

Transcriptome-wide mapping of signaling pathways and early immune responses in lumpfish leukocytes upon *in vitro* bacterial exposure

Håvard Ø. Eggestøl¹, Harald S. Lunde¹, Anita Rønneseth¹, David Fredman², Kjell Petersen², Charitra K. Mishra², Tomasz Furmanek², Duncan J. Colquhoun^{1,3}, Heidrun I. Wergeland¹ and Gyri T. Haugland^{1*}.

¹Department of Biology, University of Bergen, Bergen, Norway

²Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway

³Norwegian Veterinary Institute, Oslo, Norway

* Gyri.Haugland@bio.uib.no

Supplementary materials:

Supplemental Table 1. Overview of the most significantly regulated genes at 6 and 24 hpe.....	2
Supplemental Table 2. Overview of the 50 most up-regulated genes at 24 hpe (sorted by logFC).....	4
Supplemental Table 3. Overview of the 50 most down-regulated genes at 24 hpe (sorted by logFC).....	6
Supplemental Table 4. Overview of the 50 most up-regulated genes at 6 hpe (sorted by logFC).....	8
Supplemental Table 5. Overview of the 50 most down-regulated genes at 6 hpe (sorted by logFC).....	10
Supplemental Table 6. Overview of verified lumpfish genes in the complement cascade.....	12
Supplemental Table 7. Overview of the species included in the phylogenetic analyses.....	14
Supplemental Table 8: Accession numbers of genes in the phylogenetic analyses of TLRs	16
Supplementary Methods for Assembly and Annotation.....	25
Supplementary Figure 1.....	26
Supplementary Methods for Bioinformatic scripts.....	27
Supplementary Results of Trinity RSEM.....	51

Supplemental Table 1. Overview of the most significantly regulated genes at 6 and 24 hpe (sorted by adjpval) *

Most significantly regulated genes at 24 hpe						
gene_id	logFC_24h	adjpval_24h	Annotations	Full record name	Species	E-value
TR41627	6.90	9.11E-11	TLR5	Toll-like_receptor_5	<i>Homo sapiens</i>	6E-92
TR13890	6.73	1.16E-10	IL8	Interleukin-8-like	<i>Serola dumerili</i>	6E-21
TR87025	5.27	1.16E-10	Uncharacterized	PREDICTED: uncharacterized protein LOC108879818	<i>Lates calcarifer</i>	5E-73
TR41114	6.52	1.16E-10	C1QL3	Complement_C1q-like_protein_3	<i>Mus musculus</i>	2E-12
TR41155	4.19	1.16E-10	NFAT5	Nuclear_factor_of_activated_T-cells_5	<i>Mus musculus</i>	8E-23
TR53083	-5.90	1.16E-10	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	0
TR70021	-5.18	1.30E-10	SIAE	Sialate_O-acylesterase	<i>Mus musculus</i>	1E-19
TR22616	4.01	1.74E-10	HS74L	Heat_shock_70_kDa_protein_4L	<i>Homo sapiens</i>	5E-170
TR81788	4.50	2.43E-10	DDX24	ATP-dependent_R_helicase_DDX24	<i>Pongo</i>	4E-41
TR52303	3.06	2.53E-10	IL12R-B2	Interleukin-12 receptor subunit beta-2-like	<i>Notothenia coriiceps</i>	3E-141
TR69283	5.17	2.56E-10	Hypothetical	hypothetical protein Z043_122910	<i>Scleropages formosus</i>	9E-05
TR79163	7.34	2.67E-10	TRFE	Serotransferrin	<i>Oryzias latipes</i>	0
TR8937	3.43	2.67E-10	CEP192	PREDICTED: centrosomal protein of 192 kDa	<i>Larimichthys crocea</i>	0E+00
TR71746	4.12	2.67E-10	NFAT5	Nuclear_factor_of_activated_T-cells_5	<i>Homo sapiens</i>	3E-58
TR70589	4.88	2.67E-10	PTGES	Prostaglandin_E_synthase	<i>Equus caballus</i>	4E-28
TR37961	4.42	2.67E-10	TAF4B	Transcription_initiation_factor_TFIID_subunit_4B	<i>Mus musculus</i>	6E-19
TR81776	8.77	2.83E-10	CC chemokine 20	C-C motif chemokine 20 precursor	<i>Anoplopoma fimbria</i>	8E-45
TR14360	8.27	2.83E-10	IL1B	Interleukin-1_beta	<i>Oncorhynchus mykiss</i>	9E-73
TR8934	3.21	2.83E-10	MAST3	Microtubule-associated_serine/threonine-protein_kinase_3	<i>Xenopus laevis</i>	4E-107
TR65236	3.68	2.83E-10	TIMP2	Metalloproteinase_inhibitor_2	<i>Cavia porcellus</i>	7E-109
Most significantly regulated genes at 6 hpe						
gene_id	logFC_6h	adjpval_6h	Annotations	Full record name	Specie	E-value
TR31519	4.31	8.49E-10	NR4A1	Nuclear_receptor_subfamily_4_group_A_member_1	<i>Xenopus laevis</i>	6E-52
TR13890	6.39	9.31E-10	IL8	Interleukin-8-like	<i>Serola dumerili</i>	6E-21
TR87025	5	1E-09	Uncharacterized	PREDICTED: uncharacterized protein LOC108879818	<i>Lates calcarifer</i>	5E-73
TR41627	4.28	2.90E-09	TLR5	Toll-like_receptor_5	<i>Homo sapiens</i>	6E-92
TR52303	2.66	2.90E-09	IL12R-B2	Interleukin-12 receptor subunit beta-2-like	<i>Notothenia coriiceps</i>	3E-141
TR69283	4.27	2.90E-09	Hypothetical protein	hypothetical protein Z043_122910	<i>Scleropages formosus</i>	9E-05
TR81776	6.17	2.90E-09	CC chemokine 20	C-C motif chemokine 20 precursor	<i>Anoplopoma fimbria</i>	8E-45
TR28159	3.15	2.90E-09	PTGIS	Prostacyclin_synthase	<i>Bos taurus</i>	3E-30
TR86119	4.40	3.17E-09	TNIP2	TNFAIP3-interacting_protein_2	<i>Mus musculus</i>	2E-31
TR19609	3.24	4.21E-09	IKBA	NF-kappa-B_inhibitor_alpha	<i>Gallus gallus</i>	6E-57
TR49945	5.26	7.93E-09	GL.	Glutamine_synthetase, mitochondrial	<i>Squalus acanthias</i>	6E-22
TR87414	5.19	7.93E-09	GL.	Glutamine_synthetase	<i>Canis lupus</i>	2E-147
TR80474	2.48	7.93E-09	SIK2	Serine/threonine-protein_kinase_SIK2	<i>Mus musculus</i>	5E-143
TR59888	2.70	7.93E-09	unnamed protein	Unnamed protein product	<i>Tetraodon nigroviridis</i>	1.4E-01
TR129039	3.61	7.93E-09	HSP 90-beta-3	Putative heat shock protein HSP 90-beta-3	<i>Tupaia chinensis</i>	4E-05
TR1049	1.83	7.93E-09	DDX41	Probable_ATP-dependent_R_helicase_DDX41	<i>Mus musculus</i>	0

TR59631	2.61	7.93E-09	FOS	Proto-oncogene_c-Fos	<i>Takifugu rubripes</i>	4E-100
TR81788	3.26	7.93E-09	DDX24	ATP-dependent_R_helicase_DDX24	<i>Pongo</i>	4E-41
TR14360	6.69	7.93E-09	IL1B	Interleukin-1_beta	<i>Oncorhynchus mykiss</i>	9E-73
TR8934	2.65	7.93E-09	MAST3	Microtubule-associated_serine/threonine-protein_kinase_3	<i>Xenopus laevis</i>	4E-107

* Positive logFC values= upregulated transcripts, negative logFC values=down-regulated transcripts

Supplemental Table 2. Overview of the 50 most up-regulated genes at 24 hpe (sorted by logFC)

50 most upregulated genes at 24 hpe (sorted by logFC)						
gene_id	logFC_24h	adjpval_24h	Annotations	Full record name	Specie	E-value
TR26702 c0_g1	9.50	1.97E-06	N42L1	NEDD4-binding_protein_2-like_1	<i>Mus musculus</i>	3E-07
TR78231 c0_g1	9.07	1.21E-04	GDF15	Growth/differentiation_factor_15	<i>Mus musculus</i>	2E-24
TR81776 c4_g4	8.77	2.83E-10	CC chemokine 20	C-C motif chemokine 20 precursor	<i>Anoplopoma fimbria</i>	8E-45
TR74252 c0_g1	8.68	1.11E-07	CFAH	Complement_factor_H	<i>Homo sapiens</i>	2E-34
TR45367 c1_g1	8.35	1.85E-07	HPT	Haptoglobin	<i>Sus scrofa</i>	3E-40
TR14360 c3_g2	8.27	2.83E-10	IL1B	Interleukin-1_beta	<i>Oncorhynchus mykiss</i>	9E-73
TR30315 c0_g1	8.05	3.91E-07	CO8A	Complement component C8_alpha_chain	<i>Oryctolagus cuniculus</i>	1E-157
TR23708 c0_g1	8.03	1.17E-07	NMES1	Normal_mucosa_of_esophagus-specific_gene_1_protein	<i>Homo sapiens</i>	2E-23
TR3853 c0_g2	7.83	3.67E-06	BORG1	Cdc42_effector_protein_2	<i>Homo sapiens</i>	4E-06
TR27293 c0_g1	7.67	9.05E-07	CO8B	Complement component C8_beta_chain	<i>Paralichthys olivaceus</i>	0
TR80849 c0_g2	7.59	2.06E-08	SAA1	Serum_amyloid_A-1_protein	<i>Mus musculus</i>	3E-11
TR76870 c0_g2	7.53	7.24E-07	EPGN	Epigen	<i>Gallus gallus</i>	1E-09
TR67971 c1_g1	7.49	3.57E-06	LPP	Lipoma-preferred_partner_homolog	<i>Gallus gallus</i>	1E-28
TR82903 c2_g2	7.42	6.94E-06	PHOS	Phosducin	<i>Bos taurus</i>	1E-65
TR79163 c0_g5	7.34	2.67E-10	TRFE	Serotransferrin	<i>Orizias latipes</i>	0
TR80347 c0_g1	7.17	2.83E-10	HEMO	Hemopexin_{ECO:0000250 UniProtKB:P20058}	<i>Danio rerio</i>	9E-46
TR45367 c1_g2	7.14	4.55E-05	HPT	Haptoglobin	<i>Rattus norvegicus</i>	1E-19
TR11634 c0_g2	7.10	4.82E-06	GAB	Alpha-N-acetylgalactosaminidase	<i>Gallus gallus</i>	3E-35
TR87818 c0_g1	7.10	4.16E-10	IL6	Interleukin-6	<i>Paralichthys olivaceus</i>	7E-70
TR42629 c0_g1	7.07	5.14E-06	INHBB	Inhibin_beta_B_chain	<i>Gallus gallus</i>	1E-77
TR13042 c4_g6	7.04	8.57E-09	DPYD	Dihydropyrimidine_dehydrogenase_[.DP(+)]	<i>Danio rerio</i>	5E-38
TR14183 c0_g1	6.91	4.16E-10	RET7	Retinoid-binding_protein_7	<i>Mus musculus</i>	5E-54
TR41627 c0_g1	6.90	9.11E-11	TLR5	Toll-like_receptor_5	<i>Homo sapiens</i>	6E-92
TR13890 c0_g3	6.73	1.16E-10	IL8	interleukin-8-like	<i>Serola dumerili</i>	6E-21
TR70589 c2_g2	6.70	1.97E-04	PTGES	Prostaglandin_E_synthase	<i>Equus ferus</i>	1E-22
TR76899 c0_g1	6.63	6.22E-06	IL17F	Interleukin-17F	<i>Rattus norvegicus</i>	1E-11
TR28096 c4_g11	6.57	9.52E-06	GRM4	Metabotropic_glutamate_receptor_4	<i>Rattus norvegicus</i>	7E-63
TR41114 c0_g2	6.52	1.16E-10	C1QL3	Complement_C1q-like_protein_3	<i>Mus musculus</i>	2E-12
TR41089 c3_g7	6.52	4.31E-06	MORC2	MORC_family_CW-type_zinc_finger_protein_2	<i>Homo sapiens</i>	2E-27
TR32785 c0_g2	6.42	6.87E-07	A2GL	Leucine-rich_alpha-2-glycoprotein	<i>Homo sapiens</i>	7E-46
TR13469 c0_g1	6.33	1.47E-05	MA7D2	MAP7_domain-containing_protein_2	<i>Pongo sp.</i>	5E-08
TR47931 c1_g1	6.30	2.65E-07	MOT4	Monocarboxylate_transporter_4	<i>Gallus gallus</i>	2E-149
TR31477 c0_g1	6.28	7.76E-06	NDF4	Neurogenic_differentiation_factor_4	<i>Homo sapiens</i>	7E-68

TR74574 c0_g2	6.25	2.64E-05	RGS5	Regulator_of_G-protein_signaling_5	<i>Rattus norvegicus</i>	1E-26
TR8953 c5_g4	6.22	6.56E-08	CO5	Complement_C5	<i>Homo sapiens</i>	0
TR48860 c0_g1	6.12	3.43E-04	EMAL6	Echinoderm_microtubule-associated_protein-like_6	<i>Homo sapiens</i>	1E-98
TR45357 c0_g2	6.08	2.31E-09	TGM2	Protein-glutamine_gamma-glutamyltransferase_2	<i>Pagrus major</i>	4E-153
TR38005 c0_g2	6.07	2.99E-07	HME2A	Homeobox_protein_engrailed-2-A	<i>Xenopus laevis</i>	5E-61
TR64331 c3_g1	6	4.16E-10	DPYD	Dihydropyrimidine_dehydrogenase_[.DP(+)]	<i>Danio rerio</i>	2E-49
TR81910 c0_g1	5.89	7.57E-09	CHSP1	Calcium-regulated_heat_stable_protein_1	<i>Rattus norvegicus</i>	2E-19
TR45395 c0_g1	5.85	3.36E-05	SPIC	Transcription_factor_Spi-C	<i>Bos taurus</i>	1E-34
TR78110 c0_g1	5.83	9.67E-05	CFAB	Complement_factor_B	<i>Bos taurus</i>	3E-63
TR57748 c0_g1	5.81	1.28E-05	KC.3	Potassium_voltage-gated_channel_subfamily_A_member_3	<i>Mus musculus</i>	0
TR71513 c0_g2	5.80	7.24E-05	S6A17	Sodium-dependent_neutral_amino_acid_transporter_SLC6A17	<i>Mus musculus</i>	1E-65
TR42615 c0_g1	5.63	3.76E-05	TRI55	Tripartite_motif-containing_protein_55	<i>Rattus norvegicus</i>	2E-37
TR10806 c0_g1	5.58	2.38E-05	GRAM3	GRAM_domain-containing_protein_3	<i>Rattus norvegicus</i>	5E-10
TR145009 c0_g1	5.55	8.80E-05	SIA4B	CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase_2	<i>Homo sapiens</i>	3E-118
TR78132 c0_g1	5.51	8.84E-05	FABP7	Fatty_acid-binding_protein,_brain	<i>Mus musculus</i>	9E-08
TR23250 c0_g1	5.50	2.12E-05	FXYP6	FXYP_domain-containing_ion_transport_regulator_6	<i>Mus musculus</i>	6E-22

Supplemental Table 3. Overview of the 50 most down-regulated genes at 24 hpe (sorted by logFC)

50 most downregulated genes at 24 hpe (sorted by logFC)						
gene_id	logFC_24h	adjpval_24h	Annotations	Full record name	Specie	E-value
TR41171 c0_g1	-7.56	2.02E-06	LRC3B	Leucine-rich_repeat-containing_protein_3B	<i>Homo sapiens</i>	9E-65
TR70719 c4_g4	-7.46	4.49E-07	NSMA2	Sphingomyelin_phosphodiesterase_3	<i>Mus musculus</i>	3E-85
TR45504 c0_g2	-7.30	1.59E-07	S43A3	Solute_carrier_family_43_member_3	<i>Homo sapiens</i>	2E-72
TR51281 c0_g3	-7.26	1.44E-06	TGM2	Protein-glutamine_gamma-glutamyltransferase_2	<i>Gallus gallus</i>	8E-80
TR53083 c0_g1	-7.17	5.90E-06	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	2E-39
TR15942 c0_g2	-7.09	5.77E-07	DMBT1	Deleted_in_malignant_brain_tumors_1_protein	<i>Homo sapiens</i>	1E-73
TR80906 c0_g1	-7	1.21E-05	IL17F	Interleukin-17F	<i>Rattus norvegicus</i>	4E-11
TR31532 c0_g1	-6.74	1.13E-05	KCJ12	ATP-sensitive_inward_rectifier_potassium_channel_12	<i>Gallus gallus</i>	0
TR68049 c0_g1	-6.56	9.71E-05	CNKR2	Connector_enhancer_of_kinase_suppressor_of_ras_2	<i>Homo sapiens</i>	0
TR3259 c0_g1	-6.53	3.10E-06	TNR3	Tumor_necrosis_factor_receptor_superfamily_member_3	<i>Homo sapiens</i>	8E-09
TR53083 c0_g2	-6.45	2.87E-05	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	1E-39
TR65297 c3_g1	-6.44	1.44E-06	TMM88	Transmembrane_protein_88	<i>Mus musculus</i>	2E-12
TR69288 c0_g1	-6.36	9.32E-07	TNR11	Tumor_necrosis_factor_receptor_superfamily_member_11A	<i>Mus musculus</i>	9E-13
TR28894 c2_g1	-6.28	5.44E-06	COR2A	Coronin-2A	<i>Homo sapiens</i>	0
TR2247 c0_g1	-6.23	2.19E-06	RLBP1	Retinaldehyde-binding_protein_1	<i>Bos taurus</i>	2E-158
TR144820 c0_g2	-6.10	2.97E-05	PRDM1	PR_domain_zinc_finger_protein_1	<i>Mus musculus</i>	2E-26
TR6306 c0_g2	-6.03	1.49E-03	EVA1A	Protein_eva-1_homolog_A	<i>Danio rerio</i>	5E-37
TR87113 c1_g2	-6.01	9.36E-04	CEBPD	CCAAT/enhancer-binding_protein_delta	<i>Homo sapiens</i>	2E-56
TR53083 c1_g5	-5.90	1.16E-10	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	0
TR80012 c0_g1	-5.87	5.45E-08	S12A4	Solute_carrier_family_12_member_4	<i>Homo sapiens</i>	0
TR49079 c0_g3	-5.75	3.21E-06	MITF	Microphthalmia-associated_transcription_factor	<i>Homo sapiens</i>	1E-67
TR65368 c1_g21	-5.73	6.74E-05	ANKR1	Ankyrin_repeat_domain-containing_protein_1	<i>Gallus gallus</i>	4E-49
TR26688 c3_g1	-5.72	1.34E-02	ENOG	Gamma-enolase	<i>Rattus norvegicus</i>	2E-133
TR40542 c0_g1	-5.68	7.31E-07	HECA2	HEPACAM_family_member_2	<i>Homo sapiens</i>	1E-06
TR37184 c1_g3	-5.62	2.73E-04	FDXA1	Ferredoxin-fold_anticonodon-binding_domain-containing_protein_1_homolog	<i>Mus musculus</i>	2E-27
TR8782 c0_g1	-5.56	3.30E-05	EPD2	Ependymin-2	<i>Carassius auratus</i>	3E-22
TR68285 c0_g1	-5.50	5.35E-05	DTBP1	Dysbindin	<i>Gallus gallus</i>	6E-31
TR880 c0_g2	-5.48	2.73E-05	MAP6	Microtubule-associated_protein_6_homolog	<i>Gallus gallus</i>	1E-31
TR71817 c0_g1	-5.44	2.54E-05	TGFA1	Transforming_growth_factor-beta_receptor-associated_protein_1	<i>Mus musculus</i>	1E-123
TR101 c0_g1	-5.41	1.04E-06	VAT1	Synaptic vesicle membrane protein VAT-1 homolog	<i>Danio rerio</i>	0
TR34008 c0_g1	-5.33	2.52E-07	SEM4G	Semaphorin-4G	<i>Homo sapiens</i>	0
TR44282 c0_g1	-5.32	3.20E-06	PHEX	Metalloendopeptidase_homolog_PEX	<i>Mus musculus</i>	0

