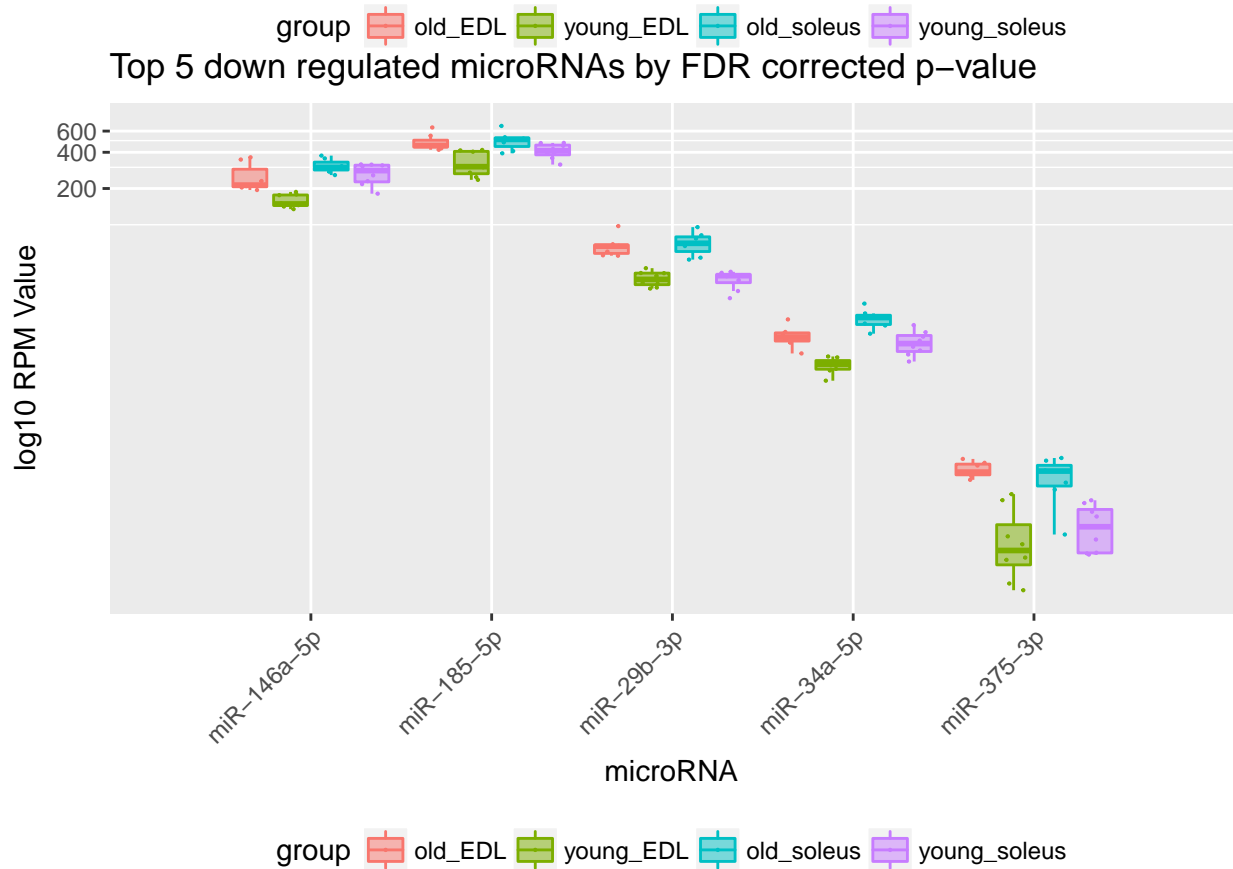
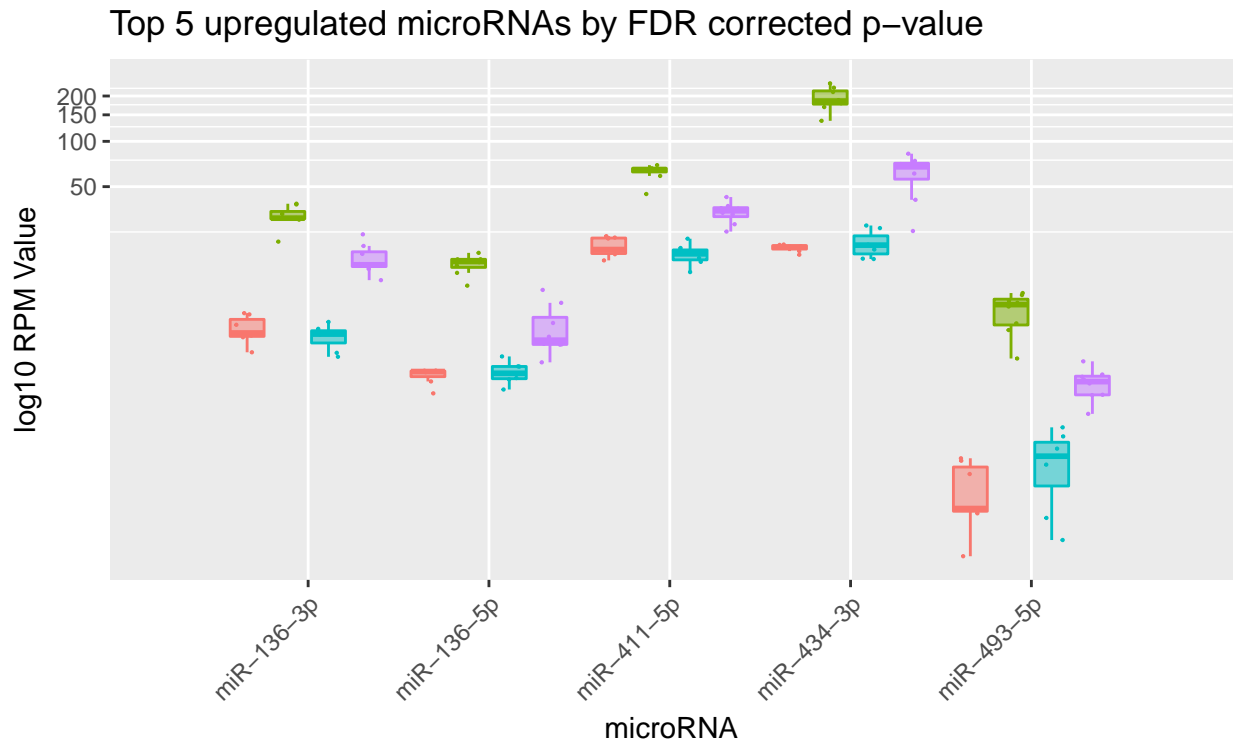
6.2 young EDL vs. old EDL

