Tiedonlouhintaa hiirien ja rottien geeniekspressiodatasta

Tiedonlouhinta 2013

Antti Kurronen ja Irene Pöllänen

Data: Rotat

- Ekspressiodataa, joissa geenien suhteellisten ilmenemisten tasoja*
- ~ 30 000 riviä -> ~ 1 rivi / geeni

ZL_E18_SD_1_150405 - linear (x3)
–18 päivän alkio

ZL_neo_SD_1_080405 - linear (x3)
-vastasyntynyt

ZL_d10-rat-heart_#3_130407 - linear (x3)
–10 päivän vanha

ZL_adult_SD_1_110505 - linear (x3)
–nuori aikuinen 3kk

ZL_adult_WS_1_100605 - linear (x3)
–nuori aikuinen, eri kantaa (Wistar, muut Sprague-Dawley)

ZL_old_WS_1_100605 - linear (x3)
–vanha Wistar 24kk

*(Kustakin rotasta kolme replikaattia)

Data:Hiiret

- Ekspressiodataa, joissa geenien suhteellisten ilmenemisten tasoja
- ~ 25 000 riviä -> ~ 1 rivi / geeni

JP_control_1_050511_(MoGene-1_0-st-v1) - linear (x3)

–Kontrolli villityypin hiiri

JP_mutant_1_050511_(MoGene-1_0-st-v1) - linear (x3)

–Twinkle-transgeeninen hiiri

JP_S1_SOD het_210512_(MoGene-1_0-st-v1) - linear (x3)

–SOD2 heterotsygootti poistogeeninen

JP_ST1_SOD_het_Tw_pos_251012_wdh_(MoGene-1_0-st-v1) - linear (x3)

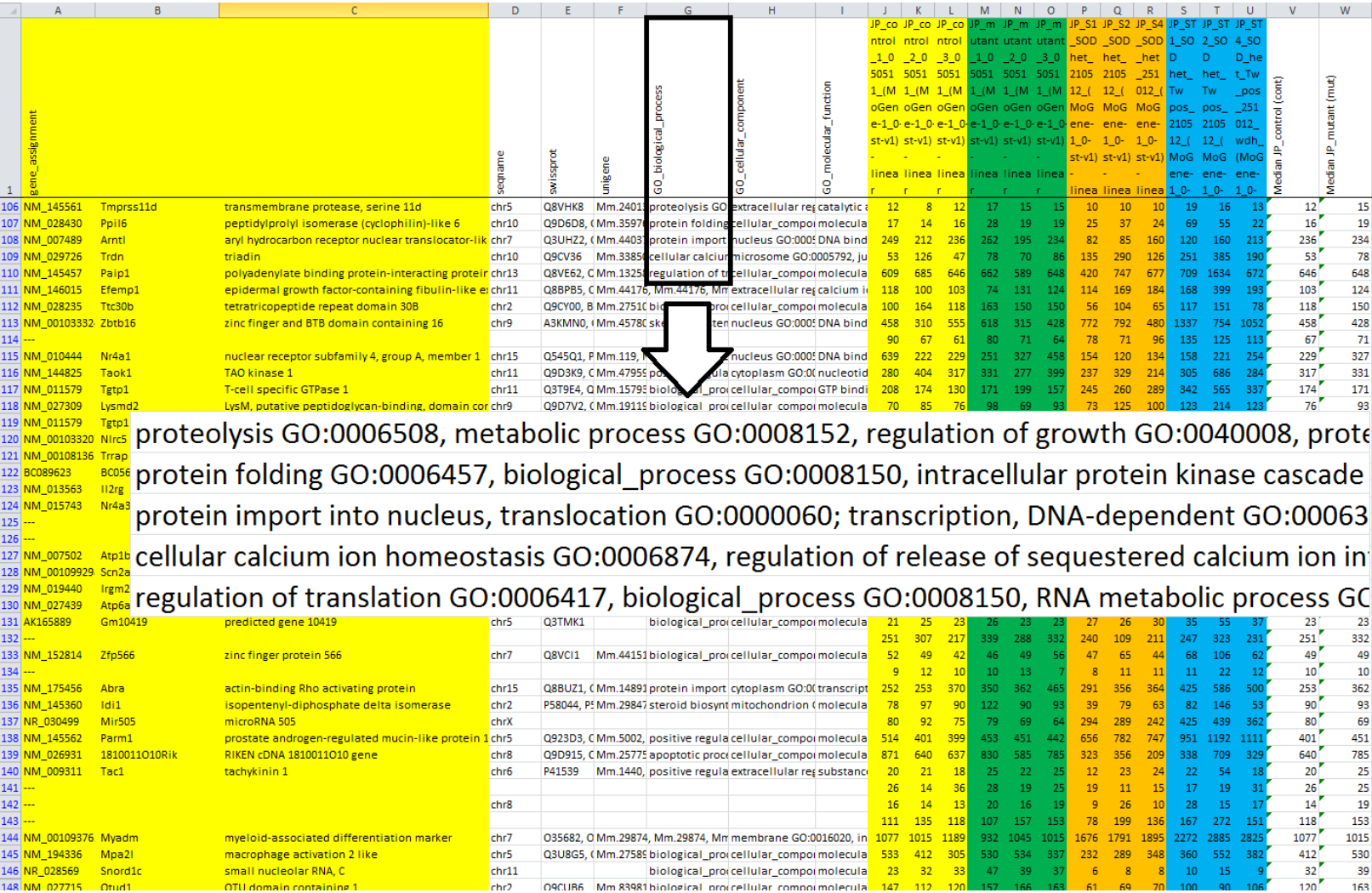
– SOD2 heterotsygootti poistogeeninen x Twinkle-transgeeni risteymä

*(Kustakin hiirestä kolme replikaattia)