TR43339 c0_g1	-5.31	1.36E-05	NFIL3	Nuclear_factor_interleukin-3-regulated_protein	<i>Rattus norvegicus</i>	2E-17
TR62246 c0_g1	-5.25	2.81E-05	JIP1	C-Jun-amino-terminal_kinase-interacting_protein_1	<i>Homo sapiens</i>	2E-141
TR14442 c0_g1	-5.21	6.43E-05	TLR13	Toll-like_receptor_13	<i>Mus musculus</i>	3E-66
TR54313 c0_g2	-5.19	5.12E-06	FABPL	Fatty_acid-binding_protein, liver	<i>Ginglymostoma cirratum</i>	2E-14
TR5820 c0_g1	-5.14	1.30E-04	SREC	Scavenger_receptor_class_F_member_1	<i>Homo sapiens</i>	9E-37
TR85479 c0_g2	-5.14	7.80E-05	S29A1	Equilibrative_nucleoside_transporter_1	<i>Rattus norvegicus</i>	3E-42
TR78540 c0_g2	-5.11	9.48E-05	PERF	Perforin-1	<i>Rattus norvegicus</i>	8E-129
TR9527 c0_g1	-4.99	3.57E-06	NUAK1	NUAK_family_SNF1-like_kinase_1	<i>Homo sapiens</i>	0
TR65997 c0_g1	-4.98	1.54E-05	EFNB2	Ephrin-B2a	<i>Danio rerio</i>	5E-40
TR156627 c0_g3	-4.96	8.03E-03	CBP	CREB-binding_protein	<i>Rattus norvegicus</i>	1E-08
TR40644 c0_g2	-4.95	1.74E-04	RGS5	Regulator_of_G-protein_signaling_5	<i>Rattus norvegicus</i>	7E-55
TR38943 c0_g1	-4.93	3.66E-04	FGF4	Fibroblast_growth_factor_4	<i>Gallus gallus</i>	6E-09
TR6306 c0_g5	-4.90	4.04E-05	EVA1A	Protein_eva-1_homolog_A	<i>Danio rerio</i>	4E-37
TR25290 c0_g4	-4.90	1.61E-04	KCNH6	Potassium_voltage-gated_channel_subfamily_H_member_6	<i>Rattus norvegicus</i>	2E-83
TR12896 c0_g2	-4.85	1.08E-03	AMPN	Aminopeptidase_N_{ECO:0000312 EMBL:ACZ95799.1}	<i>Gallus gallus</i>	4E-141
TR1713 c0_g2	-4.84	2.41E-05	APOA4	Apolipoprotein_A-IV	<i>Papio anubis</i>	2E-28
TR56616 c0_g1	-4.72	2.88E-04	TGM2	Protein-glutamine_gamma-glutamyltransferase_2	<i>Homo sapiens</i>	3E-145

Supplemental Table 4. Overview of the 50 most up-regulated genes at 6 hpe (sorted by logFC)

50 most upregulated genes at 6 hpe (sorted by logFC)						
gene_id	logFC_6h	adjpval_6h	Annotations	Full record name	Specie	E-value
TR76899 c0_g1	9.37	7.35E-06	IL17F	Interleukin-17F	<i>Rattus norvegicus</i>	1E-11
TR26702 c0_g1	9.05	8.57E-06	N42L1	NEDD4-binding_protein_2-like_1	<i>Mus musculus</i>	3E-07
TR20674 c0_g2	7.73	4.87E-06	HYAS1	Hyaluronan_synthase_1	<i>Xenopus laevis</i>	0
TR45366 c0_g1	7.13	2.38E-05	ZN648	Zinc_finger_protein_648	<i>Homo sapiens</i>	7E-112
TR78231 c0_g1	6.87	1.18E-03	GDF15	Growth/differentiation_factor_15	<i>Mus musculus</i>	2E-24
TR32785 c0_g2	6.74	7.21E-05	A2GL	Leucine-rich_alpha-2-glycoprotein	<i>Homo sapiens</i>	7E-46
TR14360 c3_g2	6.69	7.93E-09	IL1B	Interleukin-1_beta	<i>Oncorhynchus mykiss</i>	9E-73
TR87818 c0_g1	6.51	1.62E-07	IL6	Interleukin-6	<i>Paralichthys olivaceus</i>	7E-70
TR13890 c0_g3	6.39	9.31E-10	IL8	interleukin-8-like	<i>Serola dumerili</i>	6E-21
TR81776 c4_g4	6.17	2.90E-09	CC chemokine 20	C-C motif chemokine 20 precursor	<i>Anoplopoma fimbria</i>	8E-45
TR76870 c0_g2	5.85	4.53E-07	EPGN	Epigen	<i>Gallus gallus</i>	1E-09
TR45395 c0_g1	5.60	4.71E-04	SPIC	Transcription_factor_Spi-C	<i>Bos taurus</i>	1E-34
TR19933 c0_g1	5.47	1.69E-08	GL	Glutamine_synthetase	<i>Acomys cahirinus</i>	4E-109
TR67971 c1_g1	5.46	5.24E-05	LPP	Lipoma-preferred_partner_homolog	<i>Gallus gallus</i>	1E-28
TR74252 c0_g1	5.40	9.82E-06	CFAH	Complement_factor_H	<i>Homo sapiens</i>	2E-34
TR49945 c0_g1	5.26	7.93E-09	GL	Glutamine_synthetase_mitochondrial	<i>Squalus acanthias</i>	6E-22
TR87414 c0_g1	5.19	7.93E-09	GL	Glutamine_synthetase	<i>Canis familiaris</i>	2E-147
TR44307 c0_g2	5.17	2.60E-04	VAX2B	Ventral_anterior_homeobox_2b	<i>Xenopus laevis</i>	5E-63
TR27410 c1_g1	5.15	8.04E-07	NOXO1	.DPH_oxidase_organizer_1	<i>Mus musculus</i>	4E-59
TR81910 c0_g1	5.13	1.67E-06	CHSP1	Calcium-regulated_heat_stable_protein_1	<i>Rattus norvegicus</i>	2E-19
TR66238 c0_g2	5.08	1.50E-04	CN159	UPF0317_protein_C14orf159_homolog_mitochondrial	<i>Mus musculus</i>	2E-15
TR87025 c0_g2	5	1E-09	Uncharacterized protein	PREDICTED: uncharacterized protein LOC108879818	<i>Lates calcarifer</i>	5E-73
TR68048 c0_g1	5	2.42E-04	GRAM3	GRAM_domain-containing_protein_3	<i>Rattus norvegicus</i>	2E-10
TR14183 c0_g1	4.94	6.67E-08	RET7	Retinoid-binding_protein_7	<i>Mus musculus</i>	5E-54
TR11634 c0_g2	4.91	4E-06	.GAB	Alpha-N-acetylgalactosaminidase	<i>Gallus gallus</i>	3E-35
TR61339 c0_g1	4.85	6.62E-04	ISL2A	Insulin_gene_enhancer_protein_isl-2a	<i>Danio rerio</i>	0
TR80347 c0_g1	4.70	1.70E-07	HEMO	Hemopexin_{ECO:0000250 UniProtKB:P20058}	<i>Danio rerio</i>	9E-46
TR34004 c3_g1	4.50	7.93E-09	TNIP2	TNFAIP3-interacting_protein_2	<i>Mus musculus</i>	5E-20
TR78134 c0_g1	4.46	2.34E-03	SSR2	Somatostatin_receptor_type_2	<i>Sus scrofa</i>	3E-118
TR44039 c0_g1	4.41	4.89E-03	ETV6	Transcription_factor_ETV6	<i>Homo sapiens</i>	5E-20
TR86119 c0_g1	4.40	3.17E-09	TNIP2	TNFAIP3-interacting_protein_2	<i>Mus musculus</i>	2E-31
TR69814 c0_g2	4.38	2.80E-08	TNFA	Tumor_necrosis_factor	<i>Sparus auratus</i>	3E-114
TR45367 c1_g2	4.35	3.78E-03	HPT	Haptoglobin	<i>Rattus norvegicus</i>	1E-19
TR31519 c0_g1	4.31	8.49E-10	NR4A1	Nuclear_receptor_subfamily_4_group_A_member_1	<i>Xenopus laevis</i>	6E-52

TR45033 c0_g1	4.30	2.83E-03	SSR2	Somatostatin_receptor_type_2	<i>Homo sapiens</i>	1E-62
TR27348 c0_g2	4.30	2.48E-02	IKBA	NF-kappa-B_inhibitor_alpha	<i>Sus scrofa</i>	5E-06
TR41627 c0_g1	4.28	2.90E-09	TLR5	Toll-like_receptor_5	<i>Homo sapiens</i>	6E-92
TR69283 c1_g1	4.27	2.90E-09	Hypothetical protein	hypothetical protein Z043_122910	<i>Scleropages formosus</i>	9E-05
TR57748 c0_g1	4.24	4.95E-04	KC.3	Potassium_voltage-gated_channel_subfamily_A_member_3	<i>Mus musculus</i>	0
TR31477 c0_g1	4.20	6.99E-04	NDF4	Neurogenic_differentiation_factor_4	<i>Homo sapiens</i>	7E-68
TR45357 c0_g2	4.18	3.24E-07	TGM2	Protein-glutamine_gamma-glutamyltransferase_2	<i>Pagrus major</i>	4E-153
TR57132 c0_g1	4.17	9.26E-03	LORF2	LINE-1_retrotransposable_element_ORF2_protein	<i>Homo sapiens</i>	3E-20
TR80347 c0_g2	4.12	1.69E-03	HEMO	Hemopexin_{ECO:0000250 UniProtKB:P20058}	<i>Danio rerio</i>	3E-14
TR80849 c0_g2	4.11	1.47E-03	SAA1	Serum_amyloid_A-1_protein	<i>Mus musculus</i>	3E-11
TR23312 c0_g1	4.10	1.10E-08	ADTRP	Androgen-dependent_TFPI-regulating_protein	<i>Homo sapiens</i>	4E-06
TR69281 c3_g2	4.08	4.79E-03	MTSS1	Metastasis_suppressor_protein_1	<i>Homo sapiens</i>	2E-31
TR64314 c0_g1	4.05	5.05E-05	S12A2	Solute_carrier_family_12_member_2	<i>Mus musculus</i>	1E-06
TR61516 c0_g2	3.95	3.33E-03	BIRC6	Baculoviral_IAP_repeat-containing_protein_6	<i>Homo sapiens</i>	4E-125
TR45367 c1_g1	3.94	4.97E-04	HPT	Haptoglobin	<i>Sus scrofa</i>	3E-40
TR41618 c1_g3	3.94	1.11E-04	PDK2	[Pyruvate_dehydrogenase_(acetyl-transferring)]_kinase_isozyme_2_mitochondrial	<i>Rattus norvegicus</i>	7E-09

Supplemental Table 5. Overview of the 50 most down-regulated genes at 6 hpe (sorted by logFC)

50 most downregulated genes at 6 hpe (sorted by logFC)						
gene_id	logFC_6h	adjpval_6h	Annotations	Full record name	Specie	E-value
TR145793 c0_g6	-6.30	2.63E-02	ZFX	Zinc_finger_X-chromosomal_protein	<i>Mus musculus</i>	9E-117
TR53083 c0_g2	-5.29	1.61E-03	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	1E-39
TR55536 c0_g2	-4.63	2.16E-02	CCR6	C-C_chemokine_receptor_type_6	<i>Homo sapiens</i>	2E-76
TR40542 c0_g1	-4.24	7.23E-04	HECA2	HEPACAM_family_member_2	<i>Homo sapiens</i>	1E-06
TR145793 c0_g1	-4.23	4.14E-02	ZFX	Zinc_finger_X-chromosomal_protein	<i>Mus musculus</i>	1E-116
TR85757 c0_g4	-3.68	6.06E-02	ZBT37	Zinc_finger_and_BTBDomain-containing_protein_37	<i>Homo sapiens</i>	2E-79
TR62246 c0_g1	-3.61	1.49E-03	JIP1	C-Jun-amino-terminal_kinase-interacting_protein_1	<i>Homo sapiens</i>	2E-141
TR150022 c0_g1	-3.56	4.70E-03	LPAR4	Lysophosphatidic_acid_receptor_4	<i>Mus musculus</i>	5E-17
TR53083 c0_g1	-3.52	1.60E-03	MRC1	Macrophage_mannose_receptor_1	<i>Mus musculus</i>	2E-39
TR71817 c0_g1	-3.48	2.46E-03	TGFA1	Transforming_growth_factor-beta_receptor-associated_protein_1	<i>Mus musculus</i>	1E-123
TR40529 c0_g3	-3.26	1.98E-02	JDP2	Jun_dimerization_protein_2	<i>Rattus norvegicus</i>	6E-57
TR57429 c0_g9	-3.24	9.04E-03	ABI3	ABI_gene_family_member_3	<i>Homo sapiens</i>	5E-10
TR16539 c0_g2	-3.17	5.35E-02	TSN8	Tetraspanin-8	<i>Bos taurus</i>	2E-07
TR36086 c0_g2	-3.08	3E-02	CXCR5	C-X-C_chemokine_receptor_type_5	<i>Rattus norvegicus</i>	1E-36
TR67966 c4_g1	-3.04	1.09E-01	UB2G2	Ubiquitin-conjugating_enzyme_E2_G2	<i>Pongo sp.</i>	8E-10
TR33782 c1_g4	-3.04	3.81E-06	MMP14	Matrix_metalloproteinase-14	<i>Mus musculus</i>	5E-138
TR67966 c4_g4	-3.02	6.91E-02	UB2G2	Ubiquitin-conjugating_enzyme_E2_G2	<i>Pongo sp.</i>	7E-10
TR633 c0_g3	-3.02	1.51E-01	ADAT2	tR.-specific_adenosine_deaminase_2	<i>Danio rerio</i>	1E-71
TR44789 c0_g1	-3	1.69E-01	KAD2	Adenylate_kinase_2_mitochondrial_{ECO:0000255 HAMAP-Rule:MF_03168}	<i>Salmo salar</i>	2E-15
TR34008 c0_g1	-2.93	3.81E-06	SEM4G	Semaphorin-4G	<i>Homo sapiens</i>	0
TR69449 c4_g10	-2.92	1.96E-02	RTXE	Probable_R.-directed_D._polymerase_from_transposon_X-element	<i>Drosophila melongaster</i>	2E-13
TR70501 c4_g2	-2.91	7.79E-03	ABI2	Abl_interactor_2	<i>Mus musculus</i>	2E-101
TR65297 c3_g1	-2.88	5.13E-04	TMM88	Transmembrane_protein_88	<i>Mus musculus</i>	2E-12
TR25793 c0_g1	-2.85	2.12E-02	FHOD3	FH1/FH2_domain-containing_protein_3	<i>Mus musculus</i>	8E-22
TR2196 c0_g2	-2.83	3.60E-02	WDR18	WD_repeat-containing_protein_18	<i>Danio rerio</i>	1E-76
TR35910 c0_g2	-2.77	1.18E-01	ADA1D	Alpha-1D_adrenergic_receptor	<i>Sus scrofa</i>	5E-36
TR49070 c0_g1	-2.75	5.86E-05	SPTB2	Spectrin_beta_chain_non-erythrocytic_1	<i>Mus musculus</i>	3E-13
TR84473 c0_g1	-2.74	1.25E-02	TLR2	Toll-like_receptor_2	<i>Cricetulus griseus</i>	3E-15
TR87042 c0_g6	-2.73	1.65E-01	ADCK1	Uncharacterized_aarF_domain-containing_protein_kinase_1	<i>Homo sapiens</i>	8E-09
TR27448 c0_g3	-2.72	3.35E-01	SC6A8	Sodium-_and_chloride-dependent_creatine_transporter_1	<i>Rattus norvegicus</i>	0
TR15942 c0_g2	-2.70	4.75E-06	DMBT1	Deleted_in_malignant_brain_tumors_1_protein	<i>Homo sapiens</i>	1E-73
TR25104 c0_g5	-2.68	1.95E-01	FND3A	Fibronectin_type-III_domain-containing_protein_3A	<i>Homo sapiens</i>	3E-13
TR85479 c0_g2	-2.66	4.34E-03	S29A1	Equilibrative_nucleoside_transporter_1	<i>Rattus norvegicus</i>	3E-42
TR84456 c0_g3	-2.64	1.85E-01	PSPC1	Paraspeckle_component_1	<i>Danio rerio</i>	2E-134

TR65271 c0_g1	-2.63	1.26E-02	VINC	Vinculin	<i>Xenopus laevis</i>	5E-15
TR60099 c0_g2	-2.63	1.16E-01	TRIM3	Tripartite_motif-containing_protein_3	<i>Mus musculus</i>	2E-94
TR24328 c0_g1	-2.61	2.38E-03	TM158	Transmembrane_protein_158	<i>Mus musculus</i>	7E-63
TR85510 c0_g3	-2.60	7.17E-03	F264	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase_4	<i>Mus musculus</i>	6E-40
TR74312 c1_g7	-2.58	3.17E-03	PI42A	Phosphatidylinositol_5-phosphate_4-kinase_type-2_alpha	<i>Gallus gallus</i>	6E-111
TR43358 c0_g1	-2.55	5.16E-03	I22R2	Interleukin-22_receptor_subunit_alpha-2	<i>Mus musculus</i>	2E-22
TR65255 c0_g11	-2.48	3.49E-03	LORF2	LINE-1_retrotransposable_element_ORF2_protein	<i>Mus musculus</i>	2E-46
TR25130 c0_g2	-2.46	1.92E-02	AQP8	Aquaporin-8	<i>Homo sapiens</i>	2E-17
TR22880 c0_g3	-2.46	1.95E-01	MED23	Mediator_of_R._polymerase_II_transcription_subunit_23	<i>Danio rerio</i>	0
TR41097 c0_g1	-2.45	2.94E-03	MAGI1	Membrane-associated_guanylate_kinase,_WW_and_PDZ_domain-containing_protein_1	<i>Rattus norvegicus</i>	6E-44
TR65980 c0_g1	-2.42	2.11E-04	FAT3	Protocadherin_Fat_3	<i>Rattus norvegicus</i>	0
TR8375 c0_g1	-2.41	9.46E-05	P2RX5	P2X_purinoreceptor_5	<i>Rattus norvegicus</i>	2E-82
TR57480 c0_g1	-2.40	2.56E-05	SPTN5	Spectrin_beta_chain,_non-erythrocytic_5	<i>Homo sapiens</i>	4E-20
TR26681 c0_g5	-2.38	1.47E-01	SOS1	Son_of_sevenless_homolog_1	<i>Homo sapiens</i>	3E-158
TR28080 c0_g1	-2.38	2.15E-03	SVEP1	Sushi_von_Willebrand_factor_type_A,_EGF_and_pentraxin_domain-containing_protein_1	<i>Homo sapiens</i>	0
TR40831 c1_g1	-2.36	1.58E-01	CEGT	Ceramide_glucosyltransferase	<i>Mus musculus</i>	3E-69
TR59866 c0_g1	-2.36	4.71E-08	P5CS	Delta-1-pyrroline-5-carboxylate_synthase	<i>Pongo sp.</i>	1E-165