6.2.1 Volcano Plot for young EDL vs. old EDL



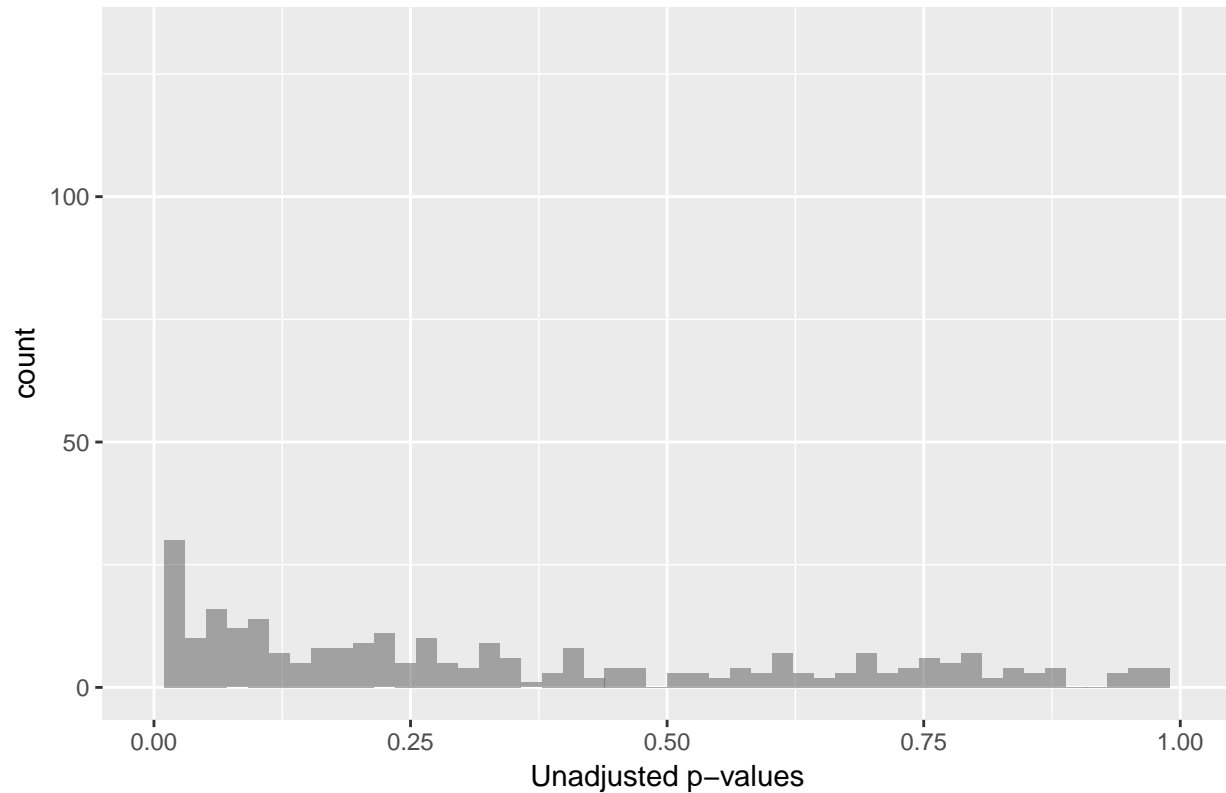
6.2.2 Count Plot for Top 5 microRNAs by FDR corrected p-value for young EDL vs. old EDL

Note that the following plots show RPM values and not the count data used internally in edgeR after library normalization.