Data (hiiri)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	gene_assignment			seqname	swissprot	unigene	GO_biological_process	GO_cellular_component	GO_molecular_function	JP_control_1_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_control_2_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_control_3_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_mutant_1_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_mutant_2_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_mutant_3_0_5051_1_(MoGene-1_0-st-v1)_-_linear	JP_S1_SOD_het_2105_12_(MoGene-1_0-st-v1)_-_linear	JP_S2_SOD_het_2105_12_(MoGene-1_0-st-v1)_-_linear	JP_S4_SOD_het_251012_(MoGene-1_0-st-v1)_-_linear	JP_ST1_SOD_het_Two_pos_2105_12_(MoGene-1_0-	JP_ST2_SOD_het_Two_pos_2105_12_(MoGene-1_0-	JP_ST4_SOD_het_Two_pos_251012_wdh_(MoGene-1_0-	Median JP_control (cont)	Median JP_mutant (mut)	Median JP_S1-3_SOD (SOD1)	Median JP_ST1,2,4_SOD (SOD2)	(mut - cont)/cont	(SOD1 - cont)/cont	(SOD2 - cont)/cont
106	NM_145561	Tmprss11d	transmembrane protease, serine 11d	chr5	Q8VHK8	Mm.24013	proteolysis GO	extracellular reg	catalytic a	12	8	12	17	15	15	10	10	10	19	16	13	12	15	10	16	0,324646	-0,11632	0,418469
107	NM_028430	Ppil6	peptidylprolyl isomerase (cyclophilin)-like 6	chr10	Q9D6D8, (	Mm.35976	protein folding	cellular_compo	molecula	17	14	16	28	19	19	25	37	24	69	55	22	16	19	25	55	0,192351	0,594374	2,507585
108	NM_007489	Arntl	aryl hydrocarbon receptor nuclear translocator-lik	chr7	Q3UHZ2, (	Mm.44037	protein import	nucleus GO:0005	DNA bind	249	212	236	262	195	234	82	85	160	120	160	213	236	234	85	160	-0,00662	-0,63922	-0,31971
109	NM_029726	Trdn	triadin	chr10	Q9CV36	Mm.33850	cellular calciur	microsome GO:0005792, ju		53	126	47	78	70	86	135	290	126	251	385	190	53	78	135	251	0,466875	1,545725	3,716757
110	NM_145457	Paip1	polyadenylate binding protein-interacting protein	chr13	Q8VE62, (	Mm.13258	regulation of tr	cellular_compo	molecula	609	685	646	662	589	648	420	747	677	709	1634	672	646	648	677	709	0,002755	0,048024	0,097129
111	NM_146015	Efemp1	epidermal growth factor-containing fibulin-like ex	chr11	Q8BPB5, (	Mm.44176,	Mm.44176, Mm	extracellular reg	calcium i	118	100	103	74	131	124	114	169	184	168	399	193	103	124	169	193	0,2015	0,637205	0,863912
112	NM_028235	Ttc30b	tetratricopeptide repeat domain 30B	chr2	Q9CY00, B	Mm.27510	biological_proc	cellular_compo	molecula	100	164	118	163	150	150	56	104	65	117	151	78	118	150	65	117	0,26906	-0,44683	-0,01295
113	NM_00103332	Zbtb16	zinc finger and BTB domain containing 16	chr9	A3KMN0, (	Mm.45780	skeletal systen	nucleus GO:0005	DNA bind	458	310	555	618	315	428	772	792	480	1337	754	1052	458	428	772	1052	-0,06611	0,684454	1,295316
114	---									90	67	61	80	71	64	78	71	96	135	125	113	67	71	78	125	0,063907	0,155978	0,864609
115	NM_010444	Nr4a1	nuclear receptor subfamily 4, group A, member 1	chr15	Q545Q1, P	Mm.119, M	transcription, D	nucleus GO:0005	DNA bind	639	222	229	251	327	458	154	120	134	158	221	254	229	327	134	221	0,42982	-0,41583	-0,03535
116	NM_144825	Taok1	TAO kinase 1	chr11	Q9D3K9, (	Mm.47959	positive regula	cytoplasm GO:0(	nucleotid	280	404	317	331	277	399	237	329	214	305	686	284	317	331	237	305	0,045368	-0,25384	-0,03788
117	NM_011579	Tgtp1	T-cell specific GTPase 1	chr11	Q3T9E4, (	Mm.15793	biological_proc	cellular_compo	GTP bindi	208	174	130	171	199	157	245	260	289	342	565	337	174	171	260	342	-0,01946	0,495667	0,967633
118	NM_027309	Lysmd2	LysM, putative peptidoglycan-binding, domain cor	chr9	Q9D7V2, (	Mm.19115	biological_proc	cellular_compo	molecula	70	85	76	98	69	93	73	125	100	123	214	123	76	93	100	123	0,23008	0,320196	0,625041
119	NM_011579	Tgtp1	T-cell specific GTPase 1	chr11	Q3T9E4, (	Mm.15793	biological_proc	cellular_compo	GTP bindi	224	193	140	173	218	168	258	293	319	369	651	354	193	173	293	369	-0,10162	0,520814	0,913008
120	NM_00103320	Nlrc5	NLR family, CARD domain containing 5	chr8	Q571H9, (	Mm.33472	defense respor	cellular_compo	nucleotid	40	34	23	28	43	20	23	23	22	25	62	27	34	28	23	27	-0,18176	-0,31415	-0,21056
121	NM_00108136	Trrap	transformation	chr5	Q80YV3, (	Mm.27357	transcription, D	histone acetyltr	transcript	142	134	133	164	152	181	70	115	113	147	160	134	134	164	113	147	0,22354	-0,15849	0,094466
122	BC089623	BC056474	cDNA sequence BC056474	chr8	Q6ZWX0	Mm.35790	biological_proc	cellular_compo	molecula	409	196	422	341	412	407	1543	649	867	1069	2055	1363	409	407	867	1363	-0,00349	1,120651	2,332936