Supplemental Table 6. Overview of verified lumpfish genes in the complement cascade

			Top BLAST hit			
Gene-ID	Name	KEGG orthology ID	Description	E-value	Species	Accession number
Complement components						
TR16469 c0_g4	C1R	K01330	complement component C1q receptor	0	<i>Monopterus albus</i>	XP_020479096
TR40321 c0_g2	C1R/1S	K01331	Ca ²⁺ -dep.complex C1R/C1S subunit	5.14E-152	<i>Perca flavescens</i>	ABU63968
TR74292 c0_g1	C1S	K01331	complement component 1s	2.38E-162	<i>Oplegnathus fasciatus</i>	AIZ96980
TR74819 c0_g1	C2	K01332	complement component 2	1.20E-143	<i>Oplegnathus fasciatus</i>	AIN76765
TR104686 c0_g1	C3	K03990	PRED: complement C3-like	7E-54	<i>Lates calcarifer</i>	XP_018528015
TR63519 c0_g1	C3	K03990	PRED: complement C3-like	8E-57	<i>Lates calcarifer</i>	XP_018528013
TR56223 c0_g1	C3	K03990	complement C3-like	2.16E-165	<i>Acanthochromis polyacanthus</i>	XP_022077167
TR1140 c0_g2	C3	K03990	complement C3-like	2E-144	<i>Oryzias latipes</i>	XP_020555289
TR78116 c0_g1	C3	K03990	complement C3-like	0	<i>Stegastes partitus</i>	XP_008298001
TR77762 c0_g1	C3	K03990	PRED: complement component C3	6E-46	<i>Paralichthys olivaceus</i>	XP_019939804
TR88215 c0_g1	C3	K03990	complement component C3	8E-47	<i>Acanthochromis polyacanthus</i>	XP_022064368
TR64828 c1_g1	C4	K03989	complement component 4	0	<i>Oplegnathus fasciatus</i>	AIN76766
TR8953 c5_g4	C5	K03994	PRED: complement C5	0	<i>Lates calcarifer</i>	XP_018549132
TR49109 c2_g3	C5AR1	K04010	C3a anaphylatoxin chemotactic receptor	3.36E-156	<i>Larimichthys crocea</i>	XP_010733778
TR67474 c0_g1	C5AR1	K04010	C3a anaphylatoxin chemotactic receptor	1.77E-131	<i>Seriola dumerili</i>	XP_022617678
TR62911 c0_g1	C5AR1	K04010	C3a anaphylatoxin chemotactic receptor	0	<i>Seriola dumerili</i>	XP_022616734
TR48847 c0_g1	C5AR1	K04010	chemokine-like receptor 1	0	<i>Notothenia coriiceps</i>	XP_010785059
TR48847 c0_g2	C5AR1	K04010	chemokine-like receptor 1	0	<i>Notothenia coriiceps</i>	XP_010785059
TR70869 c1_g1	C5AR1	K04010	chemokine-like receptor 1	0	<i>Acanthochromis polyacanthus</i>	XP_022065692
TR70745 c3_g1	C5AR1	K04010	chemokine-like receptor 1	7.7E-170	<i>Acanthochromis polyacanthus</i>	XP_022065692
TR121251 c0_g1	C6	K03995	complement component C6 isoform X1	2.08E-30	<i>Labrus bergylta</i>	XP_020491684
TR43399 c0_g1	C6	K03995	complement component C6-like protein	3.4E-155	<i>Siniperca chuatsi</i>	AKA66307
TR15677 c0_g1	C6	K03995	complement component C6	2.18E-44	<i>Kryptolebias marmoratus</i>	XP_017259971
TR69965 c1_g1	C6	K03995	complement component C6	1.16E-94	<i>Notothenia coriiceps</i>	XP_010771775
TR77069 c0_g1	C7	K03996	complement component C7	1.8E-100	<i>Oplegnathus fasciatus</i>	AFZ93893
TR40591 c0_g1	C7	K03996	PRED: complement component C7 X2	2.13E-31	<i>Larimichthys crocea</i>	XP_019120079
TR62647 c0_g1	C7	K03996	complement component C7-2	3.59E-67	<i>Miichthys miiuy</i>	AKM12676
TR37853 c0_g1	C7	K03996	PRED: complement component C7-like	5.71E-124	<i>Notothenia coriiceps</i>	XP_010792065
TR94503 c0_g1	C7	K03996	PRED: complement component C7-like	5.71E-124	<i>Notothenia coriiceps</i>	XP_010792065
TR122188 c0_g1	C7	K03996	complement component C7-like	9.5E-57	<i>Monopterus albus</i>	XP_020479067
TR53398 c0_g1	C7	K03996	complement component C7-like	4.1E-153	<i>Monopterus albus</i>	XP_020479067

TR126963 c0_g1	C7	K03996	complement component C7-like	1.17E-41	<i>Seriola dumerili</i>	XP_022600176
TR30315 c0_g1	C8A	K03997	complement component 8 alpha	0	<i>Siniperca chuatsi</i>	AKA66305
TR27293 c0_g1	C8B	K03998	complement component 8 beta	0	<i>Oplegnathus fasciatus</i>	AFZ93889
TR122188 c0_g1	C9	K04000	complement component 9	3.3E-06	<i>Oryzias latipes</i> <i>Paralichthys olivaceus</i>	XP_004074560
TR53398 c0_g1	C9	K04000	complement component 9	1.01E-57		BAA86878
TR126963 c0_g1	C9	K04000	complement component 9	2.22E-60		BAA86878
TR78110 c0_g1	CFB	K01335	PRED: complement factor B-like	0	<i>Lates calcarifer</i>	XP_018547243
TR5891 c0_g1	CFD	K01334	PRED: complement factor D-like	1.54E-162	<i>Lates calcarifer</i>	XP_018541031
TR62909 c0_g1	CFH	K04004	PRED: complement factor H-like	0	<i>Hippocampus comes</i>	XP_019740144
TR74252 c0_g1	CFH	K04004	PRED: complement factor H-like	6e-122	<i>Lates calcarifer</i>	XP_01852636
TR84448 c0_g1	CFI	K01333	complement factor I	0	<i>Labrus bergylta</i>	XP_020514187
TR62930 c0_g1	CFP	K15412	properdin	0	<i>Acanthochromis polyacanthus</i>	XP_022073264
TR63188 c0_g1	CLU	K17252	clusterin	1.08E-168	<i>Monopterus albus</i>	XP_020459633
TR13054 c1_g2	CR3	K06461	Integrin alpha-M (ITAM)	0	<i>Dicentrarchus labrax</i>	CBN81367
TR64830 c2_g2	CR3	K06461	Integrin alpha-M (ITAM)	0	<i>Dicentrarchus labrax</i>	CBN81367
TR100 c0_g1	ITGB2	K06464	integrin beta-2-like	6.70	<i>Hippocampus comes</i>	XP_019731231
TR15999 c3_g3	ITGB2	K06464	integrin beta-2-like	6.70	<i>Hippocampus comes</i>	XP_019731231
TR13103 c3_g7	ITGB2	K06464	integrin beta-2-like	0	<i>Monopterus albus</i>	XP_020462274
TR1125 c1_g2	MASP1	K03992	Mannan-binding lectin serine protease 1	0	<i>Larimichthys crocea</i>	KKF27967

Supplemental Table 7. Overview of the species included in the phylogenetic analyses

Species	Authorities	Common name	Teleost order (Nelson et al 2016)	Teleost order (Betancur-R et al.2013)	Family ¹
<i>Danio rerio</i>	Hamilton, 1822	zebrafish	Cypriniformes	Cypriniformes	Cyprinidae
<i>Takifugu rubripes</i>	Temminck & Schlegel, 1850	fugu	Tetraodontiformes	Tetraodontiformes	Tetraodontidae
<i>Gasterosteus aculeatus</i>	Linnaeus, 1758	stickleback	Gasterosteiformes	Perciformes	Gasterosteidae
<i>Oryzias latipes</i>	Temminck & Schlegel, 1850	medaka	Beloniformes	Beloniformes	Adrianichthyidae
<i>Gadus morhua</i>	Linnaeus, 1758	cod	Gadiformes	Gadiformes	Gadidae
<i>Cyclopterus lumpus</i>	Linnaeus, 1758	lumpfish	Scorpaniformes	Perciformes	Cyclopteridae
<i>Miichthys miiuy</i>	Basilewsky, 1855	croaker	Perciformes	Percomorpharia *	Sciaenidae
<i>Ictalurus punctatus</i>	Rafinesque, 1818	catfish	Siluriformes	Siluriformes	Ictaluridae
<i>Plecoglossus altivelis altivelis</i>	Temminck & Schlegel, 1846	ayu sweetfish	Osmeriformes	Osmeriformes	Plecoglossidae
<i>Oncorhynchus mykiss</i>	Walbaum, 1792	rainbow trout	Salmoniformes	Salmoniformes	Salmonidae
<i>Coregonus maraena</i>	Bloch, 1779	mareana whitefish	Salmoniformes	Salmoniformes	Salmonidae
<i>Sparus aurata</i>	Linnaeus, 1758	gilthead seabream	Perciformes	Spariformes	Sparidae
<i>Scophthalmus maximus</i>	Linnaeus, 1758	turbot	Pleuronectiformes	Pleuronectiformes	Scophthalmidae
<i>Paralichthys olivaceus</i>	Temminck & Schlegel, 1846	bastard halibut	Pleuronectiformes	Pleuronectiformes	Paralichthyidae
<i>Oplegnathus fasciatus</i>	Temminck & Schlegel, 1844	barred knifejaw	Perciformes	Percomorpharia *	Oplegnathidae
<i>Cirrhinus mrigala</i>	Hamilton, 1822	mrigal carp	Cypriniformes	Cypriniformes	Cyprinidae
<i>Carassius auratus</i>	Linnaeus, 1758	goldfish	Cypriniformes	Cypriniformes	Cyprinidae
<i>Ctenopharyngodon idella</i>	Valenciennes, 1844	grass carp	Cypriniformes	Cypriniformes	Cyprinidae
<i>Cyprinus carpio</i>	Linnaeus, 1758	common carp	Cypriniformes	Cypriniformes	Cyprinidae
<i>Megalobrama amblycephala</i>	Yih, 1955	wuchang bream	Cypriniformes	Cypriniformes	Cyprinidae
<i>Salmo salar</i>	Linnaeus, 1758	Atlantic salmon	Salmoniformes	Salmoniformes	Salmonidae
<i>Larimichthys crocea</i>	Richardson, 1846	large yellow croaker	Perciformes	Percomorpharia *	Sciaenidae
<i>Epinephelus coioides</i>	Hamilton, 1822	orange-spotted grouper	Perciformes	Perciformes	Serranidae
<i>Carassius gibelio</i>	Bloch, 1782	prussian carp	Cypriniformes	Cypriniformes	Cyprinidae
<i>Gobiocypris rarus</i>	Ye & Fu, 1983	N.A.	Cypriniformes	Cypriniformes	Cyprinidae
<i>Squaliobarbus curriculus</i>	b	barbel chub	Cypriniformes	Cypriniformes	Cyprinidae
<i>Cynoglossus semilaevis</i>	Günther, 1873	tounge sole	Pleuronectiformes	Pleuronectiformes	Cynoglossidae
<i>Fundulus heteroclitus</i>	Linnaeus, 1766	mummichug	Cyprinodontiformes	Cyprinodontiformes	Fundulidae
<i>Trematomus bernacchii</i>	Boulenger, 1902	emerald rockcod	Perciformes	Perciformes	Nototheniidae
<i>Chionodraco hamatus</i>	Lönnberg, 1905	crocodile icefishes	Perciformes	Perciformes	Channichthyidae

<i>Nothobranchius furzeri</i>	Jubb, 1971	turquoise killifish	<i>Cyprinodontiformes</i>	Not examined	<i>Notobranchiidae</i>
<i>Nothobranchius rachovii</i>	Ahl, 1926	bluefin notho	<i>Cyprinodontiformes</i>	Not examined	<i>Notobranchiidae</i>
<i>Nothobranchius kadleci</i>	Richard, 2010	N.A.	<i>Cyprinodontiformes</i>	Not examined	<i>Notobranchiidae</i>
<i>Labeo rohita</i>	Hamilton, 1822	roho labeo	<i>Cypriniformes</i>	<i>Cypriniformes</i>	<i>Cyprinidae</i>
<i>Tachysurus fulvidraco</i>	Richardson, 1846	yellow catfish	<i>Siluriformes</i>	<i>Siluriformes</i>	<i>Bagridae</i>
<i>Tetraodon nigroviridis</i>	Marion de Procé, 1822	spotted green pufferfish	<i>Tetraodontiformes</i>	<i>Tetraodontiformes</i>	<i>Tetraodontidae</i>
<i>Gymnocypris przewalskii</i>	Kessler, 1876	N.A.	<i>Cypriniformes</i>	<i>Cypriniformes</i>	<i>Cyprinidae</i>
<i>Scleropages formosus</i>	Müller & Schlegel, 1840	asian arowana	<i>Osteoglossiformes</i>	<i>Osteoglossiformes</i>	<i>Osteoglossidae</i>
<i>Siniperca chuatsi</i>	Basilewsky, 1855	mandarin fish	<i>Perciformes</i>	<i>Percomorpharia *</i>	<i>Perchichthyidae</i>
<i>Seriola lalandi</i>	Valenciennes, 1833	yellowtail amberjack	<i>Perciformes</i>	<i>Carangiformes</i>	<i>Carangidae</i>
<i>Lates calcarifer</i>	Bloh, 1790	barramundi	<i>Perciformes</i>	<i>Carangimorpharia*</i>	<i>Latidae</i>
<i>Catla catla</i>	Hamilton, 1822	catla	<i>Cypriniformes</i>	<i>Cypriniformes</i>	<i>Cyprinidae</i>
<i>Clarias batrachus</i>	Linnaeus, 1758	philippine catfish	<i>Siluriformes</i>	<i>Siluriformes</i>	<i>Clariidae</i>
<i>Oreochromis niloticus</i>	Linnaeus, 1758	nile tilapia	<i>Perciformes</i>	<i>Cichliformes</i>	<i>Cichlidae</i>
<i>Labrus bergylta</i>	Ascanius, 1767	ballan wrasse	<i>Perciformes</i>	<i>Labriformes</i>	<i>Labridae</i>
<i>Stegastes partitus</i>	Poey, 1868	bicolor damselfish	<i>Perciformes</i>	<i>Ovalentariae*</i>	<i>Pomacentridae</i>
<i>Maylandia zebra</i>	Boulenger, 1899	zebra mbuna	<i>Perciformes</i>	<i>Cichliformes</i>	<i>Cichlidae</i>
<i>Callorhynchus milii</i>	Bory de Saint-Vincet, 1823	ghost shark	<i>Chimaeriformes</i>	Not examined	<i>Callorhichidae</i>

* incertae sedis. N.A. = non applicable

References:

Nelson, J. S., Grande, T. C. & Wilson, M. V. H. *Fishes of the world, 5th Edition*. (2016).
 Betancur, R. R. *et al.* The tree of life and a new classification of bony fishes. *PLoS Curr* doi:10.1371/current.tol.53ba26640df0ccaee75bb165c8c26288 (2013)

Supplemental Table 8: Accession numbers of genes in the phylogenetic analyses of TLRs

Abbreviation	Species	Acc. No.
TLR1		
TLR1	<i>Coregonus maraena</i>	CEF90214.1
TLR1	<i>Ctenopharyngodon idella</i>	ACT68332.1
TLR1	<i>Cyprinus carpio</i>	BAU98379.1
TLR1	<i>Epinephelus coioides</i>	AEB32452.1
TLR1	<i>Gymnocypris przewalskii</i>	ANQ46688.1
TLR1	<i>Homo sapiens</i>	Q15399
TLR1	<i>Larimichthys crocea</i>	AHB51065.1
TLR1	<i>Megalobrama amblycephala</i>	APT35500.1
TLR1	<i>Miichthys miiuy</i>	AKJ66261.1
TLR1	<i>Mus musculus</i>	Q9EPQ1
TLR1	<i>Oncorhynchus mykiss</i>	ACV92063.1
TLR1	<i>Oryzias latipes</i>	XP_011478513.1
TLR1	<i>Paralichthys olivaceus</i>	AFW04264.1
TLR1	<i>Tachysurus fulvidraco</i>	ANA09008.1
TLR1	<i>Takifugu rubripes</i>	AAW69368.1
TLR1	<i>Tetraodon nigroviridis</i>	ABO15772.1
TLR2		
TLR2	<i>Carassius gibelio</i>	AGR53440.1
TLR2	<i>Chionodraco hamatus</i>	ACT64127.1
TLR2	<i>Cirrhinus mrigala</i>	AHI59129.1
TLR2	<i>Ctenopharyngodon idella</i>	ACT68333.1
TLR2	<i>Cyprinus carpio</i>	BAU98381.1
TLR2	<i>Epinephelus coioides</i>	AEB32453.1
TLR2	<i>Homo sapiens</i>	O60603
TLR2	<i>Ictalurus punctatus</i>	ABD17347.1
TLR2	<i>Labeo rohita</i>	ADQ74644.1
TLR2	<i>Larimichthys crocea</i>	AJP16420.1
TLR2	<i>Latescalcarifer</i>	XP_18558738.1
TLR2	<i>Latimeria chalumnae</i>	ENSLACT00000017309

TLR2	<i>Lepisosteus oculatus</i>	ENSLOCT00000021874
TLR2	<i>Megalobrama amblycephala</i>	APT35501.1
TLR2	<i>Mus musculus</i>	Q9QUN7
TLR2	<i>Oncorhynchus mykiss</i>	CCK73195.1
TLR2	<i>Oplegnathus fasciatus</i>	AFZ81806.1
TLR2	<i>Oryzias latipes</i>	XP_004079640.2
TLR2	<i>Paralichthys olivaceus</i>	BAD01044.1
TLR2	<i>Scophthalmus maximus</i>	AMQ35498.1
TLR2	<i>Tachysurus fulvidraco</i>	ANA09009.1
TLR2	<i>Trematomus bernacchii</i>	ACT64128.1
TLR2.1	<i>Larimichthys crocea</i>	KKF28982.1
TLR2.2	<i>Larimichthys crocea</i>	KKF15865.1
TLR2a	<i>Cyprinus carpio</i>	ACP20793.2
TLR2b-P	<i>Oncorhynchus mykiss</i>	NP_001117891.1
TLR2c	<i>Cyprinus carpio</i>	BAU98380.1
TLR2-P	<i>Danio rerio</i>	NP_00997977.1
TLR3		
TLR3	<i>Carassius auratus</i>	ABC86865.1
TLR3	<i>Carassius gibelio</i>	AGR53439.1
TLR3	<i>Ctenopharyngodon idella</i>	ABI64155.1
TLR3	<i>Cyprinus carpio</i>	ABL11473.1
TLR3	<i>Danio rerio</i>	AAT37633.1
TLR3	<i>Gadus morhua</i>	ENSGMOP00000000792
TLR3	<i>Epinephelus coioides</i>	ADZ76423.1
TLR3	<i>Gobiocypris rarus</i>	ABL11471.1
TLR3	<i>Homo sapiens</i>	O15455
TLR3	<i>Ictalurus punctatus</i>	AEI59664.1
TLR3	<i>Megalobrama amblycephala</i>	ABI83673.1
TLR3	<i>Miichthys miiuy</i>	ALJ55565.1
TLR3	<i>Mus musculus</i>	Q99MB1
TLR3	<i>Oncorhynchus mykiss</i>	AAX68425.1
TLR3	<i>Oryzias latipes</i>	XP_011475331.1
TLR3	<i>Paralichthys olivaceus</i>	BAM11216.1