6.2.3 P-values distribution for young EDL vs. old EDL

Histogram of unadjusted p-values



This plot shows a histogram of the unadjusted p-values. It might be skewed right or left, or flat as shown in the Wikipedia examples. The shape depends on the percent of features that are differentially expressed. For further information on how to interpret a histogram of p-values check David Robinson's post on this topic.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00324 0.10065 0.24926 0.41235 1.00000
```

This is the numerical summary of the distribution of the p-values.

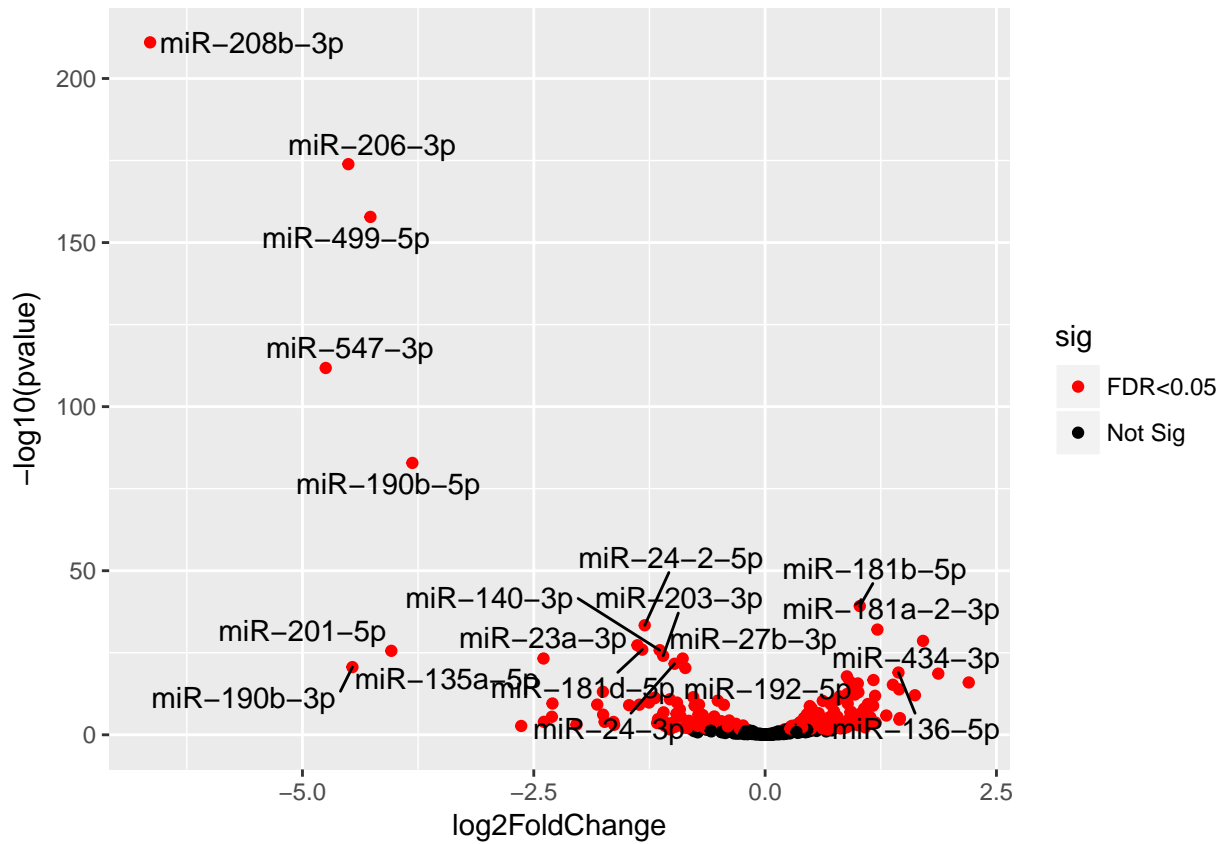
Cut	Count
0.0001	63
0.0010	83
0.0100	132
0.0250	156
0.0500	170
0.1000	207
0.2000	246
0.3000	284
0.4000	305
0.5000	323
0.6000	337
0.7000	359
0.8000	383
0.9000	398
1.0000	415

This table shows the number of features with p-values less or equal than some commonly used cutoff values.

Feature	log2FoldChange	baseMean	pvalue	padj
miR-434-3p	3.17	6.91	4.795327e-78	1.990061e-75
miR-136-3p	2.36	4.59	5.607464e-54	1.163549e-51
miR-136-5p	2.34	3.47	2.296078e-39	3.176241e-37
miR-493-5p	3.80	2.29	3.914417e-39	4.061208e-37
miR-411-5p	1.52	5.74	5.424942e-39	4.502702e-37
miR-409-5p	2.48	2.66	1.683458e-30	1.164392e-28
miR-329-3p	2.23	2.81	6.122590e-28	3.629821e-26
miR-337-5p	1.61	4.22	1.315323e-27	6.823236e-26
miR-127-5p	2.02	4.03	1.011954e-22	4.666232e-21
miR-376b-5p	2.20	1.59	6.943606e-20	2.881597e-18
miR-382-5p	1.59	3.61	3.416791e-18	1.289062e-16
miR-382-3p	1.65	3.64	5.707100e-18	1.973705e-16
miR-154-5p	2.05	1.08	7.413593e-15	2.366647e-13
miR-380-5p	2.97	0.69	1.578527e-14	4.679204e-13
miR-127-3p	1.20	8.96	7.390020e-14	2.044572e-12
miR-540-3p	1.87	1.28	1.018712e-13	2.642284e-12
miR-412-5p	3.32	-0.20	3.377884e-13	8.246010e-12
miR-653-5p	1.43	2.99	1.540578e-12	3.551888e-11
miR-494-3p	1.60	1.71	6.547912e-12	1.430202e-10
miR-434-5p	1.26	5.16	1.438757e-11	2.985420e-10

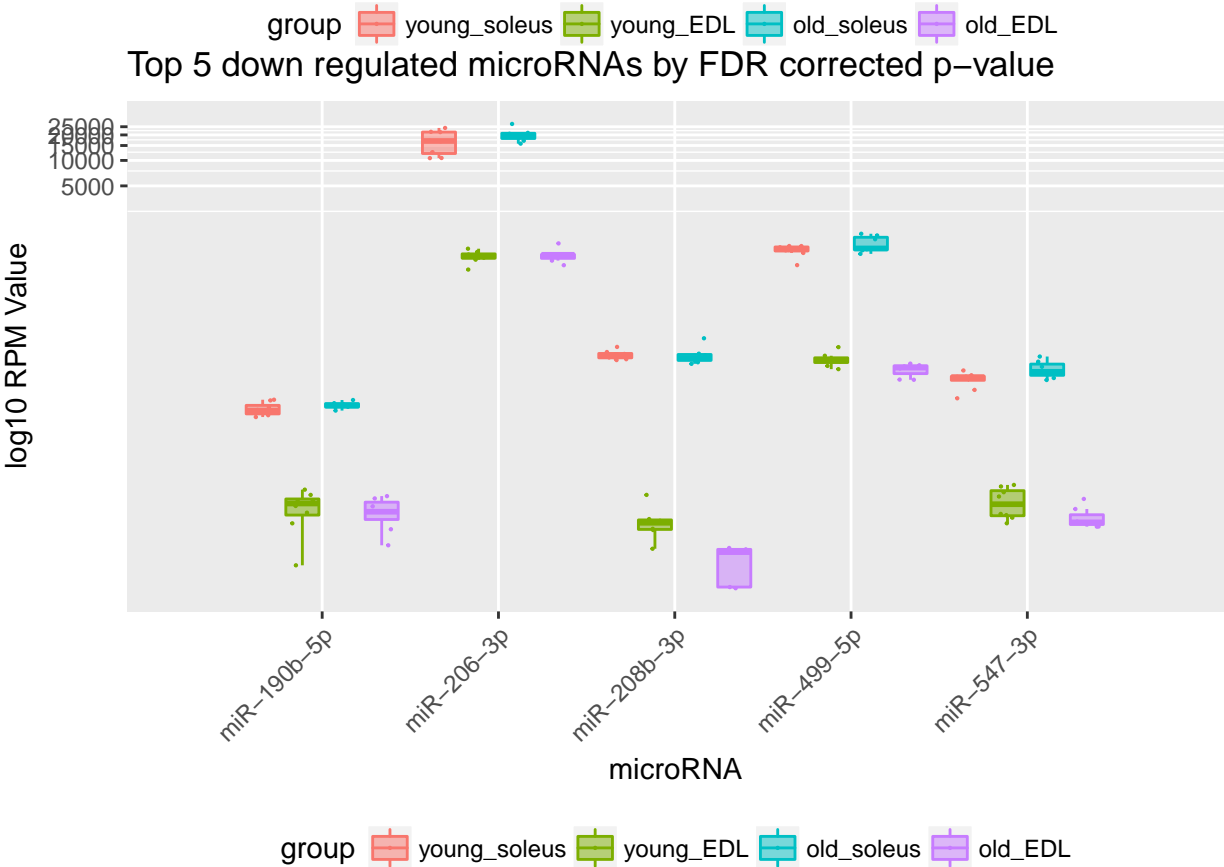
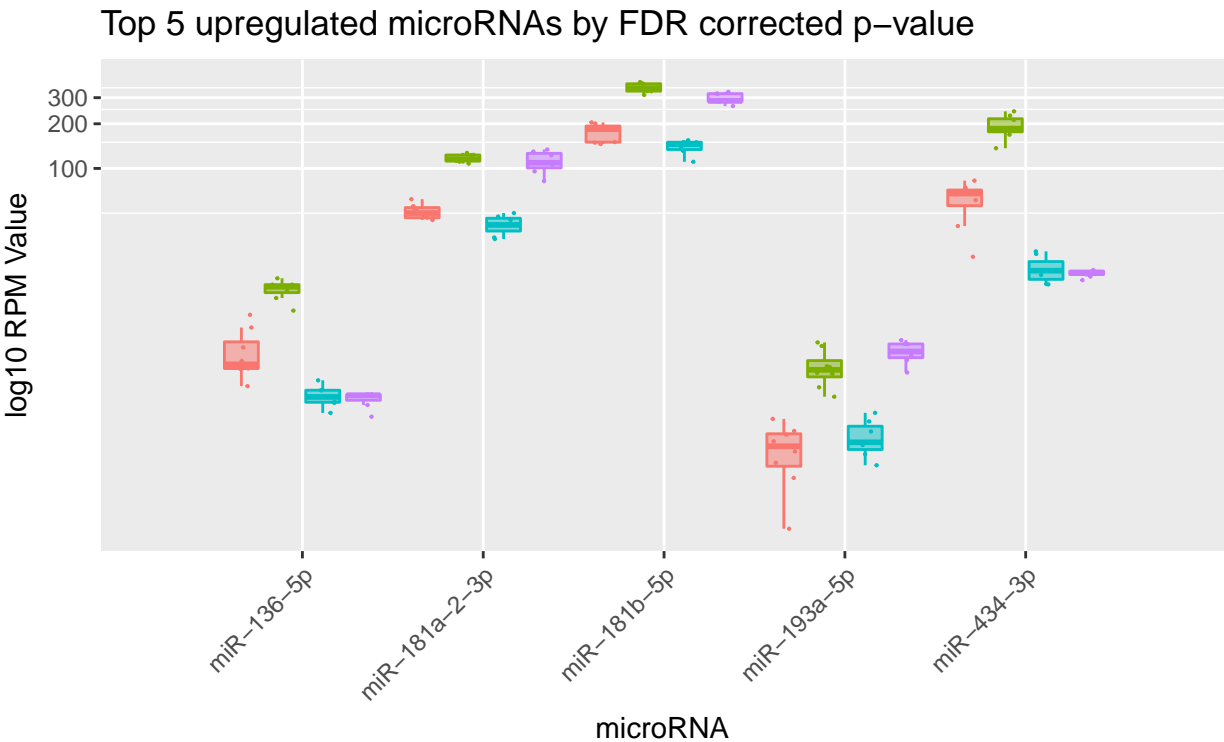
6.3 young EDL vs. young soleus

6.3.1 Volcano Plot for young EDL vs. young soleus

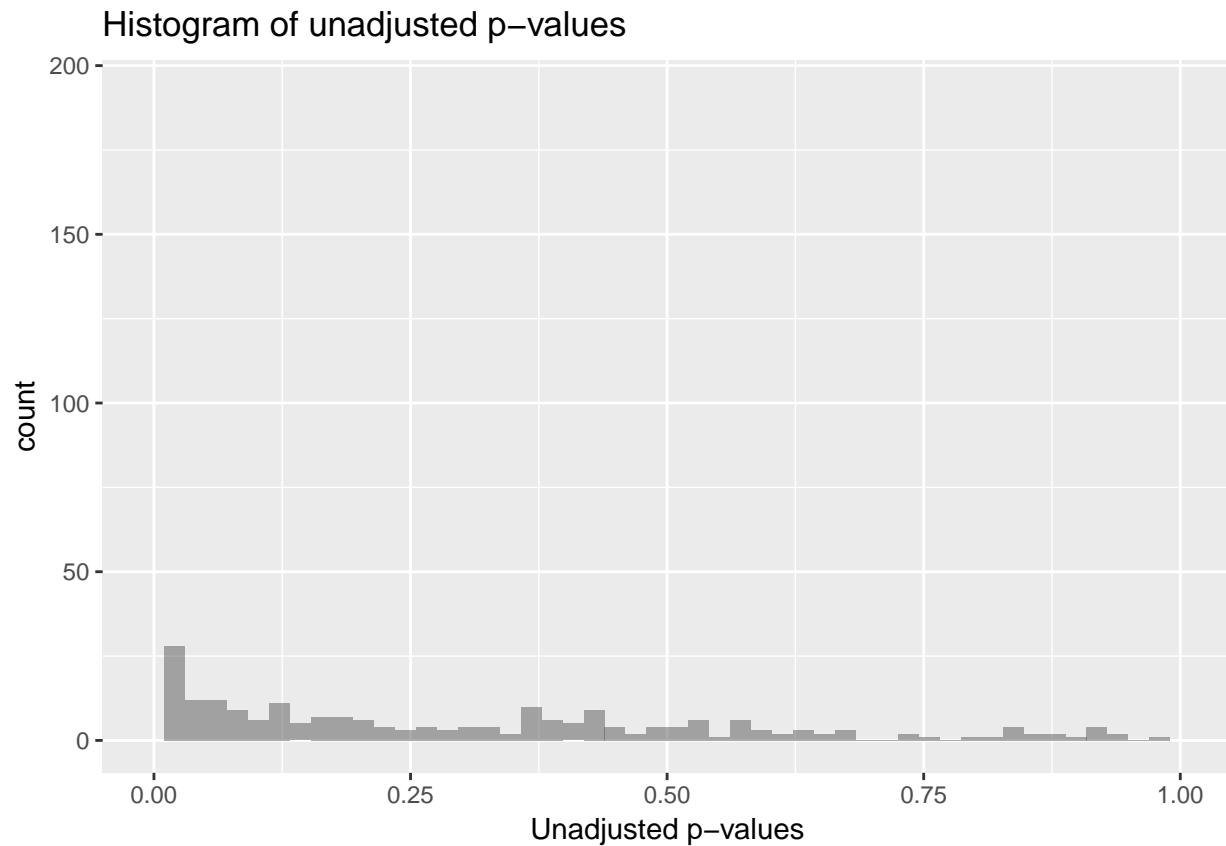


6.3.2 Count Plot for Top 5 microRNAs by FDR corrected p-value for young EDL vs. young soleus

Note that the following plots show RPM values and not the count data used internally in edgeR after library normalization.



6.3.3 P-values distribution for young EDL vs. young soleus



This plot shows a histogram of the unadjusted p-values. It might be skewed right or left, or flat as shown in the Wikipedia examples. The shape depends on the percent of features that are differentially expressed. For further information on how to interpret a histogram of p-values check David Robinson's post on this topic.

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## 0.0000000 0.0000379 0.0214147 0.1755007 0.3081942 1.0000000
```