123	NM_013563	Il2rg	interleukin 2 receptor, gamma chain	chrX	Q3UPA9, F	Mm.2923,	regulation of g	external side of	receptor a	216	230	208	206	198	200	243	229	246	313	518	290	216	200	243	313	-0,07581	0,126813	0,448067
124	NM_015743	Nr4a3	nuclear receptor subfamily 4, group A, member 3	chr4	A2VCP6, A	Mm.24726	mesoderm form	nucleus GO:0005	DNA bind	53	30	32	34	40	41	44	51	44	65	62	83	32	40	44	65	0,275308	0,387749	1,056678
125	---									27	43	36	44	27	20	42	37	39	51	82	51	36	27	39	51	-0,25669	0,099618	0,434281
126	---			chr19						253	188	289	250	249	293	136	193	122	202	338	158	253	250	136	202	-0,01166	-0,46377	-0,20143
127	NM_007502	Atp1b3	ATPase, Na+	chr9	Q544Q7, (	Mm.424, M	ATP catabolic p	cytoplasm GO:0(	sodium:p	1000	1060	962	1014	925	1153	584	896	551	922	1441	735	1000	1014	584	922	0,013603	-0,41624	-0,07757
128	NM_00109929	Scn2a1	sodium channel, voltage-gated, type II, alpha 1	chr2	Q8CAJ8, Q	Mm.22032	transport GO:0(	voltage-gated s	ion chanr	10	11	10	12	14	10	10	10	9	16	13	13	10	12	10	13	0,155502	-0,08352	0,280154
129	NM_019440	Irgm2	immunity-related GTPase family M member 2	chr11	Q6ZQL9, (	Mm.458517,	Mm.33902, M	cellular_compo	GTPase a	232	222	151	183	372	166	181	236	230	259	551	232	222	183	230	259	-0,17382	0,038475	0,168584
130	NM_027439	Atp6ap2	ATPase, H+ transporting, lysosomal accessory prot	chrX	Q9CYN9, (	Mm.25148	signal transduc	cellular_compo	molecula	138	154	152	162	122	153	111	166	122	157	285	166	152	153	122	166	0,004913	-0,20121	0,09181
131	AK165889	Gm10419	predicted gene 10419	chr5	Q3TMK1		biological_proc	cellular_compo	molecula	21	25	23	26	23	23	27	26	30	35	55	37	23	23	27	37	0,016643	0,181511	0,616782
132	---									251	307	217	339	288	332	240	109	211	247	323	231	251	332	211	247	0,321004	-0,16042	-0,0156
133	NM_152814	Zfp566	zinc finger protein 566	chr7	Q8VCI1	Mm.44151	biological_proc	cellular_compo	molecula	52	49	42	46	49	56	47	65	44	68	106	62	49	49	47	68	0,010166	-0,04466	0,380917
134	---									9	12	10	10	13	7	8	11	11	11	22	12	10	10	11	12	0,044854	0,098657	0,242474
135	NM_175456	Abra	actin-binding Rho activating protein	chr15	Q8BUZ1, (	Mm.14891	protein import	cytoplasm GO:0(	transcript	252	253	370	350	362	465	291	356	364	425	586	500	253	362	356	500	0,432226	0,406309	0,977131
136	NM_145360	Idi1	isopentenyl-diphosphate delta isomerase	chr2	P58044, P	Mm.29847	steroid biosynt	mitochondrion (	molecula	78	97	90	122	90	93	39	79	63	82	146	53	90	93	63	82	0,035782	-0,30658	-0,09097
137	NR_030499	Mir505	microRNA 505	chrX						80	92	75	79	69	64	294	289	242	425	439	362	80	69	289	425	-0,14349	2,596036	4,296773
138	NM_145562	Parm1	prostate androgen-regulated mucin-like protein 1	chr5	Q923D3, (	Mm.5002,	positive regula	cellular_compo	molecula	514	401	399	453	451	442	656	782	747	951	1192	1111	401	451	747	1111	0,124635	0,862422	1,77219
139	NM_026931	1810011O10Rik	RIKEN cDNA 1810011O10 gene	chr8	Q9D915, (	Mm.25775	apoptotic proc	cellular_compo	molecula	871	640	637	830	585	785	323	356	209	338	709	329	640	785	323	338	0,226267	-0,49606	-0,47195
140	NM_009311	Tac1	tachykinin 1	chr6	P41539	Mm.1440,	positive regula	extracellular reg	substanc	20	21	18	25	22	25	12	23	24	22	54	18	20	25	23	22	0,2581	0,141586	0,100693
141	---									26	14	36	28	19	25	19	11	15	17	19	31	26	25	15	19	-0,03516	-0,44483	-0,27604
142	---			chr8						16	14	13	20	16	19	9	26	10	28	15	17	14	19	10	17	0,370734	-0,27549	0,202212
143	---									111	135	118	107	157	153	78	199	136	167	272	151	118	153	136	167	0,296944	0,152794	0,418172
144	NM_00109376	Myadm	myeloid-associated differentiation marker	chr7	O35682, O	Mm.29874,	Mm.29874, Mm	membrane GO:0016020, in		1077	1015	1189	932	1045	1015	1676	1791	1895	2272	2885	2825	1077	1015	1791	2825	-0,05776	0,662836	1,623044
145	NM_194336	Mpa2l	macrophage activation 2 like	chr5	Q3U8G5, (	Mm.27585	biological_proc	cellular_compo	molecula	533	412	305	530	534	337	232	289	348	360	552	382	412	530	289	382	0,286173	-0,29926	-0,07251
146	NR_028569	Snord1c	small nucleolar RNA, C	chr11			biological_proc	cellular_compo	molecula	23	32	33	47	39	37	6	8	8	10	15	9	32	39	8	10	0,207483	-0,75702	-0,69518
148	NM_027715	Otud1	OTU domain containing 1	chr2	Q9CUB6	Mm.83981	biological_proc	cellular_compo	molecula	147	112	120	157	166	163	61	69	70	100	90	106	120	163	69	100	0,359394	-0,42443	-0,16705