TLR3	<i>Salmo salar</i>	AKE14222.1
TLR3	<i>Scophthalmus maximus</i>	AHW76803.1
TLR3	<i>Squaliobarbus curriculus</i>	ALO75529.1
TLR3	<i>Takifugu rubripes</i>	AAW69373.1
TLR3.2	<i>Cyprinus carpio</i>	AHE74142.1
TLR3.2	<i>Larimichthys crocea</i>	ADR01099.1
TLR3b	<i>Paralichthys olivaceus</i>	BAD01047.1
TLR3-P	<i>Ictalurus punctatus</i>	NP_001186997.1
TLR3-P	<i>Larimichthys crocea</i>	NP_001290242.1
TLR4		
TLR4	<i>Cyprinus carpio</i>	BAU98382.1
TLR4	<i>Homo sapiens</i>	O00206
TLR4	<i>Labeo rohita</i>	AOM81178.1
TLR4	<i>Mus musculus</i>	Q9QUK6
TLR4.1	<i>Ctenopharyngodon idella</i>	AEQ64877.1
TLR4.2	<i>Ctenopharyngodon idella</i>	AEQ64878.1
TLR4.3	<i>Ctenopharyngodon idella</i>	AEQ64879.1
TLR4.4	<i>Ctenopharyngodon idella</i>	AEQ64880.1
TLR4a	<i>Megalobrama amblycephala</i>	ALB39038.1
TLR4a-P	<i>Danio rerio</i>	NP_001315534.1
TLR4b	<i>Danio rerio</i>	AAH68358.1
TLR4b	<i>Megalobrama amblycephala</i>	AMH41158.1
TLR4ba	<i>Cyprinus carpio</i>	AHH85806.1
TLR4ba	<i>Danio rerio</i>	NP_001124523.1
TLR4bb	<i>Cyprinus carpio</i>	AHH85807.1
TLR4bb-P	<i>Danio rerio</i>	NP_00997978.2
TLR4c	<i>Megalobrama amblycephala</i>	AKP20514.1
TLR5		
TLR5	<i>Homo sapiens</i>	O60602
TLR5	<i>Mus musculus</i>	Q9JLF7
TLR5a	<i>Ctenopharyngodon idella</i>	AIO11757.1
TLR5b	<i>Ctenopharyngodon idella</i>	AIO11758.1
TLR5M	<i>Carassius auratus</i>	AQX43081.1

TLR5M	<i>Cirrhinus mrigala</i>	AHI59128.1
TLR5M	<i>Coregonus maraena</i>	CEF90216.1
TLR5M	<i>Cyprinus carpio</i>	BAU98383.1
TLR5M	<i>Ictalurus punctatus</i>	AEI59669.1
TLR5M	<i>Megalobrama amblycephala</i>	APT35502.1
TLR5M	<i>Miichthys miiuy</i>	ALJ55566.1
TLR5M	<i>Oplegnathus fasciatus</i>	AQT26515.1
TLR5M	<i>Oryzias latipes</i>	XP_011490072.1
TLR5M	<i>Paralichthys olivaceus</i>	BAJ16367.1
TLR5M	<i>Plecoglossus altivelis altivelis</i>	BAI68384.1
TLR5M	<i>Scophthalmus maximus</i>	AMQ35502.1
TLR5M-P	<i>Danio rerio</i>	NP_001124067.2
TLR5M-P	<i>Oncorhynchus mykiss</i>	NP_001118216.1
TLR5S	<i>Miichthys miiuy</i>	ALJ55567.1
TLR5S	<i>Paralichthys olivaceus</i>	AEN71826.1
TLR5S	<i>Scophthalmus maximus</i>	ANS71058.1
TLR5S	<i>Sparus aurata</i>	CCP37739.1
TLR5S	<i>Takifugu rubripes</i>	AAW69378.1
TLR5S-P	<i>Ictalurus punctatus</i>	NP_001187158.1
TLR5S-P	<i>Oncorhynchus mykiss</i>	NP_001117680.1
TLR6		
TLR6	<i>Homo sapiens</i>	Q9Y2C9
TLR6	<i>Mus musculus</i>	Q9EPW9
TLR7		
TLR7	<i>Cyprinus carpio</i>	BAJ19518.1
TLR7	<i>Homo sapiens</i>	Q9NYK1
TLR7	<i>Gadus morhua</i>	ENSGMOP00000001755
TLR7	<i>Ictalurus punctatus</i>	AEI59670.1
TLR7	<i>Larimichthys crocea</i>	AGO28200.1
TLR7	<i>Megalobrama amblycephala</i>	APT35503.1
TLR7	<i>Miichthys miiuy</i>	ALJ55568.1
TLR7	<i>Mus musculus</i>	P58681
TLR7	<i>Oncorhynchus mykiss</i>	ACV41797.1

TLR7	<i>Oryzias latipes</i>	XP_0011488154.1
TLR7	<i>Salmo salar</i>	CCX35457.1
TLR7	<i>Scophthalmus maximus</i>	AMQ35499.1
TLR8		
TLR8	<i>Homo sapiens</i>	Q9NR97
TLR8	<i>Mus musculus</i>	P58682
TLR8	<i>Takifugu rubripes</i>	AAW69376.1
TLR8	<i>Gadus morhua</i>	ENSGMOP00000001741
TLR8	<i>Larimichthys crocea</i>	AGO28201.1
TLR8	<i>Miichthys miiuy</i>	ALJ55569.1
TLR8	<i>Paralichthys olivaceus</i>	AOS00680.1
TLR8	<i>Scophthalmus maximus</i>	AQU15238.1
TLR8a	<i>Cyprinus carpio</i>	BAU98387.1
TLR8a	<i>Megalobrama amblycephala</i>	APT35504.1
TLR8a1	<i>Oncorhynchus mykiss</i>	ACV41799.1
TLR8a1	<i>Coregonus maraena</i>	CEF90219.1
TLR8a1	<i>Salmo salar</i>	NP_001155165.1
TLR8a2	<i>Oncorhynchus mykiss</i>	ACV41798.1
TLR8a2	<i>Salmo salar</i>	CCX35458.1
TLR8b	<i>Cyprinus carpio</i>	BAU98386.1
TLR8b2	<i>Coregonus maraena</i>	CEF90223.1
TLR9		
TLR9	<i>Chionodraco hamatus</i>	ACT64129.1
TLR9	<i>Coregonus maraena</i>	CEF90220.1
TLR9	<i>Ctenopharyngodon idella</i>	ADB96920.1
TLR9	<i>Cynoglossus semilaevis</i>	ACL68661.1
TLR9	<i>Fundulus heteroclitus</i>	JAR68992.1
TLR9	<i>Gadus morhua</i>	ENSGMOP000000012030
TLR9	<i>Homo sapiens</i>	Q9EQU3
TLR9	<i>Megalobrama amblycephala</i>	APT35506.1
TLR9	<i>Miichthys miiuy</i>	ALJ55570.1
TLR9	<i>Mus musculus</i>	Q9NR96
TLR9	<i>Paralichthys olivaceus</i>	BAE80691.1

TLR9	<i>Scophthalmus maximus</i>	AQU15239.1
TLR9	<i>Takifugu rubripes</i>	AAW69377.1
TLR9	<i>Trematomus bernacchii</i>	ACT64130.1
TLR9a	<i>Epinephelus coioides</i>	ACV04893.1
TLR9a	<i>Larimichthys crocea</i>	ACF60624.1
TLR9a	<i>Sparus aurata</i>	AAW81697.1
TLR9b	<i>Epinephelus coioides</i>	ACV04894.1
TLR9b	<i>Sparus aurata</i>	AAW81696.1
TLR9P	<i>Larimichthys crocea</i>	ACF60625.1
TLR9-P	<i>Danio rerio</i>	NP_001124066.1
TLR9-P	<i>Oncorhynchus mykiss</i>	NP_001123463.1
TLR9-P	<i>Salmo salar</i>	NP_001117125.1
TLR10, TLR11, TLR12		
TLR10	<i>Homo sapiens</i>	Q9BXR5
TLR11	<i>Mus musculus</i>	Q6R5P0
TLR12	<i>Mus musculus</i>	Q6QNU9
TLR13		
TLR13	<i>Mus musculus</i>	Q6R5N8
TLR13	<i>Miichthys miiuy</i>	ALJ55571.1
TLR13	<i>Salmo salar</i>	NP_001133860.1
TLR13	<i>Oryzias latipes</i>	XP_004078275.1
TLR13a	<i>Larimichthys crocea</i>	KKF19122.1
TLR13b	<i>Larimichthys crocea</i>	KKF22613.1
TLR13-P	<i>Larimichthys crocea</i>	NP_001290325.1
TLR13-P	<i>Oreochromis niloticus</i>	NP_001298246.1
TLR14		
TLR14	<i>Miichthys miiuy</i>	ALJ55572.1
TLR14	<i>Paralichthys olivaceus</i>	BAJ78226.1
TLR18		
TLR18	<i>Danio rerio</i>	AAI63840.1
TLR18	<i>Ictalurus punctatus</i>	AEI59674.1
TLR18	<i>Ctenopharyngodon idella</i>	AIB55030.1
TLR18	<i>Megalobrama amblycephala</i>	APT35507.1

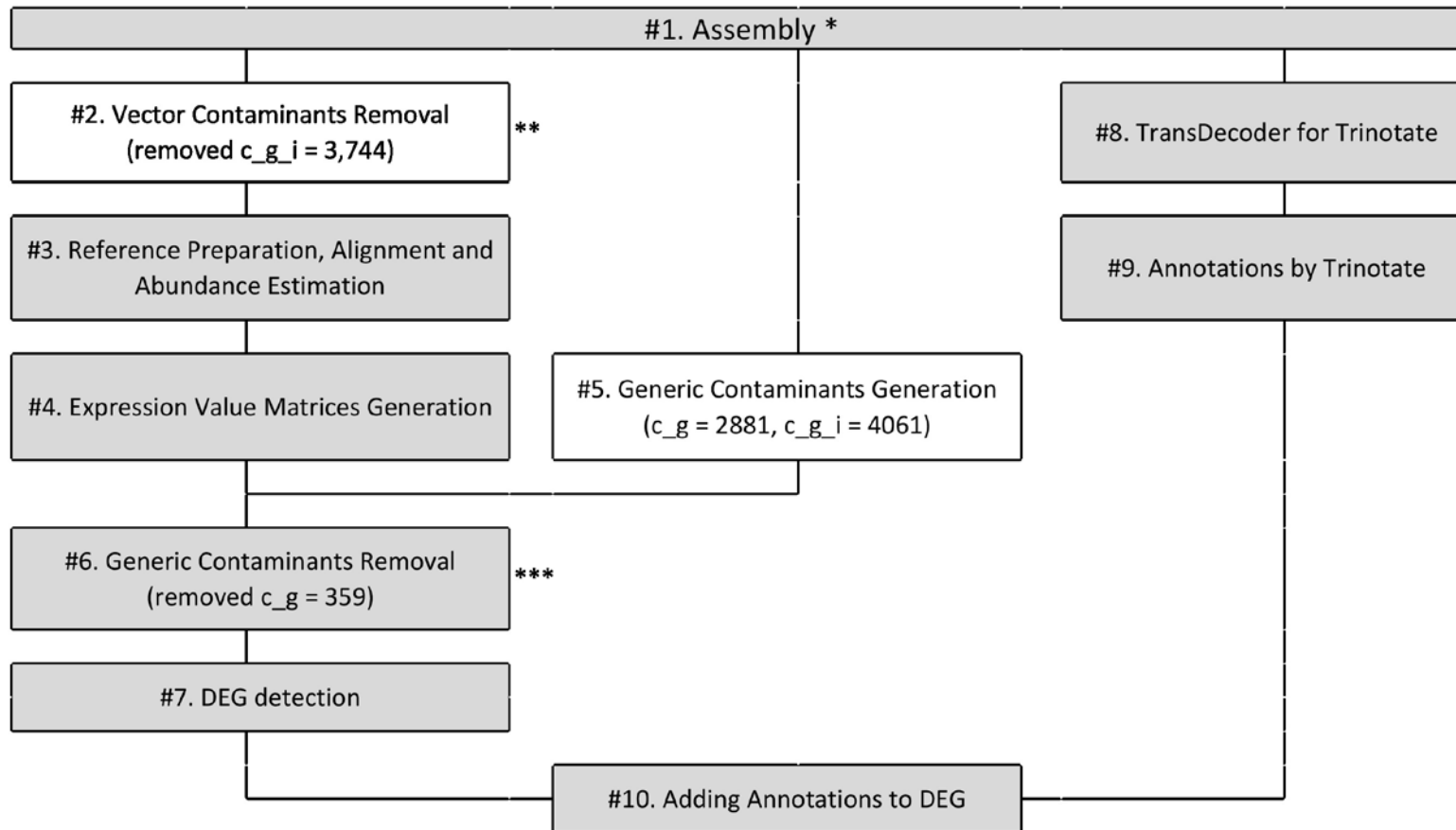
TLR18	<i>Gadus morhua</i>	ENSGMOP00000004019
TLR18	<i>Salmo salar</i>	CDK60413.1
TLR18a	<i>Cyprinus carpio</i>	BAU98389.1
TLR18b	<i>Cyprinus carpio</i>	BAU98388.1
TLR19		
TLR19	<i>Megalobrama amblycephala</i>	APT35508.1
TLR19	<i>Cyprinus carpio</i>	BAU98390.1
TLR20		
TLR20	<i>Cyprinus carpio</i>	AHH85805.1
TLR20	<i>Ctenopharyngodon idella</i>	AHN49762.1
TLR20	<i>Megalobrama amblycephala</i>	APT35509.1
TLR20.2	<i>Danio rerio</i>	NP_001170914.2
TLR20a	<i>Ictalurus punctatus</i>	NP_001187159.1
TLR21		
TLR21	<i>Epinephelus coioides</i>	ADM34974.2
TLR21	<i>Gadus morhua</i>	AFK76484.1
TLR21	<i>Paralichthys olivaceus</i>	AFW04263.1
TLR21	<i>Ctenopharyngodon idella</i>	AGM21642.1
TLR21	<i>Clarias batrachus</i>	AGM39445.1
TLR21	<i>Oplegnathus fasciatus</i>	AIT52504.1
TLR21	<i>Miichthys miiuy</i>	ALJ55573.1
TLR21	<i>Scophthalmus maximus</i>	AMQ35500.1
TLR21	<i>Larimichthys crocea</i>	AOZ21302.1
TLR21	<i>Cyprinus carpio</i>	BAU98391.1
TLR21	<i>Danio rerio</i>	CAQ13807.1
TLR21	<i>Ictalurus punctatus</i>	NP_001186994.1
TLR21/13	<i>Seriola lalandi</i>	ALI16363.1
TLR21-P	<i>Takifugu rubripes</i>	NP_001027751.1
TLR22	<i>Carassius auratus</i>	AQX43082.1
TLR22	<i>Catla catla</i>	AGW43269.2
TLR22	<i>Ctenopharyngodon idella</i>	ADX97523.2
TLR22	<i>Epinephelus coioides</i>	AGA84053.1
TLR22	<i>Labeo rohita</i>	AHV90682.1

TLR22	<i>Lates calcarifer</i>	AOV82293.1
TLR22	<i>Miichthys miiuy</i>	ALJ55574.1
TLR22	<i>Scophthalmus maximus</i>	AIC75881.1
TLR22	<i>Seriola lalandi</i>	AKN10669.1
TLR22	<i>Siniperca chuatsi</i>	AFC95889.1
TLR22	<i>Sparus aurata</i>	CDK37745.1
TLR22	<i>Takifugu rubripes</i>	AAW69372.1
TLR22		
TLR22a	<i>Cyprinus carpio</i>	BAU98393.1
TLR22a	<i>Danio rerio</i>	NP_001122147.2
TLR22a	<i>Salmo salar</i>	CAJ80696.1
TLR22a2	<i>Salmo salar</i>	CAR62394.1
TLR22a-P	<i>Oncorhynchus mykiss</i>	NP_001117884.1
TLR22b	<i>Cyprinus carpio</i>	BAU98395.1
TLR22b	<i>Danio rerio</i>	AAI63527.1
TLR22b	<i>Gadus morhua</i>	AFK76486.1
TLR22c	<i>Cyprinus carpio</i>	BAU98394.1
TLR22d	<i>Gadus morhua</i>	AFK76488.1
TLR22g	<i>Gadus morhua</i>	AFK76491.1
TLR22i	<i>Gadus morhua</i>	AFK76493.1
TLR22l	<i>Gadus morhua</i>	AFK76496.1
TLR22-P	<i>Scleropages formosus</i>	KPP60030.1
TLR23	<i>Miichthys miiuy</i>	ALJ55575.1
TLR23a	<i>Gadus morhua</i>	AFK76497.1
TLR23b	<i>Gadus morhua</i>	AFK76498.1
TLR23-TLR26		
TLR	<i>Nothobranchius furzeri</i>	SBP54052.1
TLR	<i>Nothobranchius kadlecii</i>	SBP83130.1
TLR	<i>Nothobranchius rachovii</i>	SBR74107.1
TLR25	<i>Ictalurus punctatus</i>	AEI59680.1
TLR25a	<i>Cyprinus carpio</i>	BAU98397.1
TLR25b	<i>Cyprinus carpio</i>	BAU98396.1
TLR1_25	<i>Oryzias latipes</i>	XP_004083162.1

TLR26	<i>Ictalurus punctatus</i>	AEI59681.1
TLR27		
TLR27	<i>Callorhinchus milii</i>	XP_007893881.1
TLR27	<i>Latimeria chalumnae</i>	ENSLACP000000017183
TLR27	<i>Lepisosteus oculatus</i>	ENSLOCP000000021836
TLR28		
TLR2.2	<i>Epinephelus coioides</i>	AIS23533.1
TLR2.2	<i>Maylandia zebra</i>	XP_14266344.1
TLR2.2	<i>Oreochromis niloticus</i>	XP_019215654.1
TLR2.2	<i>Oryzias latipes</i>	XP_004078522.1
TLR2.2	<i>Stegastes partitus</i>	XP_008295222.1
TLR2.2	<i>Labrus bergylta</i>	XP_020495322.1
TLR2.2	<i>Lates calcarifer</i>	XP_018558738.
TLR28	<i>Miichthys miiuy</i>	AKN63433.1

Supplementary Methods for Assembly and Annotation:

We have used Trinity v2.0.6 (<https://github.com/trinityrnaseq/trinityrnaseq/wiki>) installed on our Linux server (CentOS release 6.9). Transcripts were assembled using all the raw reads (n=13) with additional trinity parameters, defined for library type (RF), trimmomatic quality trimming and reads normalization, resulted assembly fasta file. Assembly was subjected for removal of vector contamination using blastn. Thereafter, reference preparation, abundance estimation, generation of expression value matrices, generic contaminants removal from matrices and detection of differentially expressed genes were performed. To add functional annotations, as recommended by Trinity, we used Trinotate v2.0.2 <https://trinotate.github.io/> and TransDecoder v2.0.1 (<https://github.com/TransDecoder/TransDecoder/wiki>). The steps performed during assembly and annotations are available in Supplementary Figure 1 given below. All the commands used for Trinity, contaminants removal, Trinotate and TransDecoder are available in Supplementary Methods for Bioinformatic scripts.