This is the numerical summary of the distribution of the p-values.

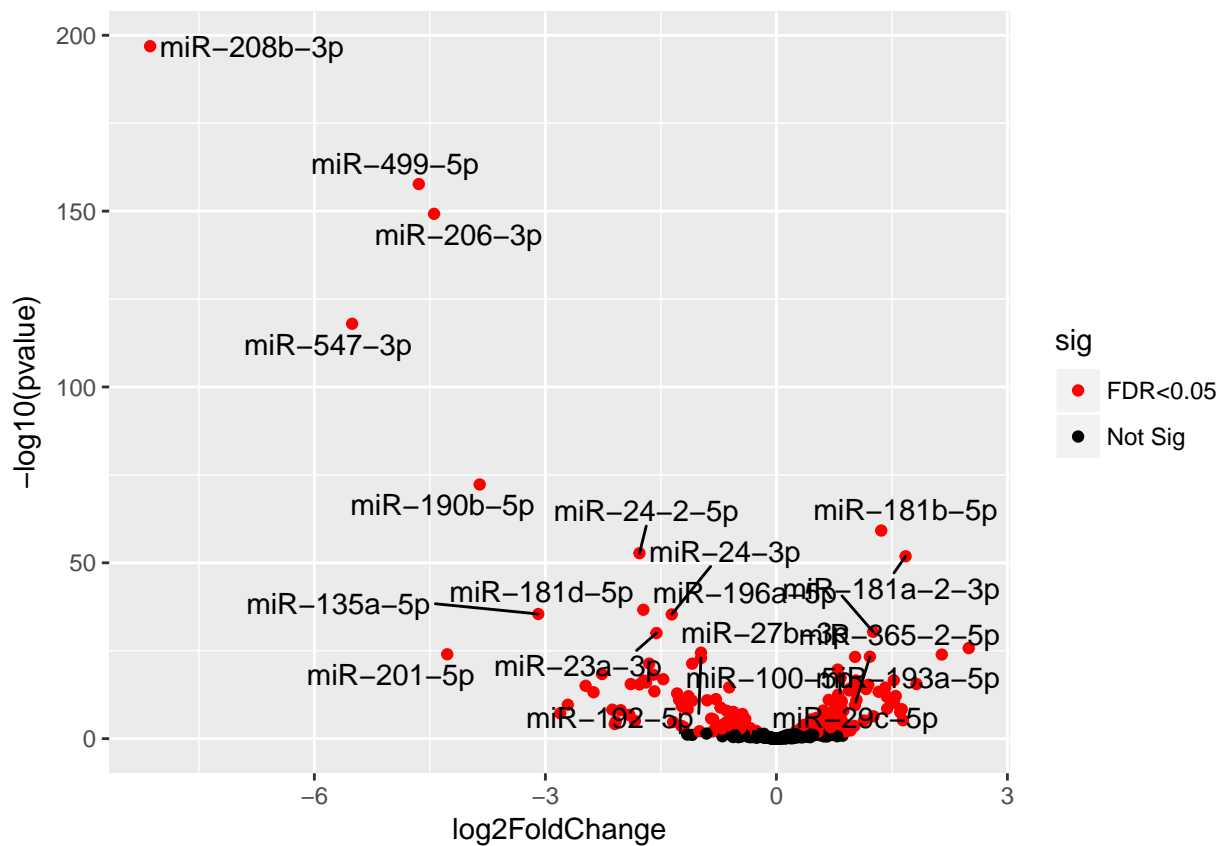
Cut	Count
0.0001	112
0.0010	149
0.0100	191
0.0250	213
0.0500	232
0.1000	257
0.2000	291
0.3000	309
0.4000	336
0.5000	359
0.6000	379
0.7000	389
0.8000	393
0.9000	403
1.0000	415