Geeniontologian termit (GO)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	gene_assignment			seqname	swissprot	unigene	GO_biological_process	GO_cellular_component	GO_molecular_function	JP_control_1_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_control_2_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_control_3_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_mutant_1_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_mutant_2_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_mutant_3_0_5051_1_(MoGene-1_0-st-v1)-linear	JP_S1_SOD_het_2105_12_(MoGene-1_0-st-v1)-linear	JP_S2_SOD_het_2105_12_(MoGene-1_0-st-v1)-linear	JP_S4_SOD_het_251012_(MoGene-1_0-st-v1)-linear	JP_ST1_SOD_het_Tw pos_2105_12_(MoGene-1_0-	JP_ST2_SOD_het_Tw pos_2105_12_(MoGene-1_0-	JP_ST4_SOD_het_Tw pos_251012_wdh_(MoGene-1_0-	Median JP_control (cont)	Median JP_mutant (mut)	Median JP_S1-3_SOD (SOD1)	Median JP_ST1,2,4_SOD (SOD2)	(mut - cont)/cont	(SOD1 - cont)/cont	(SOD2 - cont)/cont
106	NM_145561	Tmprss11d	transmembrane protease, serine 11d	chr5	Q8VHK8	Mm.2401	proteolysis GO	extracellular re	catalytic a	12	8	12	17	15	15	10	10	10	19	16	13	12	15	10	16	0,324646	-0,11632	0,418469
107	NM_028430	Ppil6	peptidylprolyl isomerase (cyclophilin)-like 6	chr10	Q9D6D8, (	Mm.3597	protein folding	cellular_compo	molecula	17	14	16	28	19	19	25	37	24	69	55	22	16	19	25	55	0,192351	0,594374	2,507585
108	NM_007489	Arntl	aryl hydrocarbon receptor nuclear translocator-lik	chr7	Q3UHZ2, (	Mm.4403	protein import	nucleus GO:000	DNA bind	249	212	236	262	195	234	82	85	160	120	160	213	236	234	85	160	-0,00662	-0,63922	-0,31971
109	NM_029726	Trdn	triadin	chr10	Q9CV36	Mm.3385	cellular calcium	microsome GO:0005792, ju		53	126	47	78	70	86	135	290	126	251	385	190	53	78	135	251	0,466875	1,545725	3,716757
110	NM_145457	Paip1	polyadenylate binding protein-interacting protein	chr13	Q8VE62, (	Mm.1325	regulation of tr	cellular_compo	molecula	609	685	646	662	589	648	420	747	677	709	1634	672	646	648	677	709	0,002755	0,048024	0,097129
111	NM_146015	Efemp1	epidermal growth factor-containing fibulin-like ex	chr11	Q8BPB5, (	Mm.44176, Mm.44176, Mm	extracellular re	calcium i		118	100	103	74	131	124	114	169	184	168	399	193	103	124	169	193	0,2015	0,637205	0,863912
112	NM_028235	Ttc30b	tetratricopeptide repeat domain 30B	chr2	Q9CY00, B	Mm.2751(	bio	cellular_compo	molecula	100	164	118	163	150	150	56	104	65	117	151	78	118	150	65	117	0,26906	-0,44683	-0,01295
113	NM_00103332	Zbtb16	zinc finger and BTB domain containing 16	chr9	A3KMN0, (	Mm.4578(	ske	ten	nucleus GO:000	DNA bind	458	310	555	618	315	428	772	792	480	1337	754	1052	458	428	772	1052	-0,06611	0,684454	1,295316
114	---									90	67	61	80	71	64	78	71	96	135	125	113	67	71	78	125	0,063907	0,155978	0,864609
115	NM_010444	Nr4a1	nuclear receptor subfamily 4, group A, member 1	chr15	Q545Q1, P	Mm.119,			nucleus GO:000	DNA bind	639	222	229	251	327	458	154	120	134	158	221	254	229	327	134	221	0,42982	-0,41583	-0,03535
116	NM_144825	Taok1	TAO kinase 1	chr11	Q9D3K9, (	Mm.4795	po	gula	cytoplasm GO:0(	nucleotid	280	404	317	331	277	399	237	329	214	305	686	284	317	331	237	305	0,045368	-0,25384	-0,03788
117	NM_011579	Tgtp1	T-cell specific GTPase 1	chr11	Q3T9E4, Q	Mm.1579	biol	al_pro	cellular_compo	GTP bindi	208	174	130	171	199	157	245	260	289	342	565	337	174	171	260	342	-0,01946	0,495667	0,967633
118	NM_027309	Lysmd2	LysM, putative peptidoglycan-binding, domain cor	chr9	Q9D7V2, (	Mm.1911	biological_pro	cellular_compo	molecula	70	85	76	98	69	93	73	125	100	123	214	123	76	93	100	123	0,23008	0,320196	0,625041
119	NM_011579	Tgtp1	T-cell specific GTPase 1	chr11	Q3T9E4, Q	Mm.1579	biol	al_pro	cellular_compo	GTP bindi	208	174	130	171	199	157	245	260	289	342	565	337	174	171	260	293	-0,10162	0,520814	0,913008
120	NM_00103320	Nlrc5	NLR family, CARD domain containing 5	chr8	Q3TMK1		biological_pro	cellular_compo	molecula	21	25	23	26	23	23	27	26	30	35	55	37	23	23	27	37	0,016643	0,181511	0,616782
121	NM_00108136	Trrap	transformation/transcription domain-associated	chr5	Q8VCI1	Mm.4415	biological_pro	cellular_compo	molecula	52	49	42	46	49	56	47	65	44	68	106	62	49	49	47	68	0,010166	-0,04466	0,380917