*, Raw reads .fastq.gz were trimmed during assembly with additional trinity parameters defined for trimmomatic quality trimming and resulted trimmed .fastq.gz.P.qtrim.gz files. These files were submitted to array express and available as .fastq.gz files

**, Assembly after vector contaminants removal submitted to Array Express and available as Cyclopterus_lumpus_assembly_post_trinity_Cleaned.fasta

***, Expression value matrices after removing generic contaminants submitted to Array Express as differential_expression_table.txt

c_g_i, contig_gene_isoform generated by trinity

Supplementary Figure 1. Steps performed during de novo assembly and annotations.

Supplementary Methods for Bioinformatic scripts

1. Assembly

#####

```
$ /trinityrnaseq-2.0.6/Trinity --seqType fq --left all_left --right all_right --output trinity_out_dir --SS_lib_type RF --trimmomatic --normalize_reads --max_memory 200G --CPU 60
```

2. Vector contaminants removal

#####

1) Vector genomes seq in fasta were downloaded from NCBI:

gi|9630635|ref|NC_001915.1| Infectious pancreatic necrosis virus segment A, complete sequence

gi|9630638|ref|NC_001916.1| Infectious pancreatic necrosis virus segment B, complete sequence

gi|336122587|ref|NC_015633.1| Vibrio anguillarum 775 chromosome I, complete sequence

gi|336125405|ref|NC_015637.1| Vibrio anguillarum 775 chromosome II, complete sequence

2) All_Vector.fsa file was generated using txt editor combining vector genomes

```
$ /ncbi-blast-2.2.31+/bin/makeblastdb -in Vector_Genome_fa/All_Vector.fsa -parse_seqids -dbtype nucl -out nt
```

```
$ /ncbi-blast-2.2.31+/bin/blastn -task megablast -db nt -perc_identity 90 -max_target_seqs 1 -outfmt "7 qacc qstart qend sacc sstart send sstrand" -query Trinity.fasta -out result_.out -num_threads 4
```

3) good_ids.txt was generated by filtering vector ids from trinity.fasta ids using linux commands and assembly was cleaned using following commands:

```
$ perl -ne 'if(/^>(\S+)/){$c=$i{$1}}$c?print:chomp;$i{$_}=1 if @ARGV' good_ids.txt Trinity.fasta > Trinity_Cleaned.fasta
```

Note: After removing vector contamination, assembly was renamed as Trinity_Cleaned.fasta

Trinity Assembly_Stat

```
$ /trinityrnaseq-2.0.6/util/TrinityStats.pl Trinity_Cleaned.fasta
```

3. Reference Preparation, Alignment and Abundance Estimation (n=13)

#####

Ref Preparation:

```
$ /trinityrnaseq-2.0.6/util/align_and_estimate_abundance.pl --transcripts Trinity_Cleaned.fasta --est_method RSEM --aln_method bowtie --trinity_mode --prep_reference --SS_lib_type RF --output_prefix --debug --thread_count 20
```

Abundance Estimation: individual run for each sample using bowtie and RSEM (below is example for first sample only)

```
$ /trinityrnaseq-2.0.6/util/align_and_estimate_abundance.pl --transcripts Trinity_Cleaned.fasta --seqType fq --left 1-Haugland_ATCACG_L008_R1_001.fastq.gz.PwU.qtrim.fq --right 1-Haugland_ATCACG_L008_R2_001.fastq.gz.PwU.qtrim.fq --est_method RSEM --aln_method bowtie --trinity_mode --SS_lib_type RF --output_dir Sample_1_out --output_prefix Sample_1_ --debug --thread_count 5
```

Detailed Assessment of Read Content of the Assembly

```
$ cd Sample_1_out
```

```
$ /trinityrnaseq-2.0.6/util/bowtie_PE_separate_then_join.pl --seqType fq --left 1-Haugland_ATCACG_L008_R1_001.fastq.gz.PwU.qtrim.fq --right 1-Haugland_ATCACG_L008_R2_001.fastq.gz.PwU.qtrim.fq --target Trinity_Cleaned.fasta --aligner bowtie --SS_lib_type RF --retain_intermediate_files -- -p 4 --all --best --strata -m 300 > bowtie_PE_separate_then_join.out
```

```
$ /trinityrnaseq-2.0.6/util/SAM_nameSorted_to_uniq_count_stats.pl bowtie_out/bowtie_out.nameSorted.bam > SAM_nameSorted_to_uniq_count_stats.out
```

4. Expression Value Matrices Generation

#####

```
$ /trinityrnaseq-2.0.6/util/abundance_estimates_to_matrix.pl --est_method RSEM --cross_sample_fpk_norm TMM --out_prefix Trinity_genes Sample_1_genes.results Sample_2_genes.results Sample_3_genes.results Sample_4_genes.results Sample_5_genes.results Sample_6_genes.results Sample_7_genes.results Sample_8_genes.results Sample_9_genes.results Sample_10_genes.results Sample_11_genes.results Sample_12_genes.results Sample_13_genes.results > genes_abundance_estimates_to_matrix.pl.out
```

5. Generic Contaminants Generation

#####

generating and removing generic contaminants contigs by blasting against NT for clearly non-eukaryote sequences

1) the contigs of the assembly was split into 20 blast jobs, example commandline

```
$ blastn -query lumpfish_unclean00.fa -db nt -num_threads 4 -outfmt '6 qseqid sseqid evalue staxids sskingdoms' 2>
lumpfish_unclean00.fa_blast_results.log | gzip -c > lumpfish_unclean00.fa_blast_results.txt.gz
```

2) the blast results were parsed to compute an Alien Index [**Gladyshev et al, 2008**], a corrected log ratio between the E-value for the best Eukaryote and best non-Eukaryote blast hits of each contig. We used an Alien Index ≥ 45 , corresponding to an E-value ≥ 20 orders of magnitude difference between the best non-metazoan hit to the best metazoan hit, to identify 4061 non-Eukaryote contigs most likely being contaminant. These were removed from the transcript set and not used in subsequent analysis.

```
$ for FN in `ls lumpfish_*blast_results.txt.gz`; do alienIndex -c 4 -e 2 -m Eukaryota -i <( zcat $FN ) > ai_output.${FN}.txt; done
```

```
$ cat ai_output* | perl -lane 'print if $F[3] > 45' | cut -f1 > trld_isBacterial.txt
```

6. Generic Contaminants Removal

```
#####
```

Note: generic contaminants were removed from expression matrices before DEG detection. This was done as part of the R-script provided below using geneld_isContaminant.txt. Ids in trld_isBacterial.txt file were in c_g_i (contig_gene_isoform) format and only isoform ids (_i) were removed to generate geneld_isContaminant.txt file. geneld_isContaminant.txt file contained only gene Ids (c_g).

7. DEG detection (p-value 0.1)

```
#####
```

The differential expression analysis was performed in R using the limma-voom method in the limma package [Ritchie et al 2015] based on the RSEM read count estimates per gene (Trinity_genes.counts.matrix). see R-script provided.

8. TransDecoder for Trinotate

```
#####
```

```
$ TransDecoder.LongOrfs -t /export/kjempetujafs/service/projects/2015-
06_Haugland_transcriptome/WD_Charitra/TransDecoder_out_dir/Trinity_Cleaned_perl.fasta
```

```
$ ncbi-blast-2.2.31+/bin/blastp -query Trinity_Cleaned_perl.fasta.transdecoder_dir/longest_orfs.pep -db uniprot_sprot.trinotate.pep -
max_target_seqs 1 -outfmt 6 -evalue 1e-5 -num_threads 30 > blastp.outfmt6
```

```
$ hmmer-3.1b2-linux-intel-x86_64/binaries/hmmscan --cpu 30 --domtblout pfam.domtblout Pfam-A.hmm longest_orfs.pep & tail -f nohup.out
```

```
$ TransDecoder-2.0.1/TransDecoder.Predict -t Trinity_Cleaned.fasta --retain_pfam_hits pfam.domtblout --retain_blastp_hits blastp.outfmt6
```

9. Annotations by Trinotate

#####

Trinotate: <https://trinotate.github.io>

Capturing BLAST Homologies

```
$ ncbi-blast-2.2.31+/blastx -query Trinity_Cleaned_perl.fasta -db uniprot_uniref90.trinotate.pep -num_threads 8 -max_target_seqs 1 -outfmt 6 > uniref90.blastx.outfmt6
```

```
$ ncbi-blast-2.2.31+/blastp -query Trinity_Cleaned_perl.fasta.transdecoder.pep -db uniprot_uniref90.trinotate.pep -num_threads 8 -max_target_seqs 1 -outfmt 6 > uniref90.blastp.outfmt6
```

Running HMMER

```
$ hmmer-3.1b2-linux-intel-x86_64/binaries/hmmscan --cpu 8 --domtblout TrinotatePFAM.out Pfam-A.hmm Trinity_Cleaned.fasta.transdecoder.pep > pfam.log
```

Running signalP to predict signal peptides

```
$ signalp-4.1/signalp -f short -n signalp.out Trinity_Cleaned.fasta.transdecoder.pep
```

Running tmHMM to predict transmembrane regions

```
$ tmhmm-2.0c/bin/tmhmm --short < Trinity_Cleaned.fasta.transdecoder.pep > tmhmm.out
```

Running RNAMMER

```
$ perl /Trinotate-2.0.2/util/rnammer_support/RnammerTranscriptome.pl --transcriptome Trinity_Cleaned.fasta --path_to_rnammer rnammer-1.2.src/rnammer
```

Loading Above Results into a Trinotate SQLite Database

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite init --gene_trans_map Trinity.fasta.gene_trans_map --transcript_fasta Trinity_Cleaned.fasta --transdecoder_pep Trinity_Cleaned.fasta.transdecoder.pep
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite LOAD_swissprot_blastp blastp.outfmt6
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite LOAD_swissprot_blastx blastx.outfmt6
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite LOAD_pfam TrinotatePFAM.out
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite LOAD_tmhmm tmhmm.out
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite LOAD_signalp signalp.out
```

```
$ Trinotate /Trinotate-2.0.2/Trinotate.sqlite report > trinotate_annotation_report.xls
```

10. Adding Annotations to DEG

#####

Annotations in .xls file were added to DEG table using simple text/excel editor.

```
### R-script
```

```
#####
```

```
``{r config}
```

```
library(limma)
```

```
library(edgeR)
```

```
library(dplyr)
```

```
library(ggplot2)
```

```
#library(DESeq2)
```

```
#setwd("")
```

```
``
```

```
Load RSEM read count estimates per gene
```

```
``{r loadData, echo="FALSE"}
```

```
x = read.delim("../Trinity_genes.counts.matrix")
```

```
names(x) = c('gene_id',
```

```
            'ctrl_6h_r1','ctrl_6h_r2','ctrl_6h_r3',
```

```
            'treat_6h_r1','treat_6h_r2','treat_6h_r3',
```



```

'ctrl_24h_r1','ctrl_24h_r2','ctrl_24h_r3',
'treat_24h_r1','treat_24h_r2','treat_24h_r3',
'ipnv_24h_r1')

```

```

row.names(x)=x$gene_id

```

```

...

```

2016-04 Remove contaminants (detected by BLAST analysis)

```

```{r rmContaminants}

```

```

isContaminant=read.table("..geneId_isContaminant.txt")

```

```

isContaminant = as.character(isContaminant$V1)

```

```

x = x[!x$gene_id %in% isContaminant,]

```

```

...

```

```

```{r, echo=FALSE}

```

```

x=x[,-1] #strip gene_id column

```

```

...

```

Remove genes that are not appreciably expressed (min 10 reads)

```
```{r dropLowAbundant}
```

```
keep=rowSums(x>10) >= 3 #at least 10 reads in at least 3 samples
```

```
table(keep) #keep 34280 genes/221k
```

```
x = x[keep,]
```

```
...
```

Set up experimental design matrix

```
```{r designMatrix}
```

```
treatment <- factor(c('ctrl_6h','ctrl_6h','ctrl_6h','treat_6h','treat_6h','treat_6h',
```

```
                  'ctrl_24h','ctrl_24h','ctrl_24h','treat_24h','treat_24h','treat_24h',
```

```
                  'invp'), levels=c('ctrl_6h','treat_6h','ctrl_24h','treat_24h','invp')) #ctrl_mo experimental design
```

```
sample_pool <- factor(c(1,2,3,1,2,3,1,2,3,1,2,3,3), levels=c(1,2,3))
```

```
design <- model.matrix(~0 + treatment + sample_pool)
```

```
data.frame(Sample=colnames(x),treatment,sample_pool)
```

```
rownames(design) <- colnames(x)
```

```
...
```

```
### limma-voom analysis
```

```
``{r limma_scale_counts}  
  
y <- DGEList(counts=x,group=treatment)  
  
y <- calcNormFactors(y)  
  
v <- voom(y,design,plot=TRUE)  
  
plotMDS(v,top=50,labels=treatment,  
        col=ifelse(treatment=="treat_6h","blue","red"),gene.selection="common")  
  
fit <- lmFit(v,design)  
  
fit <- eBayes(fit)  
  
``
```

Examine the different contrasts (treatment vs control at each time-point)

```
``{r limmaDEResults}  
  
cont.matrix <- makeContrasts(TvsC_6h=treatmenttreat_6h-treatmentctrl_6h,  
                             TvsC_24h=treatmenttreat_24h-treatmentctrl_24h,  
                             INPVvsC=treatmentinpv-treatmentctrl_24h,  
                             # TvsT=treatmenttreat_24h-treatmenttreat_6h,
```

```

#           CvsC=treatmentctrl_24h-Intercept,
           levels=design)

#cont.matrix

#row.names(cont.matrix) = c("(Intercept)","sample_pool2","sample_pool3",
#           "treatmenttreat_6h","treatmentctrl_24h","treatmenttreat_24h",
#           "treatmentinpv") #fix for (Intercept) vs Intercept

fit2 <- contrasts.fit(fit, contrast=cont.matrix)

fit2 <- eBayes(fit2)

#fit2 <- contrasts.fit(fit, contrast=c(-1,0,0,1,0,0,0)) #treat vs ctrl 6h

#fit2 <- eBayes(fit2)

top_TvsC_6h = topTable(fit2, coef="TvsC_6h", number = Inf, adjust="BH")

sum(top_TvsC_6h$adj.P.Val<0.05)

sum(top_TvsC_6h$adj.P.Val<0.0001)

#head(top_TvsC_6h)

#x['TR31519|c0_g1',]

#x['TR35037|c0_g2',]

top_TvsC_24h = topTable(fit2, coef="TvsC_24h", number = Inf, adjust="BH")

```

```
sum(top_TvsC_24h$adj.P.Val<0.05)
```

```
sum(top_TvsC_24h$adj.P.Val<0.0001)
```

```
#fit2 <- contrasts.fit(fit, contrast=c(0,0,0,0,-1,0,1)) #INPV vs ctrl 24h
```

```
#fit2 <- eBayes(fit2)
```

```
top_VcsC = topTable(fit2, coef="INPVvsC", number= Inf, adjust="BH")
```

```
#fit2 <- contrasts.fit(fit, contrast=c(0,0,0,-1,0,1,0)) #treat 24h vs treat_6h
```

```
#fit2 <- eBayes(fit2)
```

```
#topTable(fit2, coef="TvsT", adjust="BH")
```

```
#fit2 <- contrasts.fit(fit, contrast=c(-1,0,0,0,1,0,0)) #ctrl 24h vs ctrl_6h
```

```
#fit2 <- eBayes(fit2)
```

```
#topTable(fit2, coef="CvsC", adjust="BH")
```

```
###
```

```
...
```

Merge 6h and 24h results for QC and plots

merge the DE contrast tables for export

```

```{r mergeTimepoints}

top_6_24 = merge(top_TvsC_6h, top_TvsC_24h,by="row.names",all.x=TRUE)

names(top_6_24) = c('geneID',
 'logFC_6h',
 'avgExpr_6h',
 't_6h',
 'pval_6h',
 'adjpval_6h',
 'b_6h',
 'logFC_24h',
 'avgExpr_24h',
 't_24h',
 'pval_24h',
 'adjpval_24h',
 'b_24h')

```

```

QC: Examine the concordance of the two timepoints

```

```{r rankPlot}

```

```

rank_6h = top_TvsC_6h %>% mutate(rank = rank(adj.P.Val)) %>% select(rank)

rank_24h = top_TvsC_24h %>% mutate(rank = rank(adj.P.Val)) %>% select(rank)

rank_6_24 = merge(rank_6h, rank_24h, by="row.names",all.x=TRUE)

cor(rank_6_24$rank.x,rank_6_24$rank.y, method="spearman") #spearman corr 0.90

ggplot(rank_6_24, aes(x=rank.x,y=rank.y)) + geom_point(aes(alpha=0.1)) +

labs(title="Concordance of gene DE P-value rank at 6 and 24 hrs",

 x="P-value rank 6 hrs",

 y="P-value rank 24 hrs") +

theme(legend.position = "none")

#top 100

ggplot(rank_6_24, aes(x=rank.x,y=rank.y)) + geom_point(aes(alpha=0.1)) +

labs(title="Concordance of gene DE P-value rank at 6 and 24 hrs",

 x="P-value rank 6 hrs",

 y="P-value rank 24 hrs") +

theme(legend.position = "none") +

xlim(0, 100) +

ylim(0,100)

```

...