This table shows the number of features with p-values less or equal than some commonly used cutoff values.

Feature	log2FoldChange	baseMean	pvalue	padj
miR-208b-3p	-6.65	5.23	9.749828e-212	4.046179e-209
miR-206-3p	-4.50	13.77	1.273643e-174	2.642809e-172
miR-499-5p	-4.27	9.51	1.504543e-158	2.081285e-156
miR-547-3p	-4.75	4.50	1.644473e-112	1.706141e-110
miR-190b-5p	-3.81	3.29	1.428000e-83	1.185240e-81
miR-181b-5p	1.02	8.57	5.987326e-40	4.141234e-38
miR-24-2-5p	-1.30	8.15	4.187526e-34	2.482605e-32
miR-181a-2-3p	1.21	6.99	8.534098e-33	4.427063e-31
miR-434-3p	1.71	6.91	2.385994e-29	1.100208e-27
miR-23a-3p	-1.38	11.08	5.127282e-28	2.127822e-26
miR-181d-5p	-1.33	6.36	1.138988e-26	4.297091e-25
miR-140-3p	-1.14	9.66	1.643572e-26	5.684020e-25
miR-201-5p	-4.04	0.56	2.565246e-26	8.189056e-25
miR-203-3p	-1.10	6.39	8.358808e-25	2.477790e-23
miR-27b-3p	-0.89	14.11	5.732908e-24	1.486973e-22
miR-135a-5p	-2.39	2.68	5.588304e-24	1.486973e-22
miR-24-3p	-0.98	11.26	2.253563e-22	5.501344e-21
miR-190b-3p	-4.46	0.00	2.317091e-21	5.342182e-20
miR-192-5p	-0.86	8.02	4.412299e-21	9.637390e-20
miR-136-5p	1.44	3.47	1.057288e-19	2.193874e-18

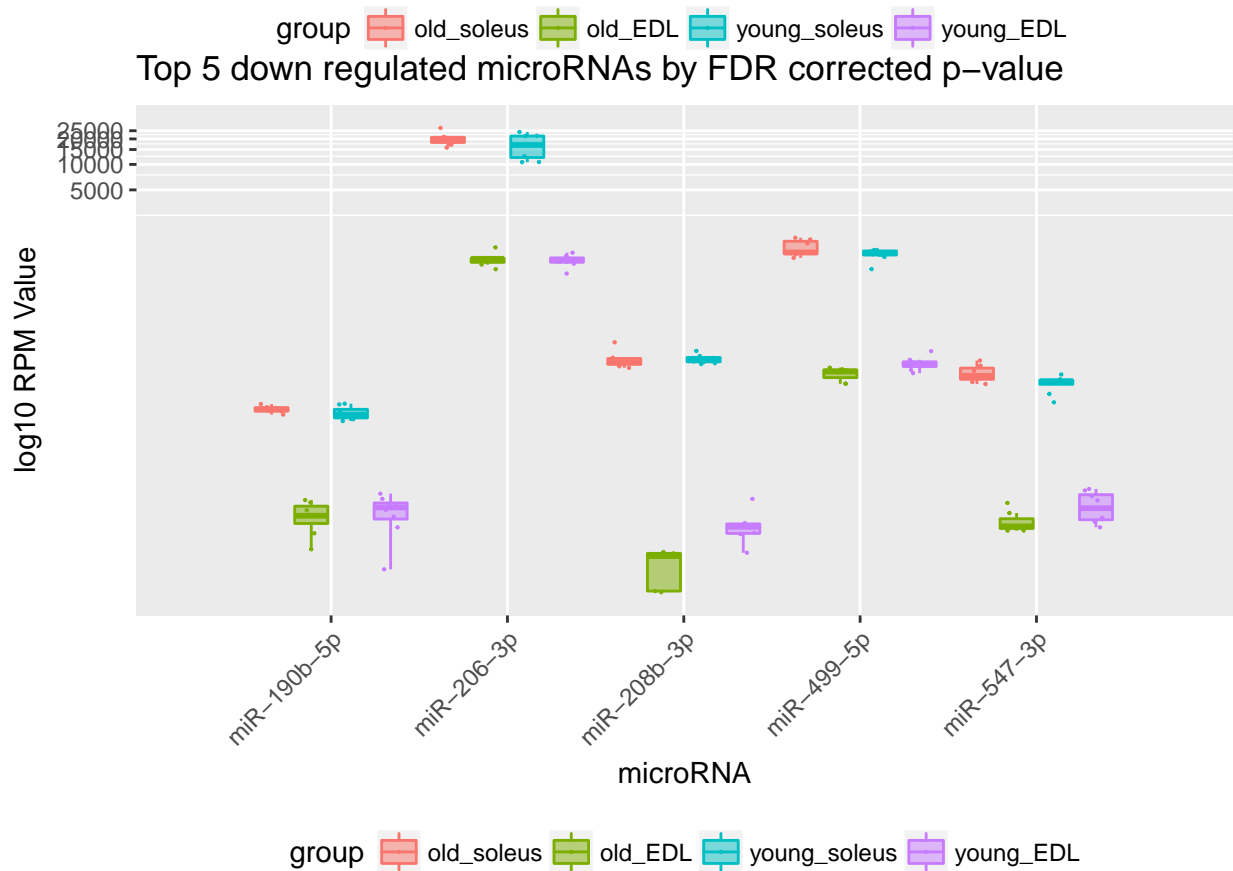
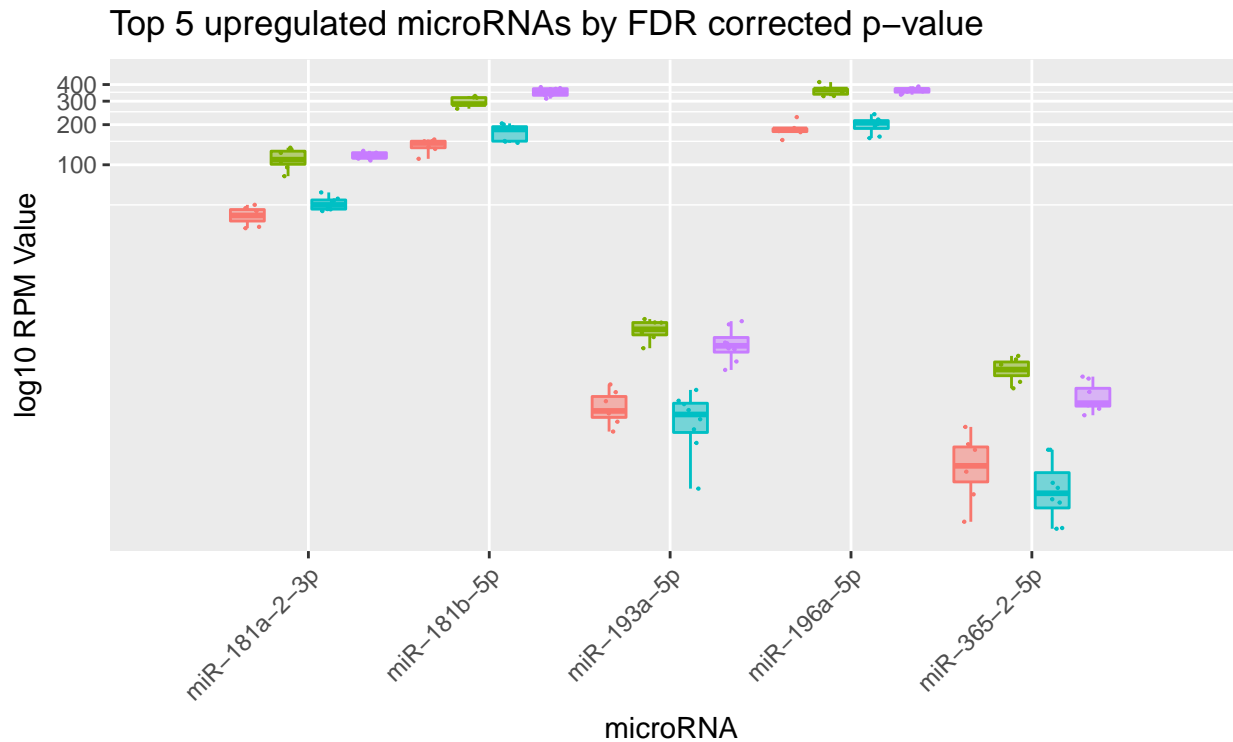
6.4 old EDL vs. old soleus