122	BC089623	BC056	cDNA sequence BC056	chr15	Q8BUZ1, (	Mm.1489	protein import	cytoplasm GO:0(	transcript	252	253	370	350	362	465	291	356	364	425	586	500	253	362	356	500	0,432226	0,406309	0,977131
123	NM_013563	Il2rg	interleukin 2 receptor, gamma chain	chr2	P58044, P	Mm.2984	steroid biosynt	mitochondrion (	molecula	78	97	90	122	90	93	39	79	63	82	146	53	90	93	63	82	0,035782	-0,30658	-0,09097
124	NM_015743	Nr4a3	nuclear receptor subfamily 4, group A, member 3	chrX						80	92	75	79	69	64	294	289	242	425	439	362	80	69	289	425	-0,14349	2,596036	4,296773
125	---			chr5	Q923D3, (	Mm.5002	positive regula	cellular_compo	molecula	514	401	399	453	451	442	656	782	747	951	1192	1111	401	451	747	1111	0,124635	0,862422	1,77219
126	---			chr8	Q9D915, (	Mm.2577	apoptotic proc	cellular_compo	molecula	871	640	637	830	585	785	323	356	209	338	709	329	640	785	323	338	0,226267	-0,49606	-0,47195
127	NM_007502	Atp1b	ATPase, Na+/K+ transporting, beta 1 polypeptide	chr6	P41539	Mm.1440	positive regula	extracellular re	substanc	20	21	18	25	22	25	12	23	24	22	54	18	20	25	23	22	0,2581	0,141586	0,100693
128	NM_00109929	Scn2a	sodium channel, voltage-gated, type II, alpha							26	14	36	28	19	25	19	11	15	17	19	31	26	25	15	19	-0,03516	-0,44483	-0,27604
129	NM_019440	Irgm2	immunity-related GTPase family M member 2							16	14	13	20	16	19	9	26	10	28	15	17	14	19	10	17	0,370734	-0,27549	0,202212
130	NM_027439	Atp6a	ATPase, H+ transporting, lysosomal V0 subunit A							111	135	118	107	157	153	78	199	136	167	272	151	118	153	136	167	0,296944	0,152794	0,418172
131	AK165889	Gm10419	predicted gene 10419	chr5	Q3TMK1		biological_pro	cellular_compo	molecula	21	25	23	26	23	23	27	26	30	35	55	37	23	23	27	37	0,016643	0,181511	0,616782
132	---									251	307	217	339	288	332	240	109	211	247	323	231	251	332	211	247	0,321004	-0,16042	-0,0156
133	NM_152814	Zfp566	zinc finger protein 566	chr7	Q8VCI1	Mm.4415	biological_pro	cellular_compo	molecula	52	49	42	46	49	56	47	65	44	68	106	62	49	49	47	68	0,010166	-0,04466	0,380917
134	---									9	12	10	10	13	7	8	11	11	11	22	12	10	10	11	12	0,044854	0,098657	0,242474
135	NM_175456	Abra	actin-binding Rho activating protein	chr15	Q8BUZ1, (	Mm.1489	protein import	cytoplasm GO:0(	transcript	252	253	370	350	362	465	291	356	364	425	586	500	253	362	356	500	0,432226	0,406309	0,977131
136	NM_145360	Idi1	isopentenyl-diphosphate delta isomerase	chr2	P58044, P	Mm.2984	steroid biosynt	mitochondrion (	molecula	78	97	90	122	90	93	39	79	63	82	146	53	90	93	63	82	0,035782	-0,30658	-0,09097
137	NR_030499	Mir505	microRNA 505	chrX						80	92	75	79	69	64	294	289	242	425	439	362	80	69	289	425	-0,14349	2,596036	4,296773
138	NM_145562	Parm1	prostate androgen-regulated mucin-like protein 1	chr5	Q923D3, (	Mm.5002	positive regula	cellular_compo	molecula	514	401	399	453	451	442	656	782	747	951	1192	1111	401	451	747	1111	0,124635	0,862422	1,77219
139	NM_026931	1810011O10Rik	RIKEN cDNA 1810011O10 gene	chr8	Q9D915, (	Mm.2577	apoptotic proc	cellular_compo	molecula	871	640	637	830	585	785	323	356	209	338	709	329	640	785	323	338	0,226267	-0,49606	-0,47195
140	NM_009311	Tac1	tachykinin 1	chr6	P41539	Mm.1440	positive regula	extracellular re	substanc	20	21	18	25	22	25	12	23	24	22	54	18	20	25	23	22	0,2581	0,141586	0,100693
141	---									26	14	36	28	19	25	19	11	15	17	19	31	26	25	15	19	-0,03516	-0,44483	-0,27604
142	---			chr8						16	14	13	20	16	19	9	26	10	28	15	17	14	19	10	17	0,370734	-0,27549	0,202212
143	---									111	135	118	107	157	153	78	199	136	167	272	151	118	153	136	167	0,296944	0,152794	0,418172
144	NM_00109376	Myadm	myeloid-associated differentiation marker	chr7	O35682, O	Mm.29874, Mm.29874, Mm	membrane GO:0016020, in			1077	1015	1189	932	1045	1015	1676	1791	1895	2272	2885	2825	1077	1015	1791	2825	-0,05776	0,662836	1,623044
145	NM_194336	Mpa2l	macrophage activation 2 like	chr5	Q3U8G5, (	Mm.2758	biological_pro	cellular_compo	molecula	533	412	305	530	534	337	232	289	348	360	552	382	412	530	289	382	0,286173	-0,29926	-0,07251
146	NR_028569	Snord1c	small nucleolar RNA, C	chr11			biological_pro	cellular_compo	molecula	23	32	33	47	39	37	6	8	8	10	15	9	32	39	8	10	0,207483	-0,75702	-0,69518
147	NM_027715	Otud1	OTU domain containing 1	chr2	Q9CUB6	Mm.83981	biological_pro	cellular_compo	molecula	147	112	120	152	166	163	61	69	70	100	90	106	120	163	69	100	0,359394	-0,42443	-0,16705