Conclusion: The genes that are differentially regulated at 6 hrs are also diff regulated at 24 hrs

Now, diagnostic exploration of of p-values - se DESeq manual for background

```
``{r pvaldist}

plot_pval_diag = function(padj){
 orderInPlot = order(padj)
 showInPlot = (padj[orderInPlot] <= 0.01)
 alpha = 0.01
 plot(seq(along=which(showInPlot)), padj[orderInPlot][showInPlot], pch=".", xlab = expression(rank(p[i])), ylab=expression(p[i]))
 abline(a=0, b=alpha/length(padj), col="red3", lwd=2)
}

plot_pval_diag(top_6_24$adjpval_6h)
plot_pval_diag(top_6_24$adjpval_24h)

plot_pval_cutoff_comp = function(padj){
 padj.log <- -log10(padj)
```



```

orderInPlot = order(padj.log)

plot(padj.log[orderInPlot], type="l")

sigline <- c(.05, .01, 5, 1,05, 01)

sigline <- -log10(sigline)

sigcolors <- c("red", "blue", "green", "yellow", "pink", "purple")

sapply(1:length(sigline), function(x){abline(h=sigline[padj.log], col=sigcolors[padj.log])})

}

plot_pval_cutoff_comp(top_6_24$adjpval_6h)

plot_pval_cutoff_comp(top_6_24$adjpval_24h)

plot_pval_histograms = function(pval){

 hist(top_6_24$adjpval_6h, breaks=5000, ylim = c(0,100))

 hist(top_6_24$adjpval_6h, breaks=2000, xlim = c(0,0.1))

 hist(top_6_24$adjpval_6h, breaks=1000000, xlim = c(0,0.01))

 hist(top_6_24$adjpval_6h, breaks=1000000, xlim = c(0,0.02))

 hist(top_6_24$adjpval_6h, breaks=1000000, xlim = c(0,0.005))

 hist(top_6_24$adjpval_6h, breaks=1000000, xlim = c(0,0.0005))

}

ggplot(top_6_24, aes(adjpval_2h)) + geom_histogram(bins=5000) + ylim(0,100)

```

```
ggplot(top_6_24, aes(adjpval_6h)) + geom_histogram(bins=2000) + xlim(0,0.1)
ggplot(top_6_24, aes(adjpval_6h)) + geom_histogram(bins=2000) + xlim(0,0.01)
ggplot(top_6_24, aes(adjpval_6h)) + geom_histogram(bins=2000) + xlim(0,0.02)
ggplot(top_6_24, aes(adjpval_6h)) + geom_histogram(bins=2000) + xlim(0,0.002)
ggplot(top_6_24, aes(adjpval_6h)) + geom_histogram(bins=2000) + xlim(0,0.0002)
```

```
plot_pval_histograms(top_6_24$adjpval_6h)
plot_pval_histograms(top_6_24$adjpval_24h)
```

```
...
```

Volcano plot

```
```{r volcanoPlot, eval=FALSE, echo=FALSE}
res <- top_6_24 %>% dplyr::select(geneID, logFC_6h, pval_6h, adjpval_6h)
names(res) = c("geneID", "log2FoldChange", "pvalue", "padj")

# Make a basic volcano plot
with(res, plot(log2FoldChange, -log10(pvalue), pch=20, main="Volcano plot", xlim=c(-2.5,2)))
```

```

# Add colored points: red if padj<0.05, orange if log2FC>1, green if both)

with(subset(res, padj<0.05), points(log2FoldChange, -log10(pvalue), pch=20, col="red"))

with(subset(res, abs(log2FoldChange)>1), points(log2FoldChange, -log10(pvalue), pch=20, col="orange"))

with(subset(res, padj<0.05 & abs(log2FoldChange)>1), points(log2FoldChange, -log10(pvalue), pch=20, col="green"))

# Label points with the textxy function from the calibrate plot

library(calibrate)

with(subset(res, padj<0.05 & abs(log2FoldChange)>1), textxy(log2FoldChange, -log10(pvalue), labs=geneID, cex=.8))

```

#volcano plot

#scattering the M values (log2 ratio) on the x axis against the p value (-log10 the p value).

```
ggplot(top_6_24, aes(x=logFC_6h, y = -log10(adjpval_6h))) + geom_point(aes(alpha=0.1))
```

```
detach("package:calibrate", unload=TRUE) #calibrate select conflicts with dplyr
```

```
```
```

Export DE Tables

```
```{r exportDETables}
```

```
top_6_24_out = top_6_24 %>% select(geneID,logFC_6h,avgExpr_6h,adjpval_6h,logFC_24h,avgExpr_24h,adjpval_24h) %>% arrange(adjpval_24h)
```

```
write.table(top_6_24_out, file="differential_expression_table.txt", quote=FALSE, sep = "\t")
```

```
#down 6h
```

```
downreg_6 = top_6_24 %>% filter(adjpval_6h < 1e-5 & logFC_6h < 0) %>%select(geneID,logFC_6h,avgExpr_6h,adjpval_6h) %>%  
arrange(adjpval_6h)
```

```
write.table(downreg_6, file="differential_expression_down_6h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE)
```

```
upreg_6 = top_6_24 %>% filter(adjpval_6h < 1e-5 & logFC_6h > 0) %>%select(geneID,logFC_6h,avgExpr_6h,adjpval_6h) %>% arrange(adjpval_6h)
```

```
write.table(upreg_6, file="differential_expression_up_6h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE)
```

```
downreg_24 = top_6_24 %>% filter(adjpval_24h < 1e-5 & logFC_24h < 0) %>%select(geneID,logFC_24h,avgExpr_24h,adjpval_24h) %>%  
arrange(adjpval_24h)
```

```
write.table(downreg_24, file="differential_expression_down_24h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE)
```

```
upreg_24 = top_6_24 %>% filter(adjpval_24h < 1e-5 & logFC_24h > 0) %>%select(geneID,logFC_24h,avgExpr_24h,adjpval_24h) %>%  
arrange(adjpval_24h)
```

```
write.table(upreg_24, file="differential_expression_up_24h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE)
```

```
# write labels file for Trinity GO analysis (label\tgeneid)
```

```
# because row.names (genelds) can not be repeated, we do 6h and 24h separately
```

```
labels_6h = rbind(  
data.frame('label' = rep("down_6",nrow(downreg_6)), 'geneID' = downreg_6$geneID),  
data.frame('label' = rep("up_6",nrow(upreg_6)), 'geneID' = upreg_6$geneID))  
write.table(labels_6h, file="differential_expression_labels_for_go_6h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE, col.names = FALSE)
```

```
labels_24h = rbind(  
data.frame('label' = rep("down_24",nrow(downreg_24)), 'geneID' = downreg_24$geneID),  
data.frame('label' = rep("up_24",nrow(upreg_24)), 'geneID' = upreg_24$geneID)  
)  
write.table(labels_24h, file="differential_expression_labels_for_go_24h.txt",  
            quote=FALSE, sep = "\t", row.names = FALSE, col.names = FALSE)
```

```
...
```

Running the Go overrep code through trinity

```
```{r gooverrep, engine="bash", eval=FALSE}
```

```
/trinity/Analysis/DifferentialExpression/run_GOseq.pl --genes_single_factor=differential_expression_labels_6h.txt --
GO_assignments=go_annotations.txt --lengths=gene_length.txt > go_overrep_6h.txt
```

```
/trinity/Analysis/DifferentialExpression/run_GOseq.pl --genes_single_factor=differential_expression_labels_24h.txt --
GO_assignments=go_annotations.txt --lengths=gene_length.txt > go_overrep_24h.txt
```

...

DESeq2 plots for QA

```
```{r DESeq2Plots}  
countData = round(x)  
colData = data.frame(sample_pool,treatment)  
row.names(colData) = colnames(countData)  
library(DESeq2)  
dds <- DESeqDataSetFromMatrix(countData = countData,  
                               colData = colData,  
                               design = ~ sample_pool + treatment)  
dds <- DESeq(dds)  
  
rld <- rlog(dds)
```

```

vsd <- varianceStabilizingTransformation(dds)

rlogMat <- assay(rld)

vstMat <- assay(vsd)

...

```{r gene_heatmap_1-30}

library("RColorBrewer")

library("gplots")

select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]

hmccl <- colorRampPalette(brewer.pal(9, "GnBu"))(100)

heatmap.2(assay(vsd)[select,], col = hmccl,

 Rowv = FALSE, Colv = FALSE, scale="none",

 dendrogram="none", trace="none", margin=c(10, 6))

...

```{r gene_heatmap_1-100}

library("RColorBrewer")

library("gplots")

```

```

select <- order(rowMeans(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:100]

hmccl <- colorRampPalette(brewer.pal(9, "GnBu"))(100)

heatmap.2(assay(vsd)[select,], col = hmccl,
          Rowv = FALSE, Colv = FALSE, scale="none",
          dendrogram="none", trace="none", margin=c(10, 6))

...

```{r gene_heatmap_genefilter_rowVars_1-30}

library("RColorBrewer")

library("gplots")

select <- order(genefilter::rowVars(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:30]

hmccl <- colorRampPalette(brewer.pal(9, "GnBu"))(100)

heatmap.2(assay(vsd)[select,], col = hmccl,
 Rowv = FALSE, Colv = FALSE, scale="none",
 dendrogram="none", trace="none", margin=c(10, 6))

...

```



```

```{r gene_heatmap_genefilter_rowVars_1-100}

library("RColorBrewer")

library("gplots")

select <- order(genefilter::rowVars(counts(dds,normalized=TRUE)),decreasing=TRUE)[1:100]

hmcol <- colorRampPalette(brewer.pal(9, "GnBu"))(100)

heatmap.2(assay(vsd)[select,], col = hmcol,

          Rowv = FALSE, Colv = FALSE, scale="none",

          dendrogram="none", trace="none", margin=c(10, 6))

...

```{r sample_heatmap}

library(ggplot2)

distsRL <- dist(t(assay(rld)))

mat <- as.matrix(distsRL)

rownames(mat) <- colnames(mat) <- with(colData(dds),

 paste(sample_pool, treatment, sep=" : "))

hc <- hclust(distsRL)

```

```

heatmap.2(mat, Rowv=as.dendrogram(hc),
 symm=TRUE, trace="none",
 col = rev(hmcol), margin=c(13, 13))

#plotPCA(rld, intgroup=c("treatment", "sample_pool"))

data <- plotPCA(rld, intgroup=c("treatment", "sample_pool"), returnData=TRUE)
percentVar <- round(100 * attr(data, "percentVar"))

ggplot(data, aes(PC1, PC2, color=sample_pool, shape=treatment)) +
 geom_point(size=3) +
 xlab(paste0("PC1: ",percentVar[1],"% variance")) +
 ylab(paste0("PC2: ",percentVar[2],"% variance"))

```

# Supplementary Results of Trinity RSEM

**Table 1.** Read count matrix from RSEM output

								Genes expressed		
Sample		proper_pairs	total aligned	improper_pairs	left_only	right_only	≥ 2 FPKM	≥ 1 FPKM	Total	
Control-1	1	count	24,189,368	29,980,237	3,976,458	1,137,983	676,428	39,506	66,391	221,659
		pct	80.68	--	13.26	3.8	2.26	--	--	--
	2	count	27,178,664	32,937,998	4,168,232	956,390	634,712	36,077	65,219	221,659
		pct	82.51	--	12.65	2.9	1.93	--	--	--
	3	count	28,659,252	37,114,067	6,626,260	1,097,874	730,681	36,250	66,468	221,659
		pct	77.22	--	17.85	2.96	1.97	--	--	--
Treated-1	4	count	24,899,860	31,065,905	4,472,704	1,028,504	664,837	36,935	65,105	221,659
		pct	80.15	--	14.4	3.31	2.14	--	--	--
	5	count	21,999,490	27,276,150	3,793,870	886,000	596,790	37,203	63,440	221,659
		pct	80.65	--	13.91	3.25	2.19	--	--	--
	6	count	24,627,776	30,329,773	4,052,902	985,884	663,211	37,970	66,678	221,659
		pct	81.2	--	13.36	3.25	2.19	--	--	--
Control-2	7	count	26,683,856	33,281,629	4,862,120	1,062,160	673,493	36,291	62,730	221,659
		pct	80.18	--	14.61	3.19	2.02	--	--	--
	8	count	23,200,556	29,769,028	4,832,936	1,025,958	709,578	38,483	64,273	221,659
		pct	77.94	--	16.23	3.45	2.38	--	--	--
	9	count	26,758,676	32,816,055	4,297,730	1,060,466	699,183	36,797	63,158	221,659
		pct	81.54	--	13.1	3.23	2.13	--	--	--
Treated-2	10	count	25,713,468	31,888,148	4,358,192	1,117,602	698,886	35,393	60,377	221,659
		pct	80.64	--	13.67	3.5	2.19	--	--	--
	11	count	23,517,108	29,484,868	4,255,492	1,038,233	674,035	37,826	64,233	221,659
		pct	79.76	--	14.43	3.52	2.29	--	--	--
	12	count	22,513,278	27,487,774	3,394,564	949,623	630,309	39,037	67,158	221,659
		pct	81.9	--	12.35	3.45	2.29	--	--	--
Treated-3	13	count	20,204,506	26,674,188	4,519,114	1,212,334	738,234	39,580	65,727	221,659
		pct	75.75	--	16.94	4.54	2.77	--	--	--

Note: Vector contaminated seq were removed before RSEM output.

#####

## Counts of transcripts, etc.

#####

Total trinity 'genes': 221659

Total trinity transcripts: 346430

Percent GC: 46.65

#####

Stats based on ALL transcript contigs:

#####

Contig N10: 6782

Contig N20: 4988

Contig N30: 3947

Contig N40: 3144

Contig N50: 2502

Median contig length: 585

Average contig: 1250.14

Total assembled bases: 433087424

#####

## Stats based on ONLY LONGEST ISOFORM per 'GENE':

#####

Contig N10: 5514

Contig N20: 3647

Contig N30: 2446

Contig N40: 1608

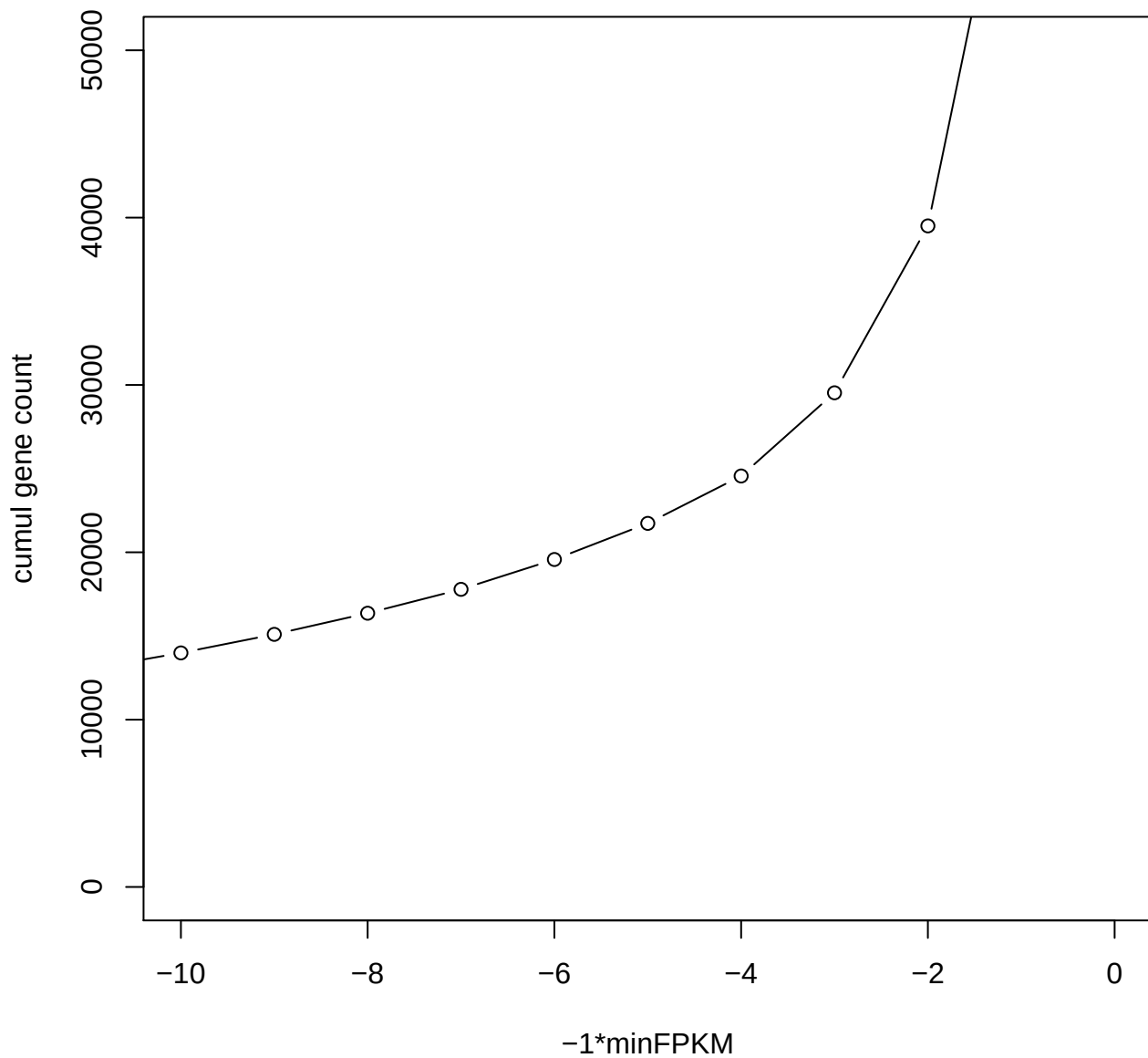
Contig N50: 1071

Median contig length: 401

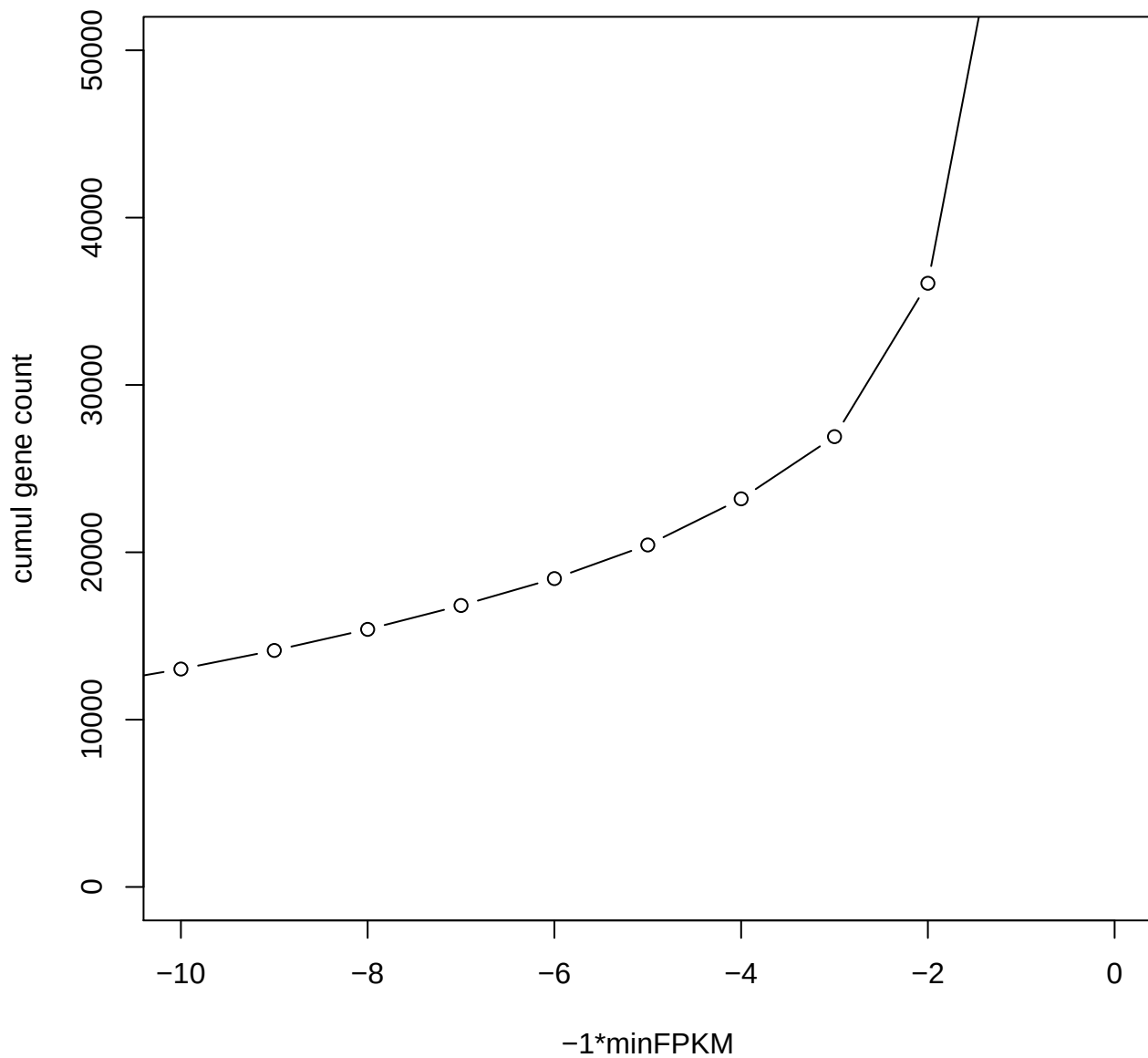
Average contig: 735.04

Total assembled bases: 162928446

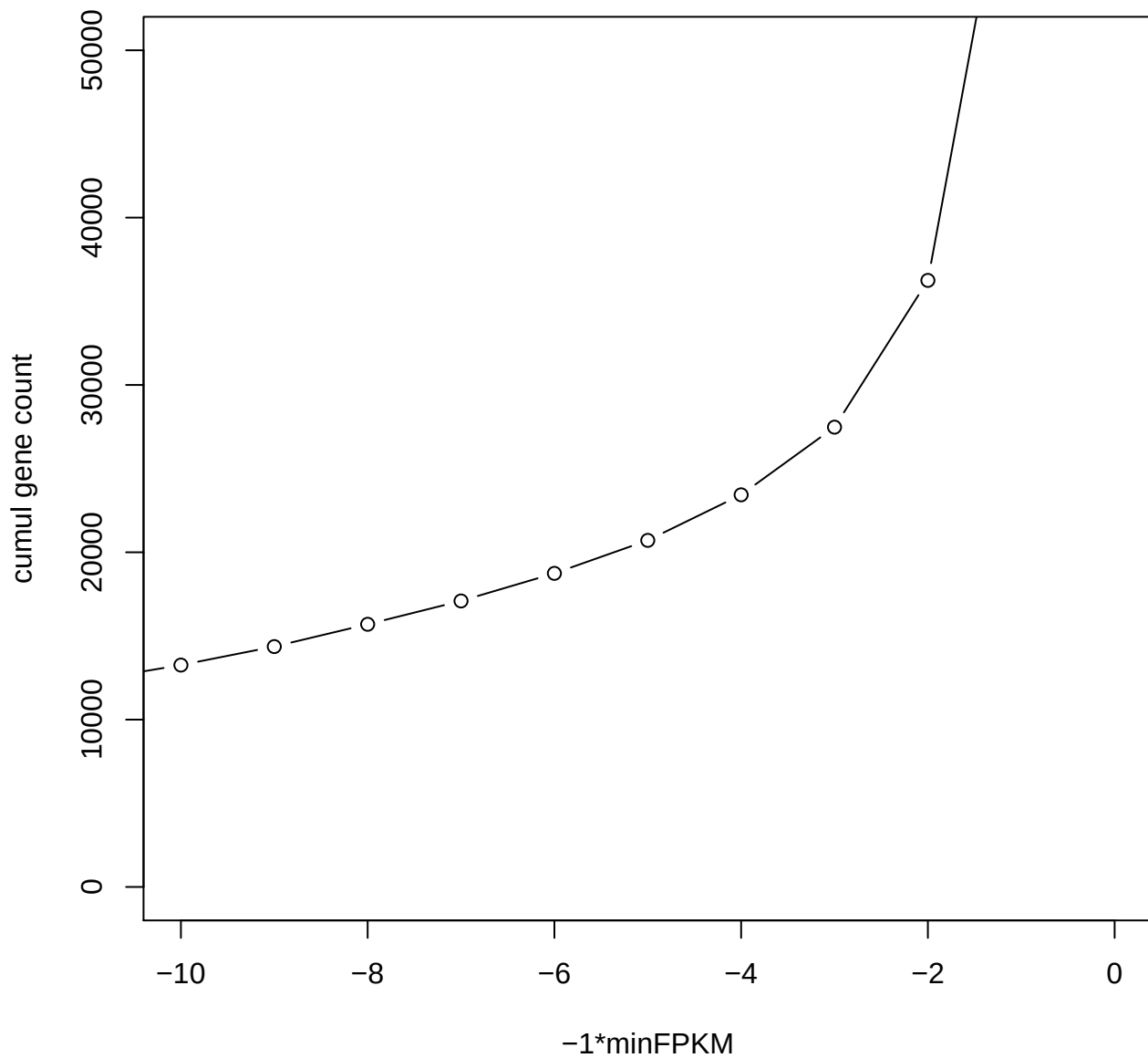
**Sample\_1 (gene count vs. minFPKM)**



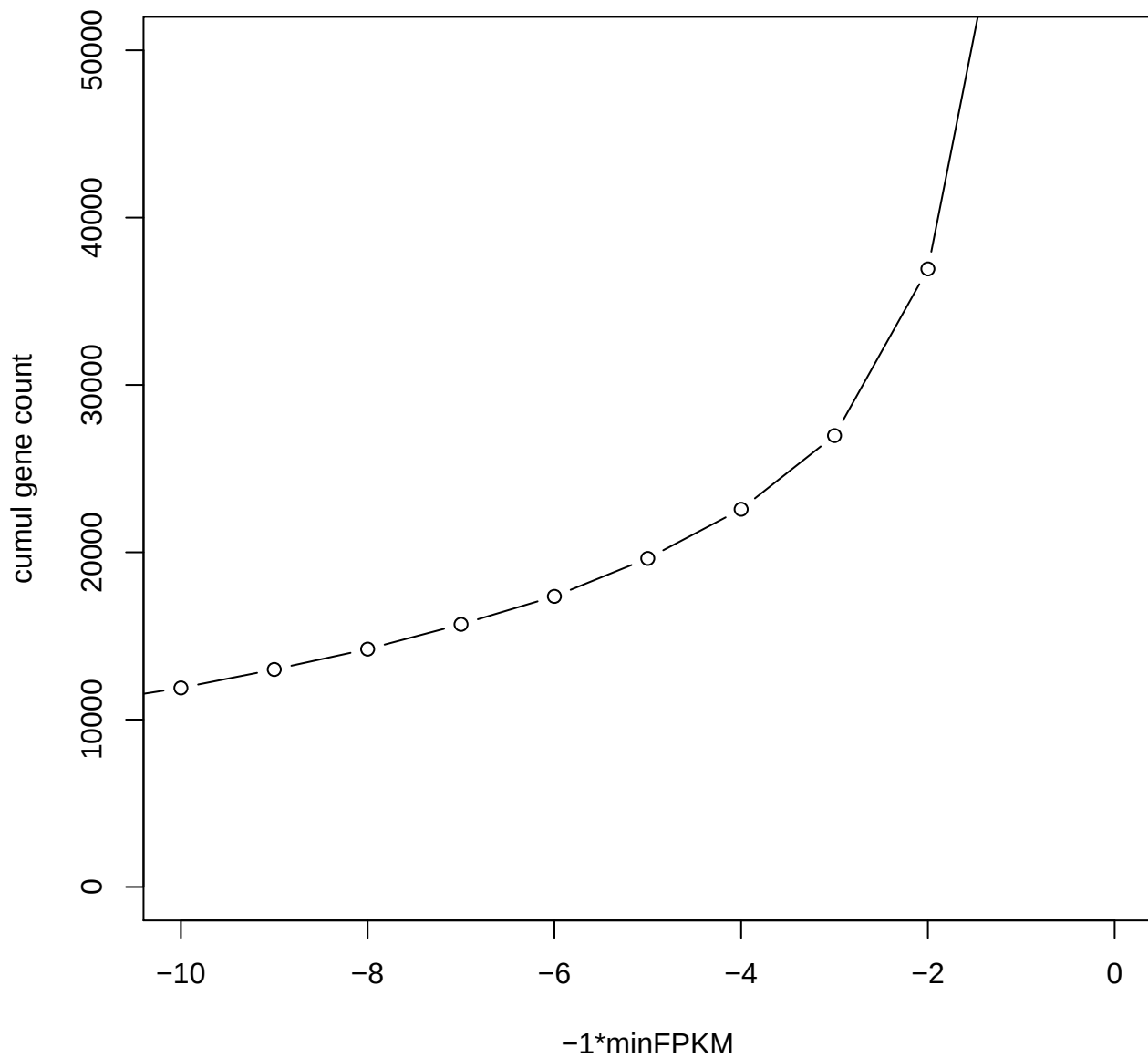
**Sample\_2 (gene count vs. minFPKM)**



**Sample\_3 (gene count vs. minFPKM)**

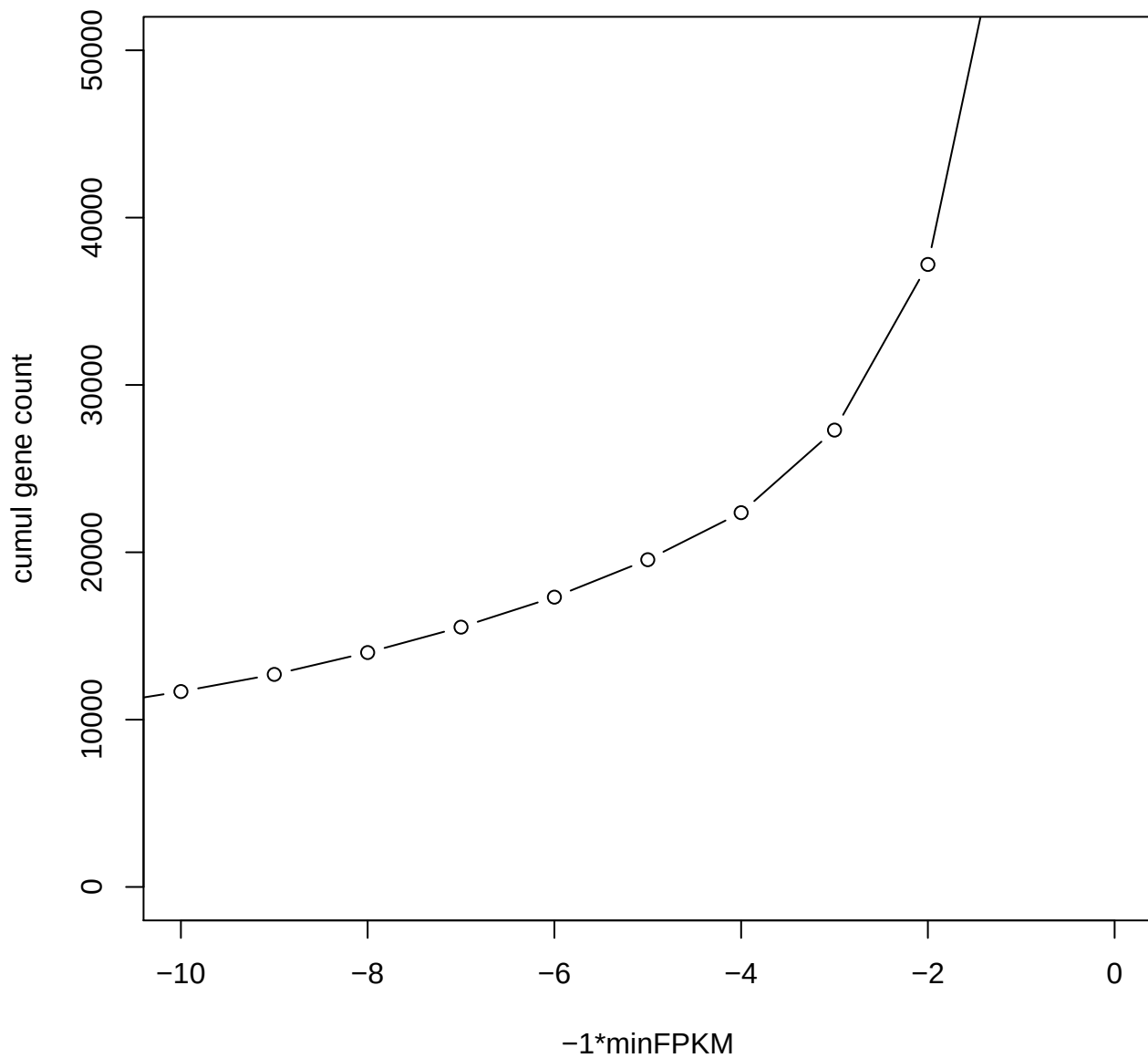