6.4.1 Volcano Plot for old EDL vs. old soleus



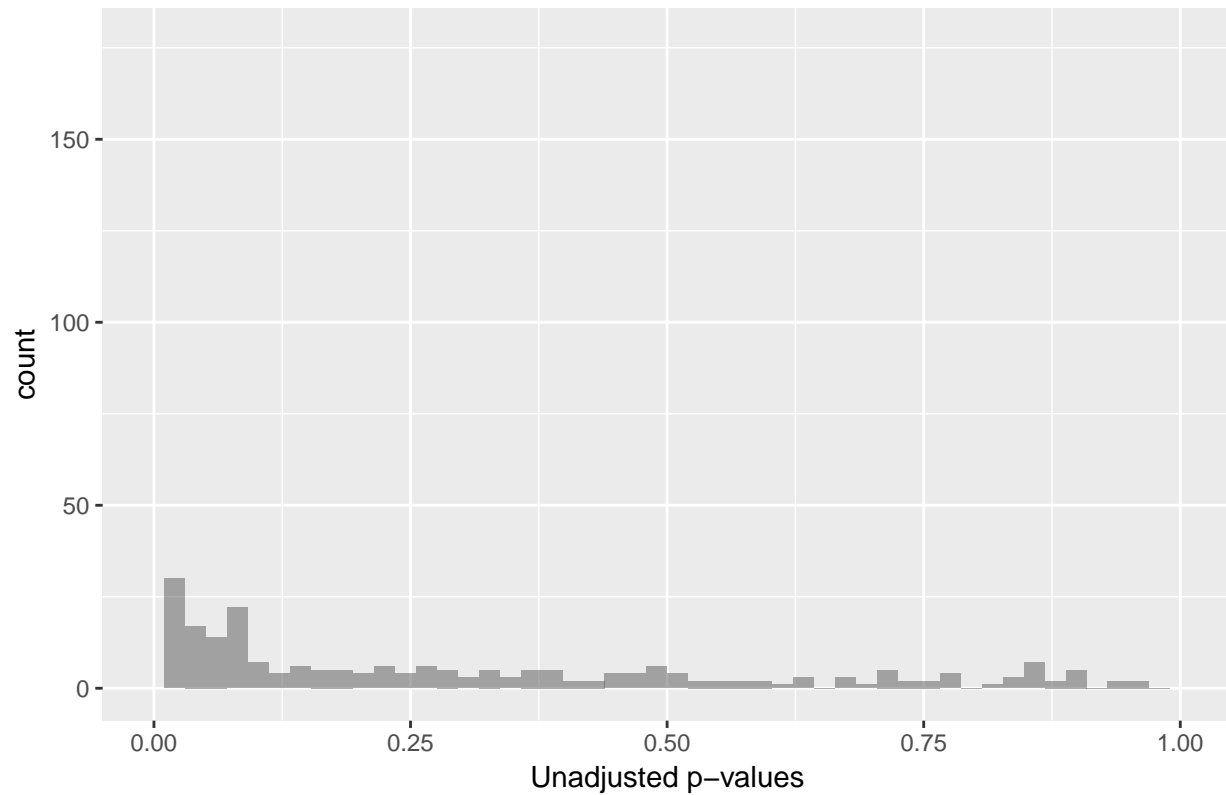
6.4.2 Count Plot for Top 5 microRNAs by FDR corrected p-value for old EDL vs. old soleus

Note that the following plots show RPM values and not the count data used internally in edgeR after library normalization.



6.4.3 P-values distribution for old EDL vs. old soleus

Histogram of unadjusted p-values



This plot shows a histogram of the unadjusted p-values. It might be skewed right or left, or flat as shown in the Wikipedia examples. The shape depends on the percent of features that are differentially expressed. For further information on how to interpret a histogram of p-values check David Robinson's post on this topic.

