A Defined Zebrafish Line for High-Throughput Genetics and Genomics: NHGRI-1

Matthew C. LaFave,* Gaurav K. Varshney,* Meghana Vemulapalli,† James C. Mullikin,†,‡ and Shawn M. Burgess*,[1]

*Developmental Genomics Section, Translational and Functional Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, and †National Institutes of Health Intramural Sequencing Center, and ‡Comparative Genomics Analysis Unit, Cancer Genetics and Comparative Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20814

ABSTRACT Substantial intrastrain variation at the nucleotide level complicates molecular and genetic studies in zebrafish, such as the use of CRISPRs or morpholinos to inactivate genes. In the absence of robust inbred zebrafish lines, we generated NHGRI-1, a healthy and fecund strain derived from founder parents we sequenced to a depth of ~50×. Within this strain, we have identified the majority of the genome that matches the reference sequence and documented most of the variants. This strain has utility for many reasons, but in particular it will be useful for any researcher who needs to know the exact sequence (with all variants) of a particular genomic region or who wants to be able to robustly map sequences back to a genome with all possible variants defined.

THE zebrafish (*Danio rerio*) is a powerful tool for understanding vertebrate biology. The usefulness of this model organism is bolstered by the availability of a "finished" sequenced and annotated genome (Howe *et al.* 2013; Flicek *et al.* 2014). As a natural extension of this resource, there are several high-throughput efforts to systematically mutagenize all zebrafish protein-coding genes (Moens *et al.* 2008; Kettleborough *et al.* 2013; Varshney *et al.* 2013a,b).

In addition to such projects, the combination of a sequenced genome and developments in targeted nuclease technology mean that the zebrafish community is now able to rapidly take advantage of custom genome-editing technologies (Doyon *et al.* 2008; Bedell *et al.* 2012; Hruscha *et al.* 2013; Hwang *et al.* 2013; Jao *et al.* 2013). CRISPRs in particular provide an efficient, easy, and inexpensive means of manipulating and interrogating the genome (Jinek *et al.* 2012; Cong *et al.* 2013; Mali *et al.* 2013). However, because there are very few hardy inbred zebrafish lines (overinbreeding tends to result in unhealthy stocks) and polymorphism rates are close to 1 every 100 bases, variants frequently have the potential to interfere with target site design (Stickney *et al.* 2002; Guryev *et al.* 2006; Bowen *et al.* 2012) or with regions of homology used for homologous recombination. In general, genome targeting is heavily dependent on an exact match to the primary sequence. Depending on the sequence, even a single mismatch can severely reduce the cutting efficiency (Hsu *et al.* 2013). In addition, other techniques such as RNA-Seq or ChIP-Seq are substantially less accurate without having fully characterized variants in the background strain. Therefore, it is preferable to carry out studies in a zebrafish strain in which the regions of invariant sequence are known with a high degree of confidence and all variants are categorized to allow for robust genomic mapping.

With these concerns in mind, we derived the zebrafish line NHGRI-1. NHGRI-1 fish were derived from an original strain known as "TAB-5" made from a hybrid cross between fish from two of the most commonly used zebrafish lines: Tübingen and AB (Streisinger *et al.* 1981; Haffter *et al.* 1996). The F_1 fish from this cross were inbred and screened to be clear of any mutations affecting the first 5 days of development. Since its initial isolation in 1997, we have carried the strain in the laboratory until the present day without introducing other outside genetic diversity. We selected several mating pairs from the TAB-5 pool, and the most robust mating pair was chosen as the founding pair

doi: 10.1534/genetics.114.166769
Manuscript received May 30, 2014; accepted for publication June 28, 2014; published Early Online July 9, 2014.
Available freely online through the author-supported open access option.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.166769/-/DC1.
[1]Corresponding author: NHGRI, National Institutes of Health, Bldg. 50, Room 5537, 50 South Dr., Bethesda, MD 20892-8004. E-mail: burgess@mail.nih.gov