**Sample\_4 (gene count vs. minFPKM)**

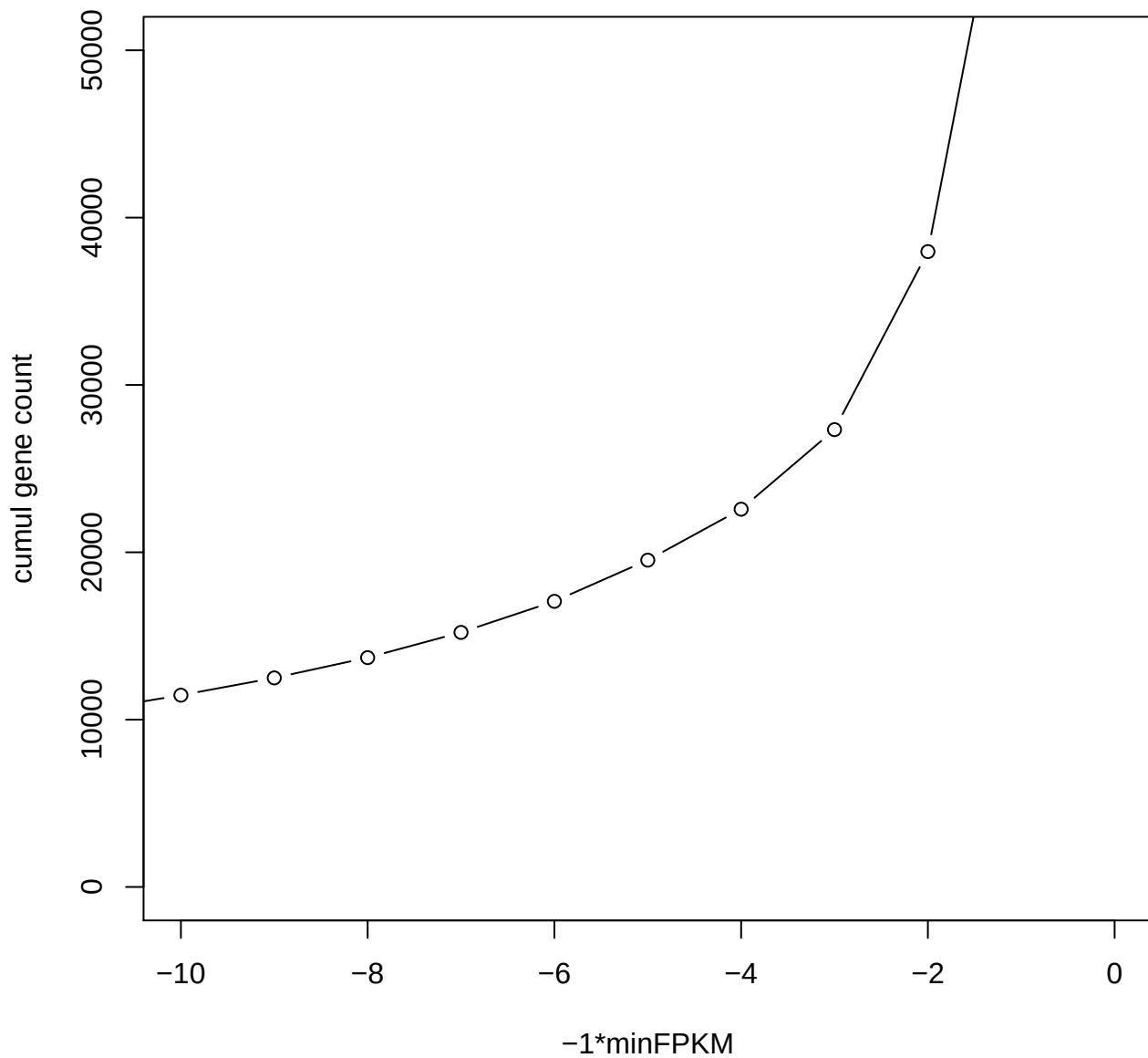




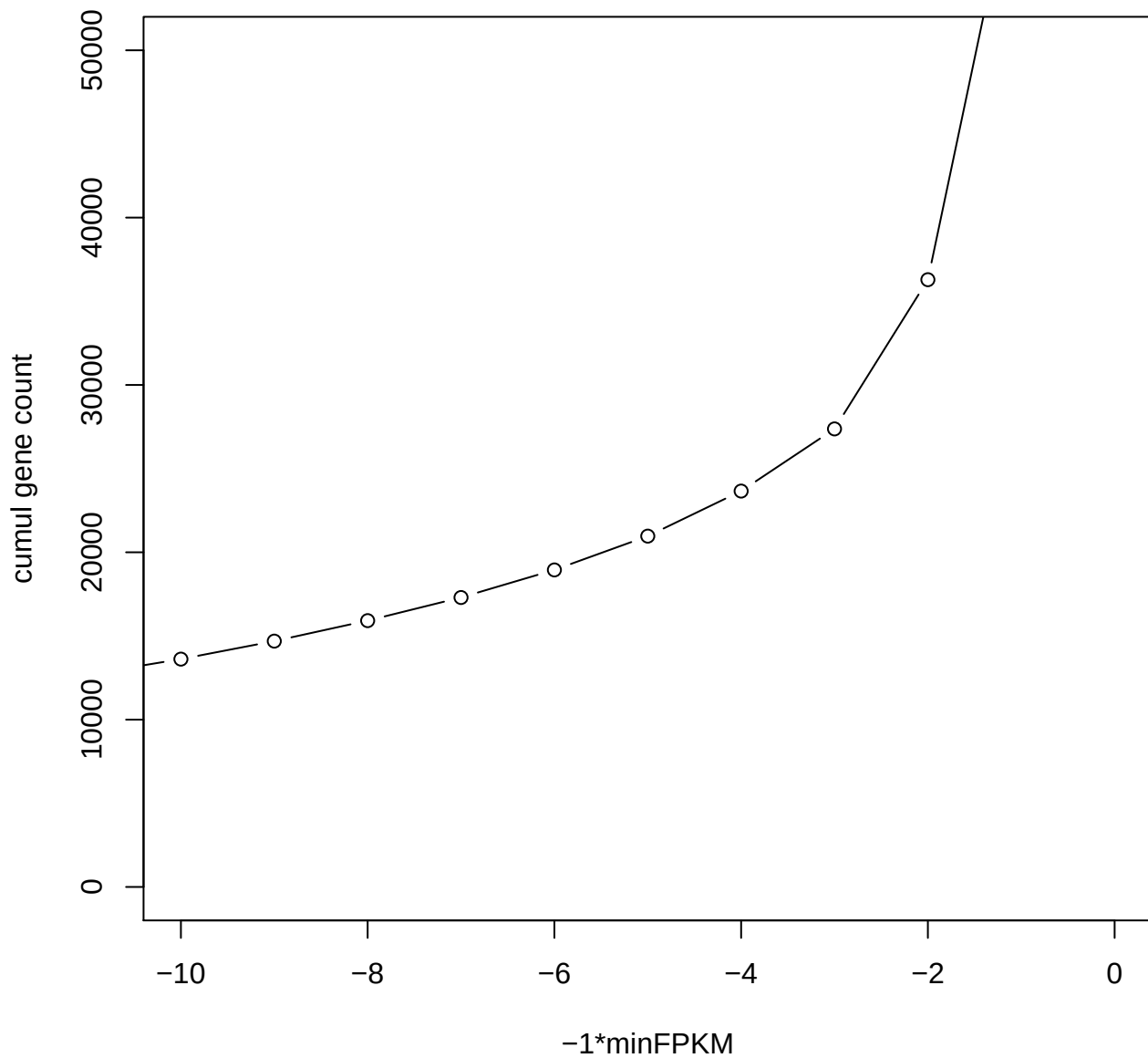
**Sample\_5 (gene count vs. minFPKM)**



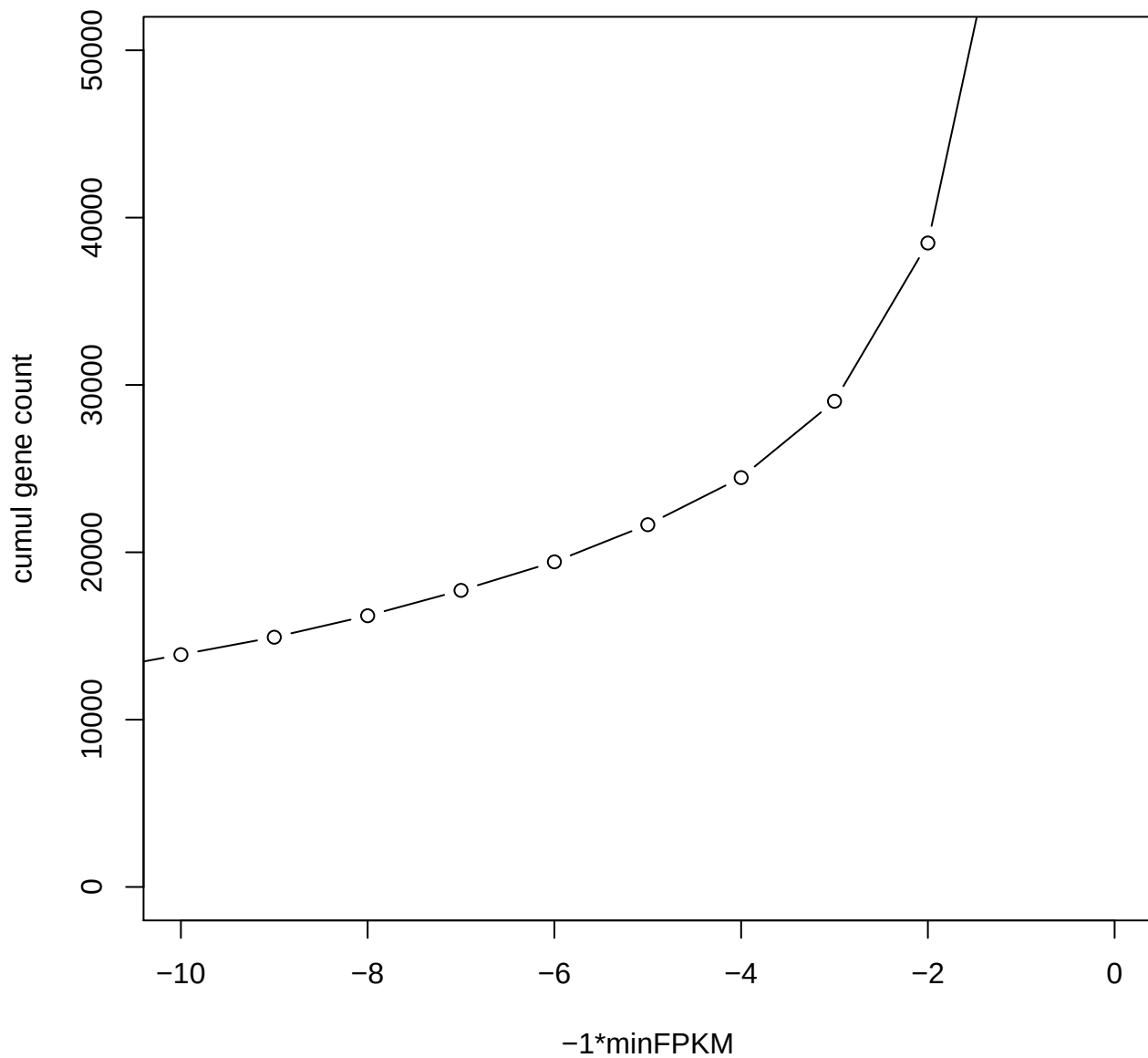
**Sample\_6 (gene count vs. minFPKM)**



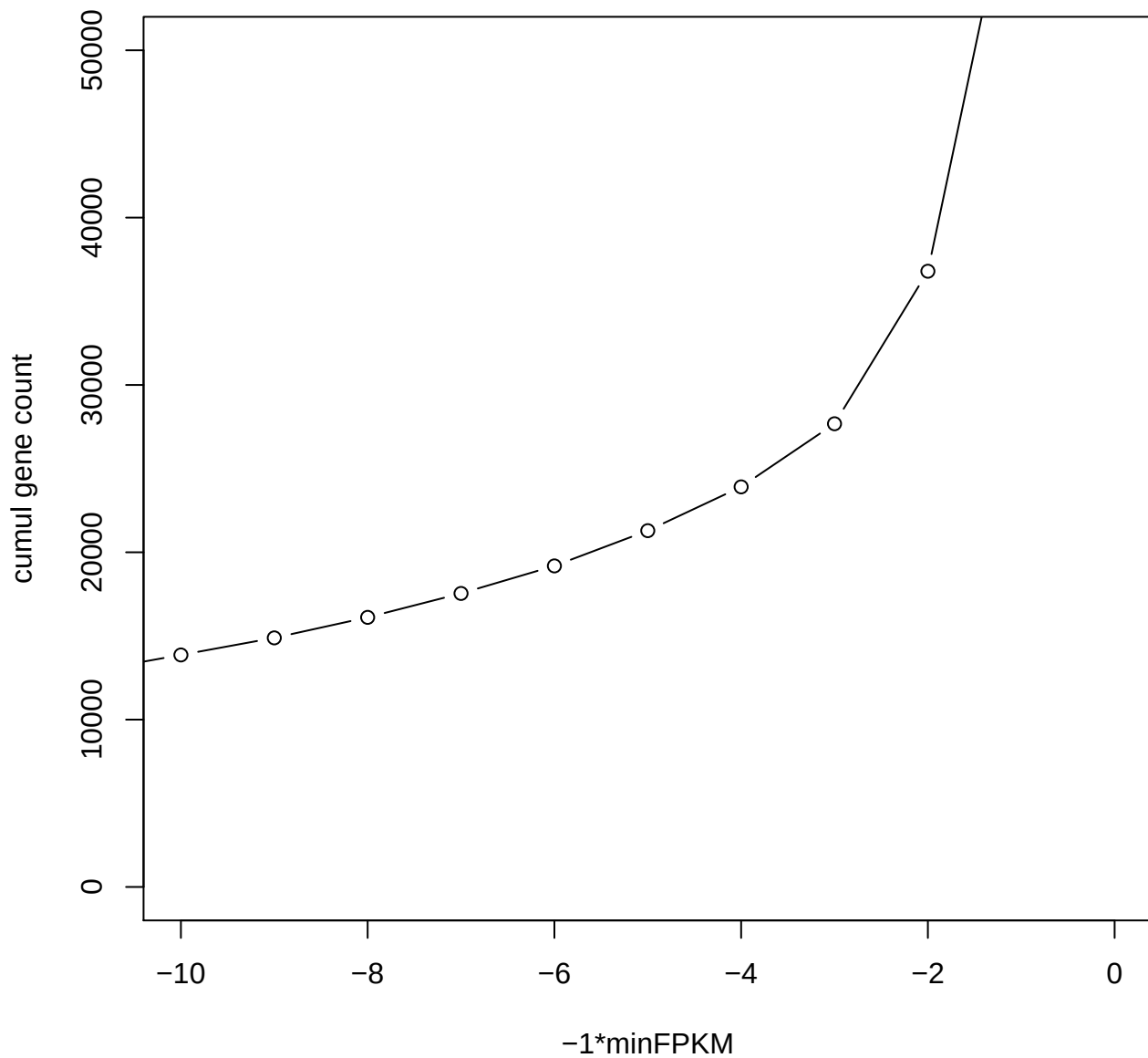
**Sample\_7 (gene count vs. minFPKM)**



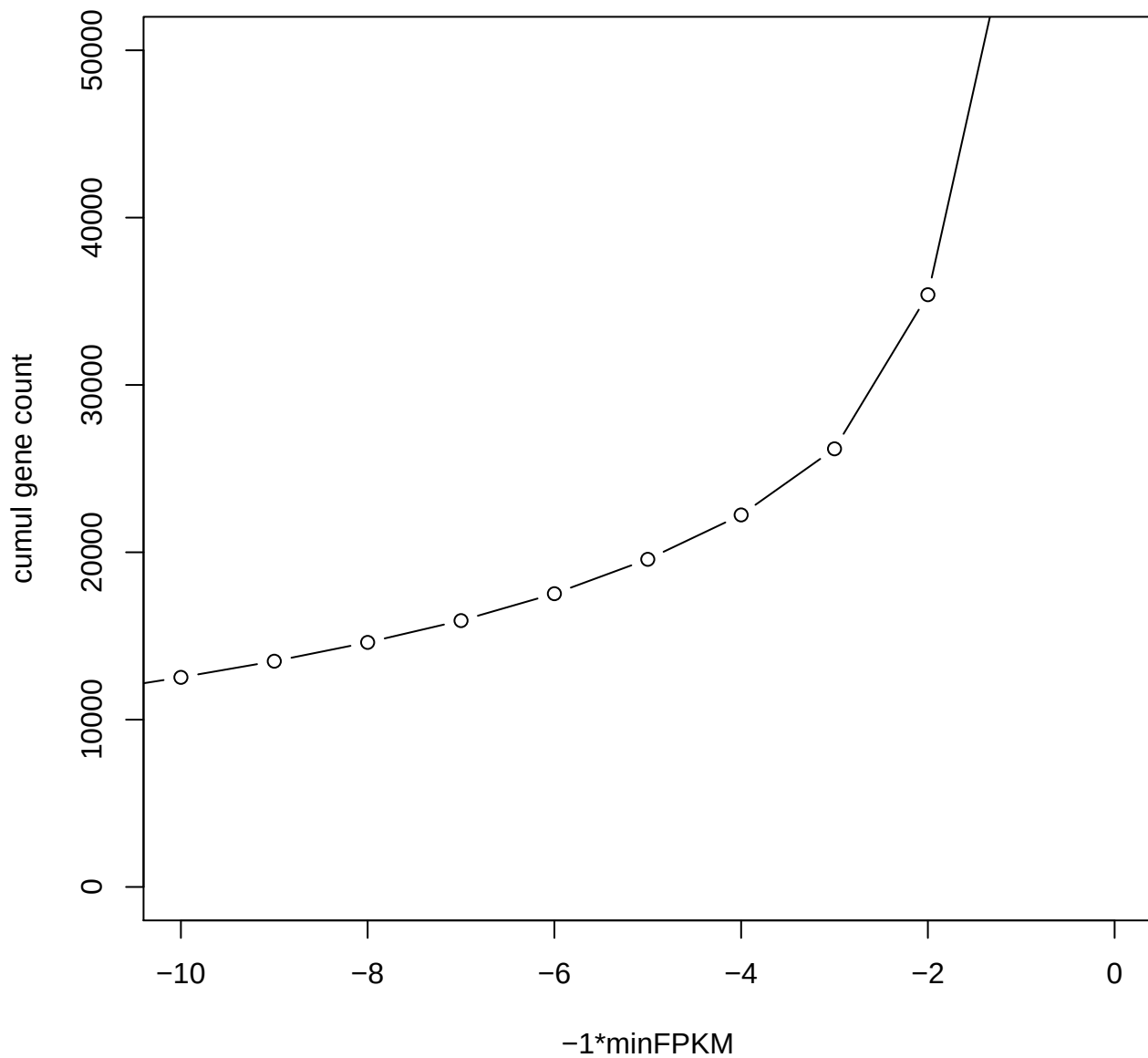
Sample\_8 (gene count vs. minFPKM)



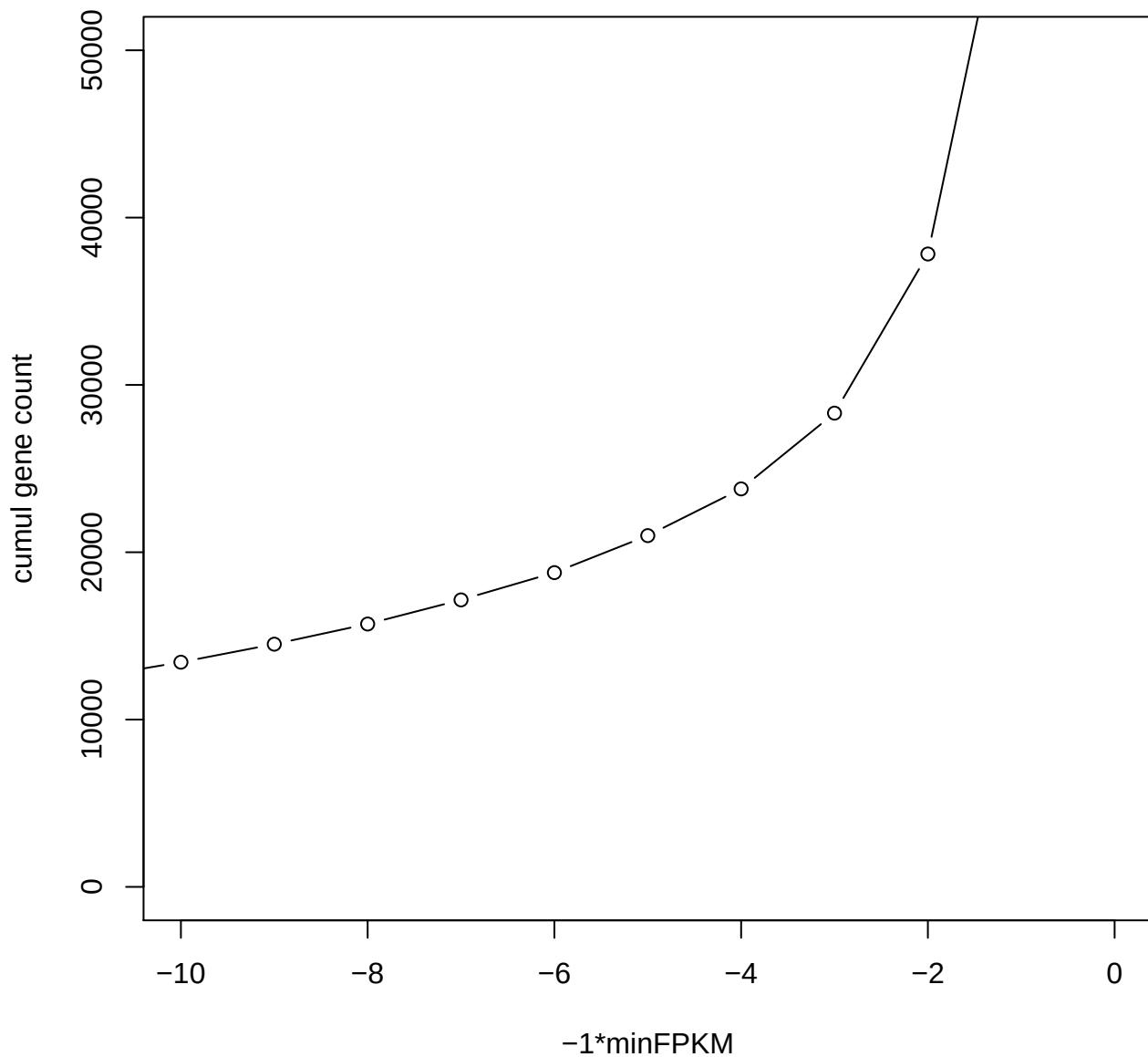
Sample\_9 (gene count vs. minFPKM)



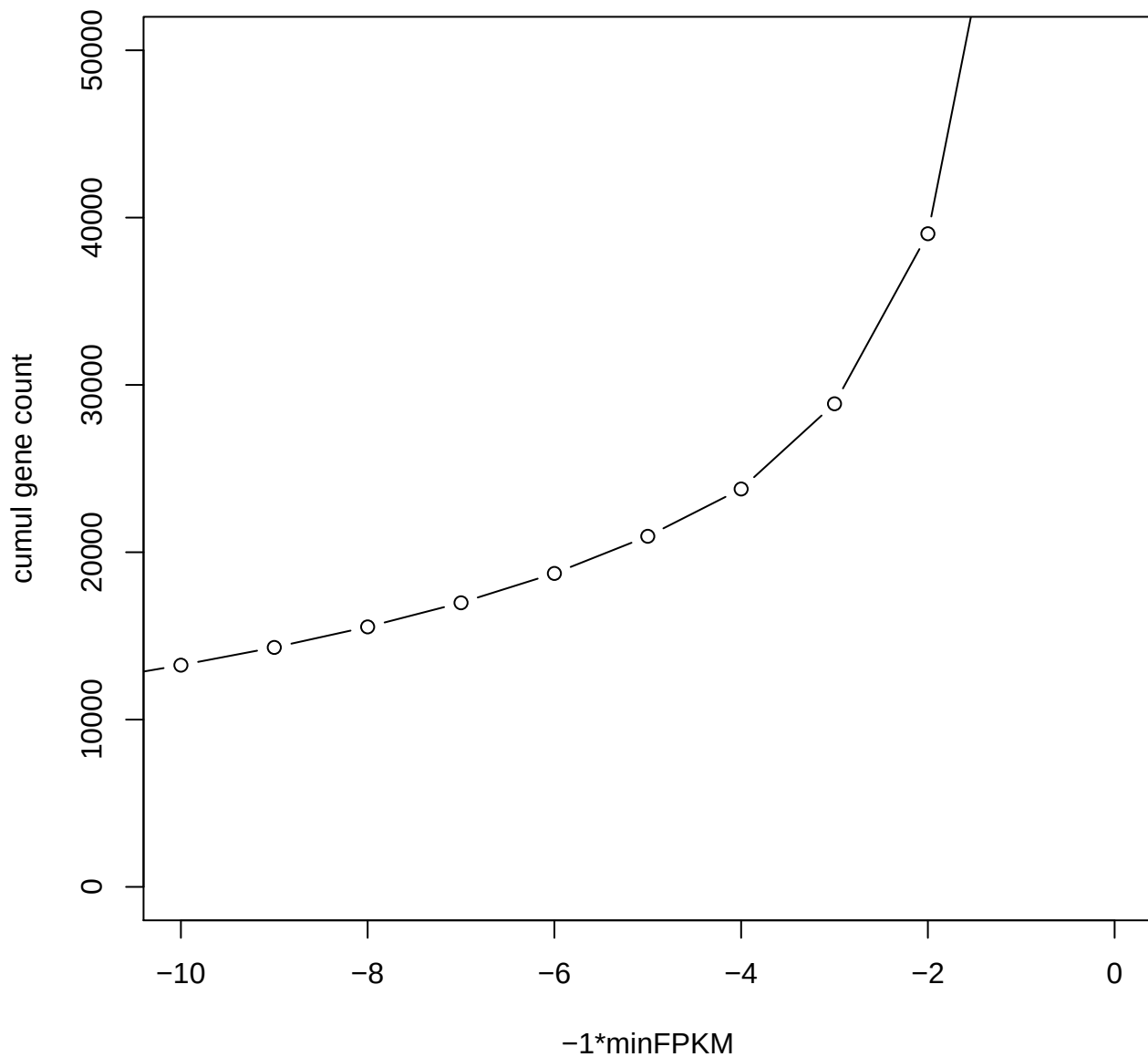
**Sample\_10 (gene count vs. minFPKM)**



**Sample\_11 (gene count vs. minFPKM)**



Sample\_12 (gene count vs. minFPKM)





Sample\_13 (gene count vs. minFPKM)