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## 0.0000000 0.0000329 0.0319179 0.1911677 0.2878284 1.0000000
```

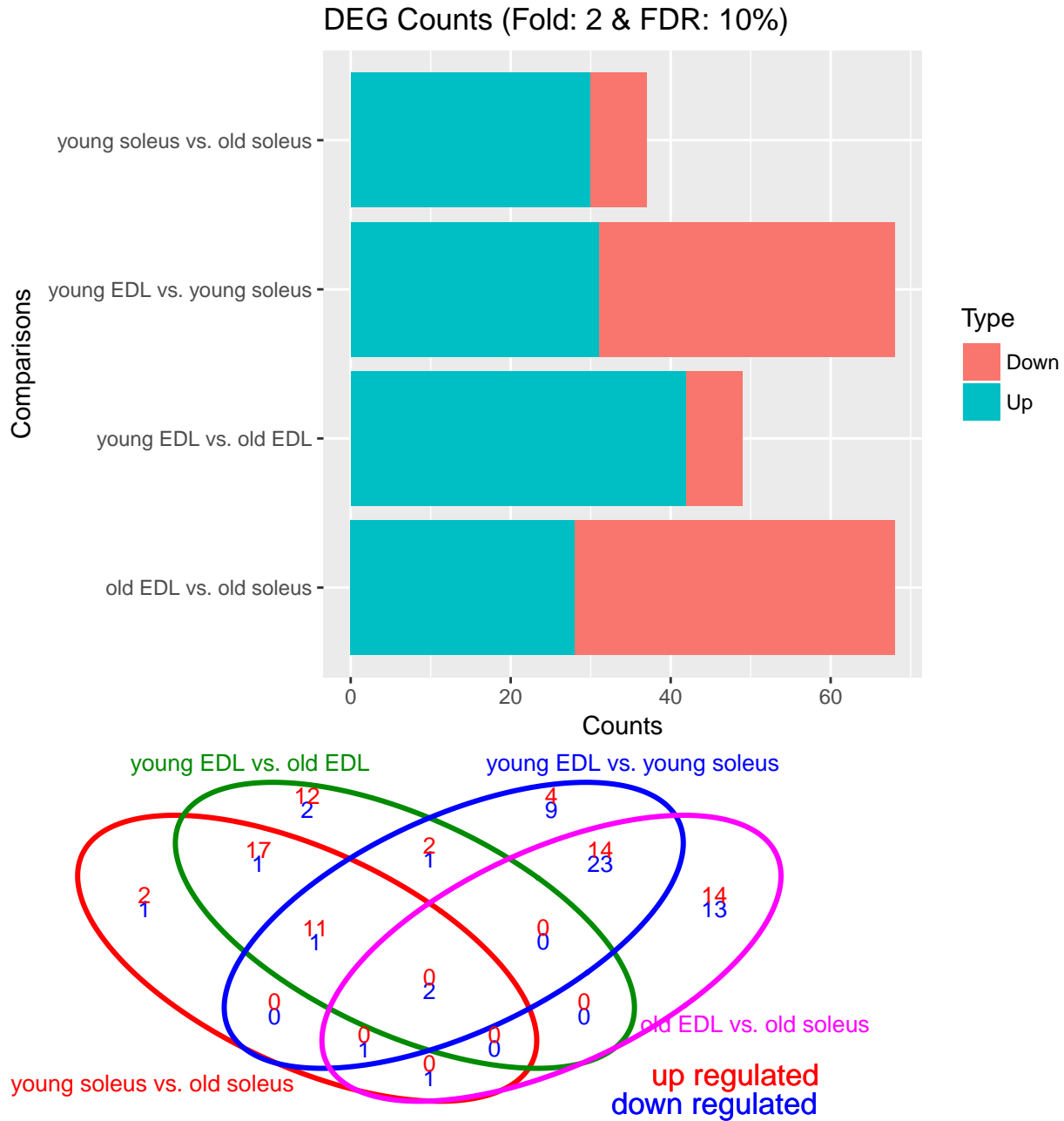
This is the numerical summary of the distribution of the p-values.

Cut	Count
0.0001	111
0.0010	139
0.0100	176
0.0250	204
0.0500	223
0.1000	262
0.2000	289
0.3000	313
0.4000	334
0.5000	351
0.6000	363
0.7000	370
0.8000	384
0.9000	400
1.0000	415

This table shows the number of features with p-values less or equal than some commonly used cutoff values.

Feature	log2FoldChange	baseMean	pvalue	padj
miR-208b-3p	-8.13	5.23	1.275776e-197	5.294471e-195
miR-499-5p	-4.64	9.51	2.042869e-158	4.238953e-156
miR-206-3p	-4.45	13.77	5.914150e-150	8.181240e-148
miR-547-3p	-5.51	4.50	1.079603e-118	1.120088e-116
miR-190b-5p	-3.85	3.29	5.355290e-73	4.444891e-71
miR-181b-5p	1.36	8.57	6.728434e-60	4.653834e-58
miR-24-2-5p	-1.78	8.15	1.874355e-53	1.111224e-51
miR-181a-2-3p	1.68	6.99	1.396162e-52	7.242593e-51
miR-181d-5p	-1.73	6.36	2.297578e-37	1.059439e-35
miR-135a-5p	-3.09	2.68	3.653031e-36	1.516008e-34
miR-24-3p	-1.36	11.26	4.775201e-36	1.801553e-34
miR-196a-5p	1.26	8.77	4.276760e-31	1.479046e-29
miR-23a-3p	-1.56	11.08	9.589708e-31	3.061330e-29
miR-365-2-5p	2.50	1.35	1.840735e-26	5.456465e-25
miR-27b-3p	-0.98	14.11	3.592945e-25	9.940482e-24
miR-201-5p	-4.27	0.56	1.023214e-24	2.653961e-23
miR-193a-5p	2.15	2.46	1.238375e-24	3.023091e-23
miR-29c-5p	1.21	4.63	4.955932e-24	1.142618e-22
miR-100-5p	1.02	12.36	5.636708e-24	1.231176e-22
miR-192-5p	-0.98	8.02	1.246806e-23	2.587123e-22

7 Summary of Differential Expression



8 Reproducibility

Software used in analyses pipeline

```
## [1] "samtools 1.7"
## [1] "cutadapt 1.16"
## [1] "bowtie-align-s version 1.2.2"
## [1] "sRNAbench version 1.5 - 6/2018"
```

```
## [1] "mirBase library version used: mature_22.fa"
```

The input for this report was generated with edgeR (Robinson, McCarthy, and Smyth, 2010; McCarthy, J., Chen, Yunshun, et al., 2012; Chen, Lun, and Smyth, 2014) . This report was generated in path /media/sf_VM_shared/vbcbf_musmus.

Date the report was generated.

```
## [1] "2018-07-06 00:20:03 CEST"
```

Wallclock time spent generating the report.

```
## Time difference of 1.524 mins
```

R session information.

```
## R version 3.4.4 (2018-03-15)
```

```
## Platform: x86_64-pc-linux-gnu (64-bit)
```

```
## Running under: Ubuntu 16.04.4 LTS
```

```
##
```

```
## Matrix products: default
```

```
## BLAS: /usr/lib/libblas/libblas.so.3.6.0
```

```
## LAPACK: /usr/lib/lapack/liblapack.so.3.6.0
```

```
##
```

```
## locale:
```

```
## [1] LC_CTYPE=en_US.UTF-8
```

```
LC_NUMERIC=C
```

```
LC_TIME=en_US.UTF-8
```

```
## [4] LC_COLLATE=en_US.UTF-8
```

```
LC_MONETARY=en_US.UTF-8
```

```
LC_MESSAGES=en_US.UTF-8
```

```
## [7] LC_PAPER=en_US.UTF-8
```

```
LC_NAME=en_US.UTF-8
```

```
LC_ADDRESS=en_US.UTF-8
```

```
## [10] LC_TELEPHONE=en_US.UTF-8
```

```
LC_MEASUREMENT=en_US.UTF-8
```

```
LC_IDENTIFICATION=en_US.UTF-8
```

```
##
```

```
## attached base packages:
```

```
## [1] tools      parallel  stats4    methods  stats      graphics  grDevices  utils      datasets  base
```