proteolysis GO:0006508, metabolic process GO:0008152, regulation of growth GO:0040008, prote
protein folding GO:0006457, biological_process GO:0008150, intracellular protein kinase cascade
protein import into nucleus, translocation GO:0000060; transcription, DNA-dependent GO:00063
cellular calcium ion homeostasis GO:0006874, regulation of release of sequestered calcium ion in
regulation of translation GO:0006417, biological_process GO:0008150, RNA metabolic process GO

Tiedonlouhinta tavoitteita:

Geenien ekspression voimistuminen ja pienentyminen geeniontologialuokittain

- Tietyissä biologisissa prosesseissa (Biological Process)
- Molekulaarisissa toiminnoissa (Molecular Function)
- Solun osissa (Cellular Component)

-> Samankaltaisuudet ekspressiotasojen muutoksissa hiirillä ja rotilla?

Esim. Proteolyysi on yksi biologisen prosessin luokka, johon kuuluvat geenit hajoittavat proteiineja

Hiiren ja rotan ekspressiotasojen vertailu uusilla muuttujilla

- Vaikuttaako hiirien mutaatiot geenien ekspressiotasoihin samalla tavoin kuin rotilla vanheneminen

Toteutus

Datan esikäsittelyyn

- R + omia funktioita
- Excel + VBA skriptejä
- Gnumeric

Tilastolliseen analysointiin

- Kingfisher
- Jokin rinnakkainen menetelmä (R tai Matlab)

Datan esikäsittely:

Poistetaan

- Datarivit, joista puuttuu geenitiedot
- Identtiset rivit (poistetaan kopiot)

Yhdistetään

- Datarivit, jotka ovat saman geenin tuloksia

-> valitaan arvot, joilla hajonta pientä

Muuttujat

Luodaan ryhmien välisiä eroja kuvaavia muuttujia

Datan esikäsittely:

Normalisointi

- Lowessin menetelmällä

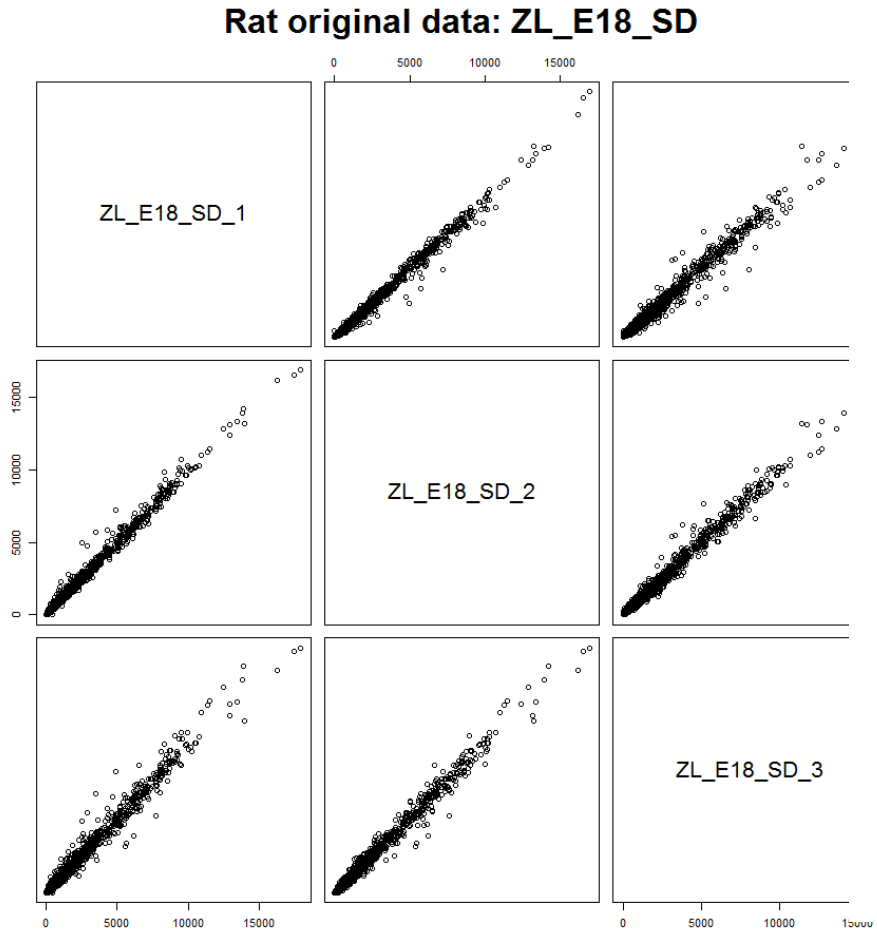
Numeerisille arvoille log2 -muunnos

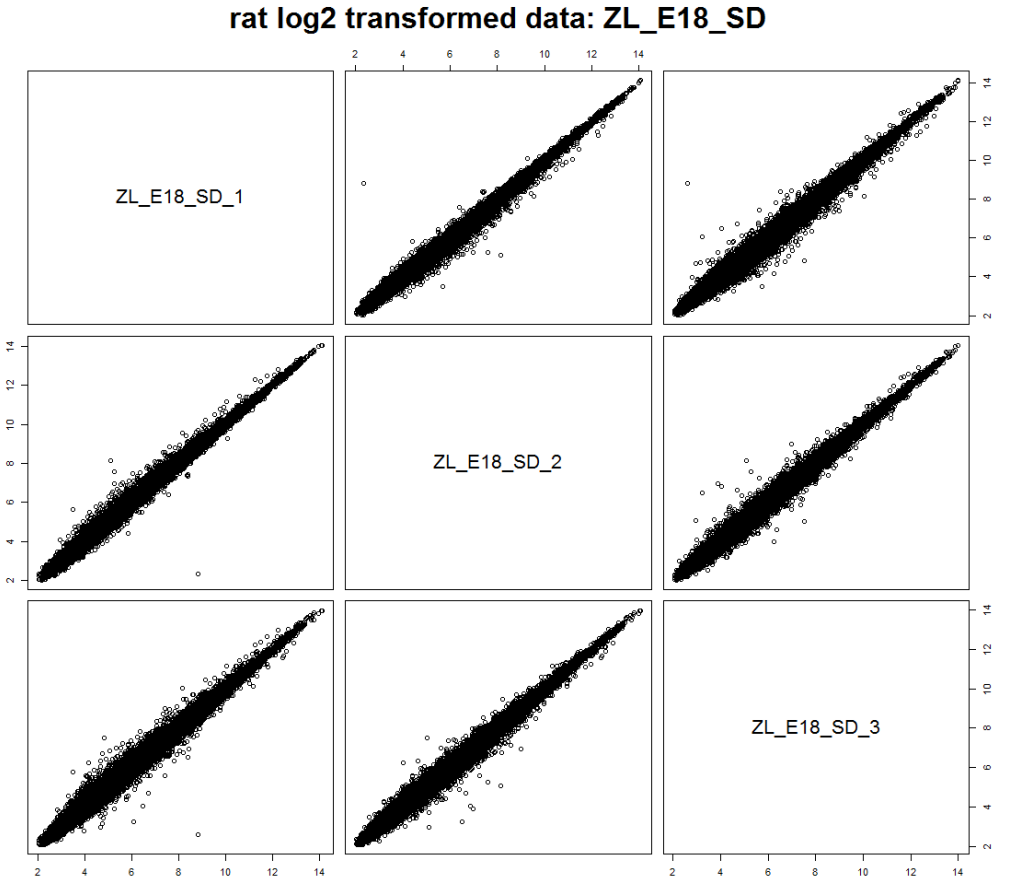
-> mukavemman näköinen data

Korkeimman 5% poisto – tarvitaanko

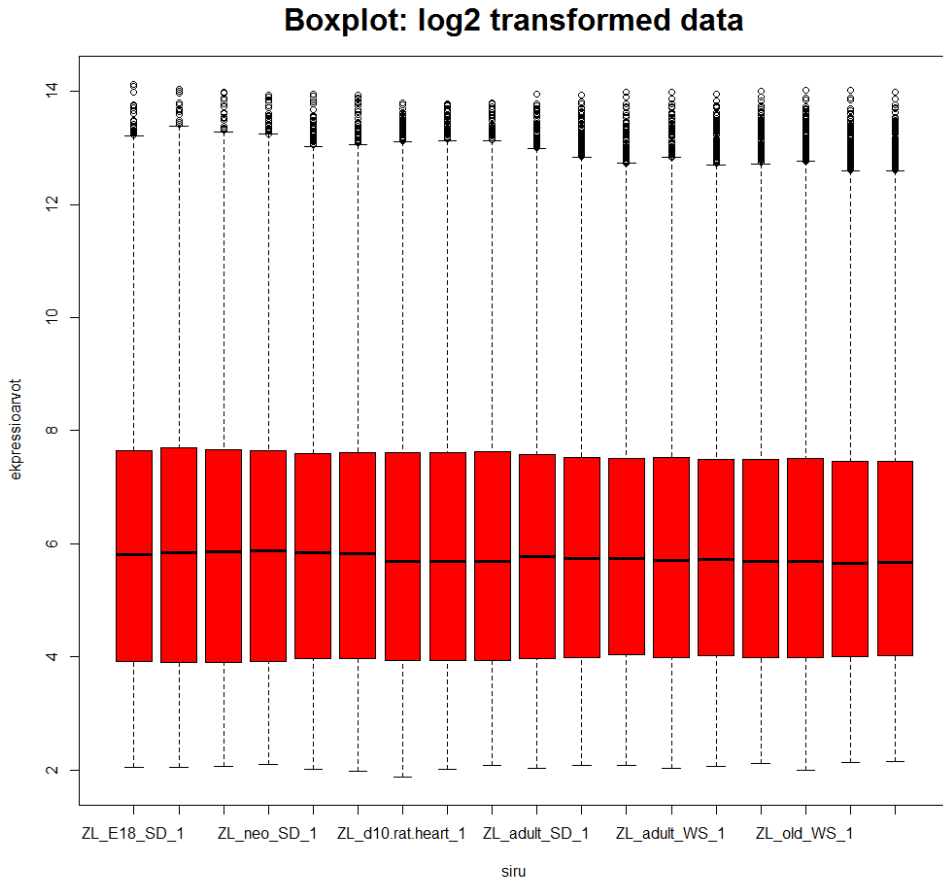
-laitteiston aiheuttamaa virhettä?

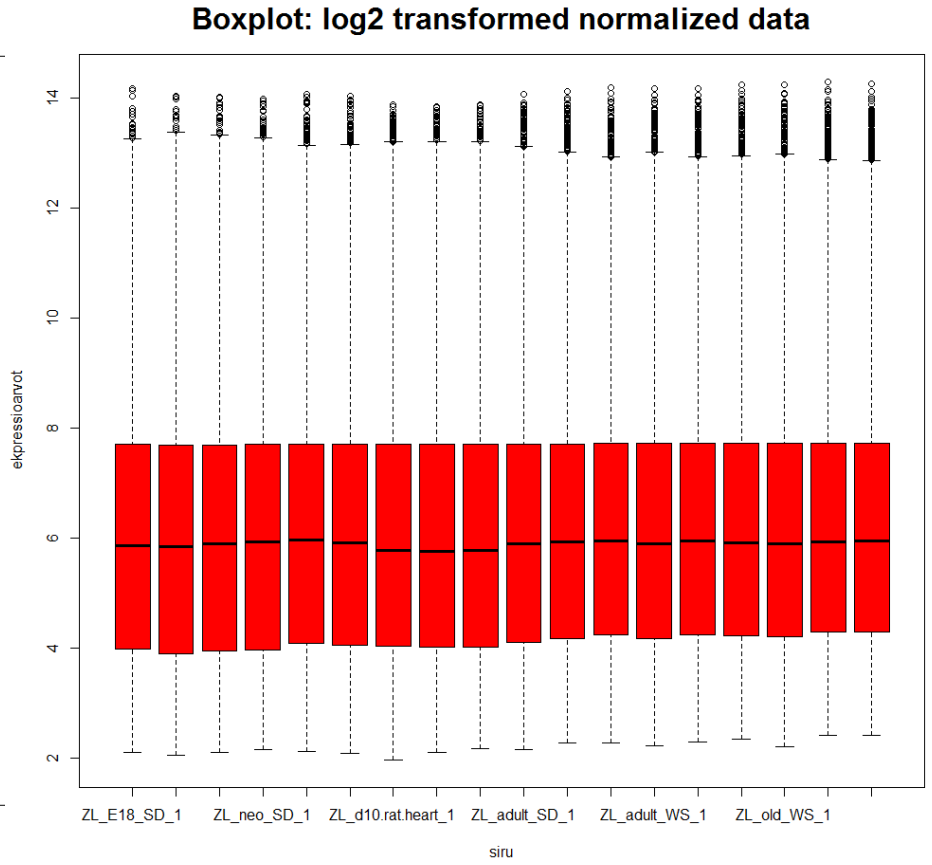
Datan esikäsittely:



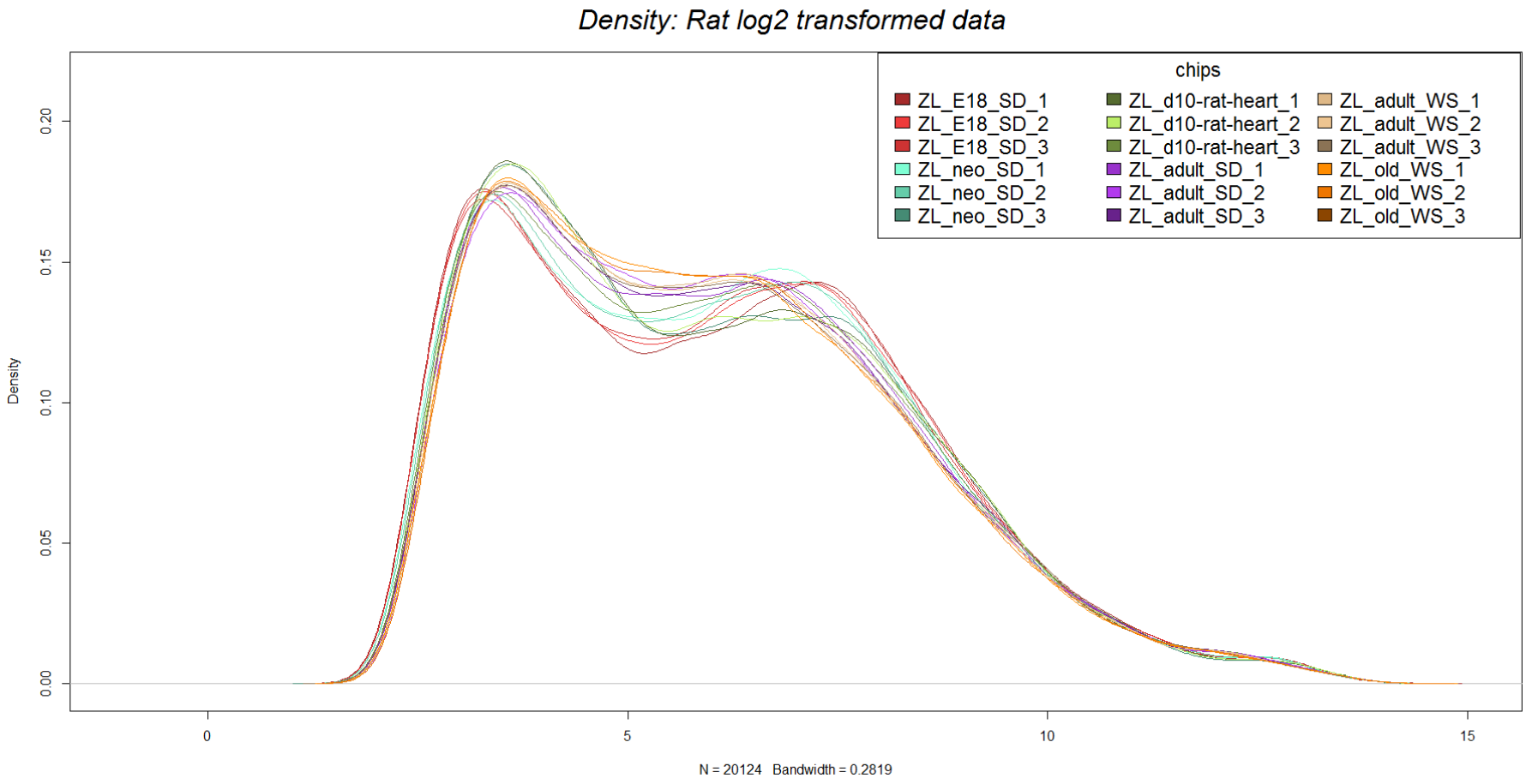


Datan esikäsittely: Normalisointi?

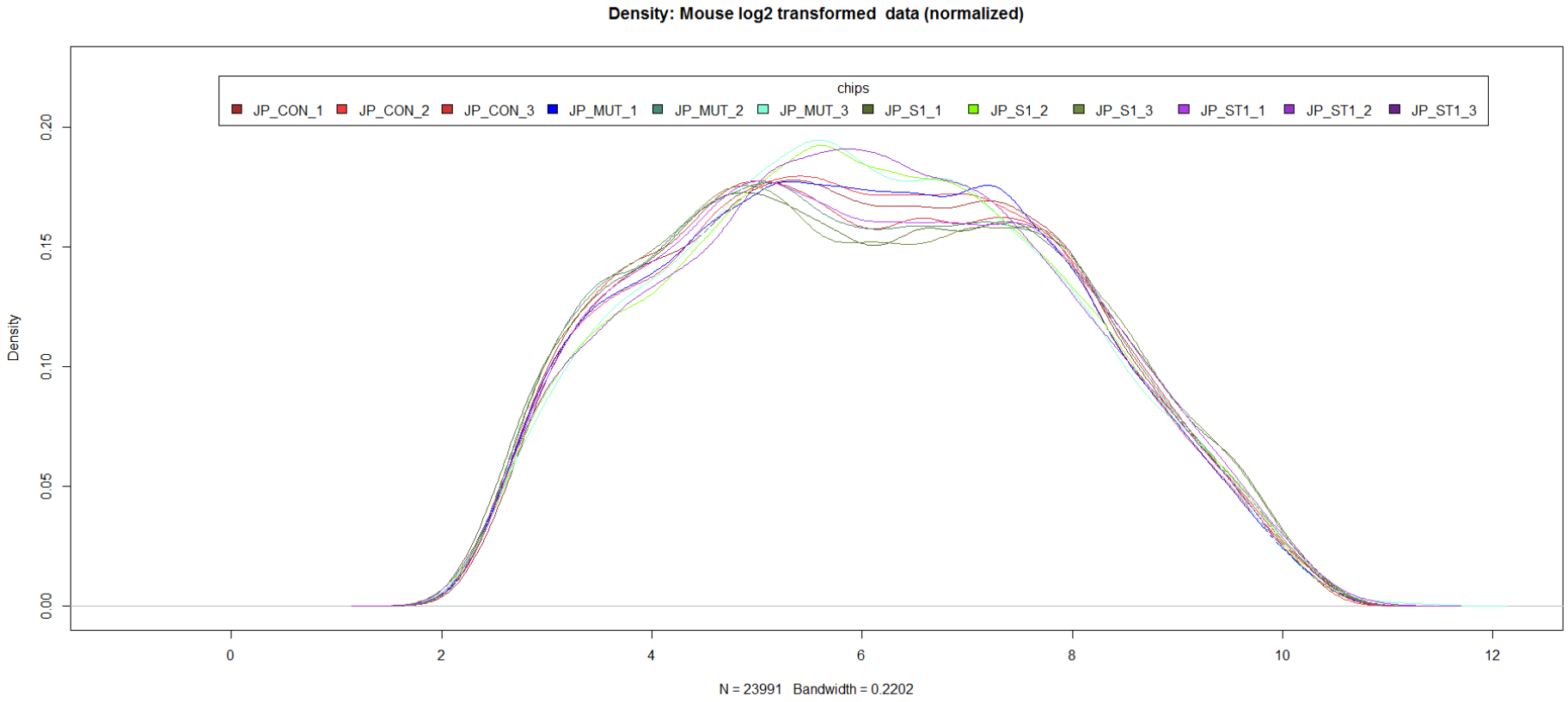




Datan esikäsittely: jakauma



Datan esikäsittely: jakauma



Alustavia tuloksia

Tilastolliset analyysit vasta alkaaneet

- Esitarkasteltu dataa, jota ei ole täysin esiprosessoitu
- GO-luokkien ja ekspressiotasojen välisiä korrelaatioita ei vielä todettu hiiridatalla (Kingfisher ohjelmalla)

Loppupäätelmät

- Työläintä oli ymmärtää itse ongelmaa samalla tavalla muiden kanssa
- Datassa oli paljon esikäsiteltävää
- Data on hyvää -> kaikki replikaatit saadaan valittua analyysiin.
- Myöhään löydetyt hiiren datassa olevat identtiset rivikopiot viivästyttivät analyysia