Details of Full-length Fraction Analysis of *X. laevis* UniGene Libraries

UniGene libraries listed in order of the proportion of distinct sequences with full-length inserts (%FLDiver, column 6). The ratio of 'know-gene-complete ESTs' / 'known-gene-complete ESTs plus known-gene-incomplete ESTs' (%CDS-OK; third column) is the fraction of full-length sequences in the library. Libraries with a full-length fraction of greater than 40% are worth further consideration. The number of distinct full-length clones expected per 100 ESTs (%FLDiver; sixth column) is the library's full-length fraction (%CDS-OK) times the library's diversity (%Divers). For libraries with a %CDS-OK greater than 40%, this figure provides that best indication that a library has many distinct full-length inserts.

Seqs	%Divers	%CDS-OK	InLen	AlnLen	%FLDivers	Library
1128	76.7	0.917	2977.9	346.8	70.3339	NICHD_XGC_Ov1
17	64.7	1.000	1892.5	514.5	64.7	*Xenopus*_laevis_ZAP_Express_endodermal_cDNA_library
1430	64.0	1.000	1638.0	329.5	64	RIKEN_*Xenopus*_egg
3552	132.1	0.419	1209.8	358.0	55.3499	*Xenopus*_laevis_oocyte
458	69.4	0.737	1447.5	419.5	51.1478	Wellcome_CRC_pRN3_St10_5
63	60.3	0.750	1493.5	629.5	45.225	Harland_stage_19-23_*Xenopus*_laevis_cDNA
1731	68.4	0.639	1950.4	405.1	43.7076	NICHD_XGC_Emb2
1538	53.4	0.750	1786.2	395.1	40.05	Wellcome_CRC_pRN3_oocyte
460	95.7	0.412	1257.2	409.3	39.4284	*Xenopus*_neurula_plasmid_library
1917	51.1	0.737	1923.8	329.4	37.6607	*Xenopus*_EST_library
4697	53.7	0.662	1531.7	332.5	35.5494	NICHD_XGC_Sp1
3527	54.8	0.643	2279.2	343.4	35.2364	NICHD_XGC_OO1
2389	50.8	0.673	1396.4	436.6	34.1884	Wellcome_CRC_pRN3_St19_26
2937	60.2	0.565	1829.4	339.9	34.013	NICHD_XGC_Emb1
5520	51.5	0.650	2100.6	338.9	33.475	NICHD_XGC_Emb4
2636	51.8	0.625	1363.6	414.7	32.375	Wellcome_CRC_pRN3_head
1711	53.9	0.574	1476.2	399.9	30.9386	Wellcome_CRC_pRN3_dorsal_lip
2807	53.9	0.558	1421.0	392.8	30.0762	Harland_stage_19-23
805	50.0	0.520	1465.4	431.6	26	Xenla_13LiCl
5452	41.4	0.615	1364.2	329.4	25.461	NICHD_XGC_Lu1
2709	51.2	0.485	1991.7	404.9	24.832	Wellcome_CRC_pSK_egg
7686	47.3	0.508	1638.0	471.0	24.0284	NIBB_Mochii_normalized_*Xenopus*_early_gastrula_library
736	45.6	0.526	1408.6	286.8	23.9856	Cho_Li_treated_gastrula
2862	55.9	0.429	1814.8	330.6	23.9811	NICHD_XGC_Emb3
2545	56.2	0.426	1766.3	359.4	23.9412	*Xenopus*_laevis_oocyte_non_normalized
2329	53.5	0.438	1255.8	399.9	23.433	Kirschner_embryo_St10_14
3551	35.8	0.644	1382.8	368.8	23.0552	NICHD_XGC_Li1
2892	54.0	0.413	1699.0	380.5	22.302	*Xenopus*_laevis_unfertilized_egg_cDNA_library
1985	38.8	0.561	1428.9	361.2	21.7668	Wellcome_CRC_pRN3_St13_17_egg_animal_cap
28718	35.7	0.551	1702.3	468.1	19.6707	NIBB_Mochii_normalized_*Xenopus*_neurula_library
3522	50.4	0.356	1130.5	333.1	17.9424	*Xenopus*_laevis_gastrula_non_normalized
469	41.9	0.368	1288.8	381.2	15.4192	Wellcome_CRC_pCS2+_st19-26
2383	53.3	0.281	1377.9	376.0	14.9773	Wellcome_CRC_pSK_animal_cap
1585	66.7	0.222	1004.1	308.6	14.8074	normalized_*Xenopus*_laevis_gastrula
35547	35.7	0.401	1642.3	490.3	14.3157	NIBB_Mochii_normalized_*Xenopus*_tailbud_library
3451	32.4	0.426	1519.5	308.2	13.8024	NICHD_XGC_He1
1023	54.1	0.250	1073.4	385.2	13.525	Wellcome_CRC_pSK_St_10_5
2277	47.0	0.255	1186.6	391.7	11.985	Wellcome_CRC_pcDNAI_egg
17678	28.6	0.142	1077.4	442.7	4.0612	Blackshear/Soares_normalized_*Xenopus*_egg_library

Definitions:

Seqs:	Sequences; Number of EST sequences
%Divers:	Diversity; 100 X Genes/Clones
CDS-OK:	Coding Sequence; Fraction of 5' ESTs which are CDS-complete by comparison with Reference Sequences
InLen:	Insert Length; Mean predicted insert length of those clones matching Reference Sequences
AlLen:	Alignment Length; Observed alignment length of those clones matching Reference Sequences
%FLDivers:	CDS-OK X %Divers; figure of merit for MGC libraries