```
##
```

```
## other attached packages:
```

```
## [1] bindrcpp_0.2          rmarkdown_1.9          knitcitations_1.0.8    kableExtra_0.7
```

```
## [5] knitr_1.20            RColorBrewer_1.1-2     pheatmap_1.0.8         DT_0.4
```

```
## [9] xlsx_0.5.7            xlsxjars_0.6.1         rJava_0.9-10           systemPipeR_1.
```

```
## [13] ShortRead_1.36.1      GenomicAlignments_1.14.1 BiocParallel_1.12.0    Rsamtools_1.30
```

```
## [17] Biostrings_2.46.0     XVector_0.18.0         pcaMethods_1.70.0      ggrepel_0.7.3
```

```
## [21] ggplot2_2.2.1.9000    dplyr_0.7.4            vsn_3.46.0             DESeq2_1.18.1
```

```
## [25] SummarizedExperiment_1.8.1 DelayedArray_0.4.1      matrixStats_0.53.1     Biobase_2.38.0
```

```
## [29] GenomicRanges_1.30.3  GenomeInfoDb_1.14.0    IRanges_2.12.0         S4Vectors_0.16
```

```
## [33] BiocGenerics_0.24.0   edgeR_3.20.9           limma_3.34.9
```

```
##
```

```
## loaded via a namespace (and not attached):
```

```
## [1] backports_1.1.2        GOstats_2.44.0          Hmisc_4.1-1            plyr_1.8.4
```

```
## [5] lazyeval_0.2.1         GSEABase_1.40.1         splines_3.4.4          BatchJobs_1.7
```

```
## [9] digest_0.6.15          foreach_1.4.4           BiocInstaller_1.28.0    htmltools_0.3.6
```

```
## [13] GO.db_3.5.0            magrittr_1.5            checkmate_1.8.5         memoise_1.1.0
```

```
## [17] BBmisc_1.11            BSgenome_1.46.0         cluster_2.0.6           readr_1.1.1
```

```
## [21] annotate_1.56.1         prettyunits_1.0.2       colorspace_1.3-2        rvest_0.3.2
```

```
## [25] blob_1.1.0             RCurl_1.95-4.10         jsonlite_1.5            graph_1.56.0
```

```
## [29] genefilter_1.60.0       bindr_0.1.1             brew_1.0-6              survival_2.41-3
```

```
## [33] sendmailR_1.2-1        VariantAnnotation_1.24.5 iterators_1.0.9          glue_1.2.0
```

```
## [37] registry_0.5           gtable_0.2.0            DEFormats_1.6.1         zlibbioc_1.24.0
```

```
## [41] Rgraphviz_2.22.0       scales_0.5.0.9000       DBI_0.8                 rngtools_1.2.4
```

```
## [45] bibtex_0.4.2           derfinder_1.12.6        derfinderHelper_1.12.0  Rcpp_0.12.16
```

## [49] viridisLite_0.3.0	xtable_1.8-2	progress_1.1.2	htmlTable_1.11.2
## [53] bumpHunter_1.20.0	foreign_0.8-69	bit_1.1-12	preprocessCore_1.40
## [57] Formula_1.2-2	AnnotationForge_1.20.0	htmlwidgets_1.0	httr_1.3.1
## [61] acepack_1.4.1	pkgconfig_2.0.1	XML_3.98-1.10	nnet_7.3-12
## [65] locfit_1.5-9.1	labeling_0.3	rlang_0.2.0.9000	reshape2_1.4.3
## [69] AnnotationDbi_1.40.0	munsell_0.4.3	RSQLite_2.0	evaluate_0.10.1
## [73] stringr_1.3.0	yaml_2.1.18	RefManageR_0.14.20	bit64_0.9-7
## [77] RBGL_1.54.0	doRNG_1.6.6	xml2_1.2.0	myReport_1.0
## [81] biomaRt_2.34.2	BiocStyle_2.6.1	compiler_3.4.4	rstudioapi_0.7
## [85] affyio_1.48.0	tibble_1.4.2	geneplotter_1.56.0	stringi_1.1.7
## [89] highr_0.6	GenomicFeatures_1.30.3	GenomicFiles_1.14.0	lattice_0.20-35
## [93] Matrix_1.2-12	markdown_0.8	pillar_1.2.1	data.table_1.10.4-3
## [97] bitops_1.0-6	rtracklayer_1.38.3	qvalue_2.10.0	R6_2.2.2
## [101] latticeExtra_0.6-28	affy_1.56.0	hwriter_1.3.2	RMySQL_0.10.14
## [105] gridExtra_2.3	knitrBootstrap_1.0.1	codetools_0.2-15	assertthat_0.2.0
## [109] pkgmaker_0.22	Category_2.44.0	rprojroot_1.3-2	rjson_0.2.15
## [113] withr_2.1.2	GenomeInfoDbData_1.0.0	hms_0.4.2	grid_3.4.4
## [117] rpart_4.1-13	lubridate_1.7.3	base64enc_0.1-3	

Pandoc version used: 2.2.1.

9 Bibliography

This report was created using **rmarkdown** (Allaire, Xie, McPherson, Luraschi, et al., 2018) while **knitr** (Xie, 2014) and **DT** (Xie, 2018) were running behind the scenes. **pheatmap** (Kolde, 2015) was used to create the sample distances heatmap. Several plots were made with **ggplot2** (Wickham, 2016).

Citations made with **knitcitations** (Boettiger, 2017). The BibTeX file can be found [here](#).

[1] J. Allaire, Y. Xie, J. McPherson, J. Luraschi, et al. **rmarkdown: Dynamic Documents for R**. R package version 1.9. 2018. URL: <https://CRAN.R-project.org/package=rmarkdown>.

[1] C. Boettiger. **knitcitations: Citations for ‘Knitr’ Markdown Files**. R package version 1.0.8. 2017. URL: <https://CRAN.R-project.org/package=knitcitations>.

[1] Y. Chen, A. T. L. Lun and G. K. Smyth. “Differential expression analysis of complex RNA-seq experiments using edgeR”. In: *Statistical Analysis of Next Generation Sequencing Data*. Ed. by S. Datta and D. Nettleton. New York: Springer, 2014, pp. 51-74.

[1] R. Kolde. **pheatmap: Pretty Heatmaps**. R package version 1.0.8. 2015. URL: <https://CRAN.R-project.org/package=pheatmap>.

[1] McCarthy, D. J., Chen, Yunshun, et al. “Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation”. In: *Nucleic Acids Research* 40.10 (2012), pp. 4288-4297.

[1] M. D. Robinson, D. J. McCarthy and G. K. Smyth. “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data”. In: *Bioinformatics* 26.1 (2010), pp. 139-140.

[1] H. Wickham. **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4. URL: <http://ggplot2.org>.

[1] Y. Xie. **DT: A Wrapper of the JavaScript Library ‘DataTables’**. R package version 0.4. 2018. URL: <https://CRAN.R-project.org/package=DT>.

[1] Y. Xie. “knitr: A Comprehensive Tool for Reproducible Research in R”. In: *Implementing Reproducible Computational Research*. Ed. by V. Stodden, F. Leisch and R. D. Peng. ISBN 978-1466561595. Chapman and Hall/CRC, 2014. URL: <http://www.crcpress.com/product/isbn/9781466561595>.