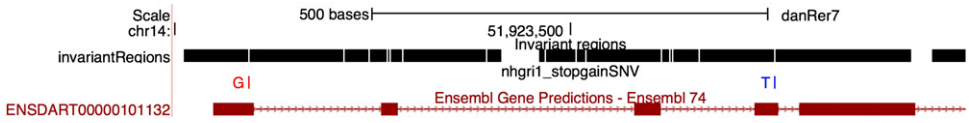

Figure 1 Screenshot of the UCSC browser custom tracks for NHGRI-1. Twenty mating pairs from 6-month-old TAB-5 fish were screened to select a robust founding pair with good clutch size and healthy progeny; the most fecund pair was renamed NHGRI-1. Fin clips from the NHGRI-1 male and female were prepared as separate genomic DNA libraries and sequenced on the Illumina HiSeq 2000 by the National Institutes of Health (NIH) Intramural Sequencing Center. Both libraries were subjected to paired-end sequencing with 101-bp reads. We aligned the sequence to the zebrafish genome [Zv9 (Howe *et al.* 2013)] with Novoalign version 2.08.02 (http://www.novocraft.com/). We removed PCR duplicates via SAMtools version 0.1.18 (Li *et al.* 2009). We used bam2mpg to identify the most probable genotype (MPG) for nucleotides in both parents (Teer *et al.* 2010). Bases that did not have an MPG score of at least 10, coverage of at least 20×, and a ratio of MPG score to coverage >0.5 were discarded. Regions of low sequence complexity were not specifically excluded from the analysis unless they failed to meet these criteria. The bases that matched the reference and met the above criteria in both fish were used to build the BED track of invariant nucleotides. The top track indicates the bases that were invariant in both fish sequenced. The white regions indicate either variation in at least one fish or insufficient read depth to confidently call the region as invariant. The second track indicates two nonsense mutations detected in this region. The letter indicates the alternative allele, and the color indicates whether the mutation was homozygous (red) or heterozygous (blue) in the NHGRI-1 population. Both tracks are available on the ZebrafishGenomics track hub, which is hosted at http://research.nhgri.nih.gov/manuscripts/Burgess/zebrafish/downloads/NHGRI-1/hub.txt and accessible through http://genome.ucsc.edu/cgi-bin/hgHubConnect.

for NHGRI-1. We are now on the third generation of NHGRI-1 and their fecundity and overall health remain strong.

We carried out high-throughput sequencing to a depth of ~50× for each parent. The male and female sequencing libraries had a combined 1,289,142,362 nonduplicate reads, with a median coverage of 52× and 47×, respectively. By doing so, we identified >10 million previously unreported single-nucleotide variants (SNVs). The raw sequence data have been deposited in the NCBI Sequence Read Archive [BioProject ID: 246102]. In addition, we have identified nearly all the regions of the genome that are invariant relative to the Zv9 reference sequence. We generated a browser extensible data (BED) file of invariant nucleotides, which indicates the regions in which there were both a lack of alternative alleles and a lack of sufficient read depth and genotype confidence to call bases as invariant (Figure 1). Seventy-one percent of the genome fits these criteria. The invariant file is hosted on the NHGRI-1 website at http://research.nhgri.nih.gov/manuscripts/Burgess/zebrafish/download.shtml, a University of California, Santa Cruz (UCSC) data hub called "ZebrafishGenomics" has been established at http://genome.ucsc.edu/cgi-bin/hgHubConnect, and data have been transferred to http://zfin.org/. Information on the variants themselves can be downloaded from dbSNP (submitter handle, NHGRI_DGS; submitter batch ID, NHGRI-1_founders). The invariant regions are easily identified by using the BED file, simplifying the design of CRISPR targets, amplicon primers, finding regions for homologous recombination, Morpholino design, or essentially any experiment that requires high confidence in the exact sequence of the genomic region of interest.

We detected >17 million total variants upon merging the variant calls from the two libraries. Of that total, 236,301 were in exons of Ensembl transcripts (Table 1). Variants were called as homozygous only if they were homozygous in both fish; such variants will stably retain the variant allele in future generations.

To underscore the issues related to background variation in the commonly used zebrafish lines, we detected 669 variants that formed premature stop codons in at least one transcript, 105 of which were homozygous mutant in both sexes (Table 2). We have generated a BED track of these variants, indicating the location, the alternative allele, and the homo/heterozygosity. This track is available on the ZebrafishGenomics hub and the NHGRI-1 website (Figure 1). A list of affected genes can also be found in supporting information, Table S1.

We detected 3160 deletion or insertion variants (DIVs) in exons. DIVs of a length divisible by three were highly represented and comprise ~60% of the DIVs (Figure 2A). Presumably, this is because the resultant nonframeshift mutations would be less likely to be selected against than those that produce frameshifts. A similar profile has been reported in human indels (Chen *et al.* 2007). This trend is not present in the genome-wide set of 2,210,080 NHGRI-1 DIVs (Figure 2B).

We compared the SNVs identified in NHGRI-1 with dbSNP (Build ID: 139) and a publically available data set

Table 1 Raw counts of variants in NHGRI-1

Variants	SNV	DIV	Total
Total variants	14,917,339	2,210,080	17,127,419
Heterozygous	12,245,715	1,953,277	14,198,992
Homozygous	2,642,908	225,347	2,868,255
Unknown	28,716	31,456	60,172
Exon variants	233,141	3,160	236,301
Heterozygous	190,626	2,815	193,441
Homozygous	42,153	311	42,464
Unknown	362	34	396

Single-nucleotide variants and deletion and insertion variants were annotated using ANNOVAR version 2012-10-16 (Wang *et al.* 2010). Our annotation used the ensGene track hosted on the UCSC genome browser, which corresponded to Ensembl release 74 (Flicek *et al.* 2014). We annotated the male and female fish separately and then combined the ANNOVAR output to determine overall homozygosity and heterozygosity. Variants were considered homozygous in NHGRI-1 only if they were independently called as homozygous in both sexes. We identified a variant as unknown if it was called as (1) unknown in both sexes or (2) unknown in one fish and homozygous reference in the other. All remaining variants were considered to be heterozygous in NHGRI-1, even if they were called as homozygous in one of the sexes. In cases in which deletion or insertion variants (DIVs) of different lengths were reported at the same position, both were counted as separate variants.

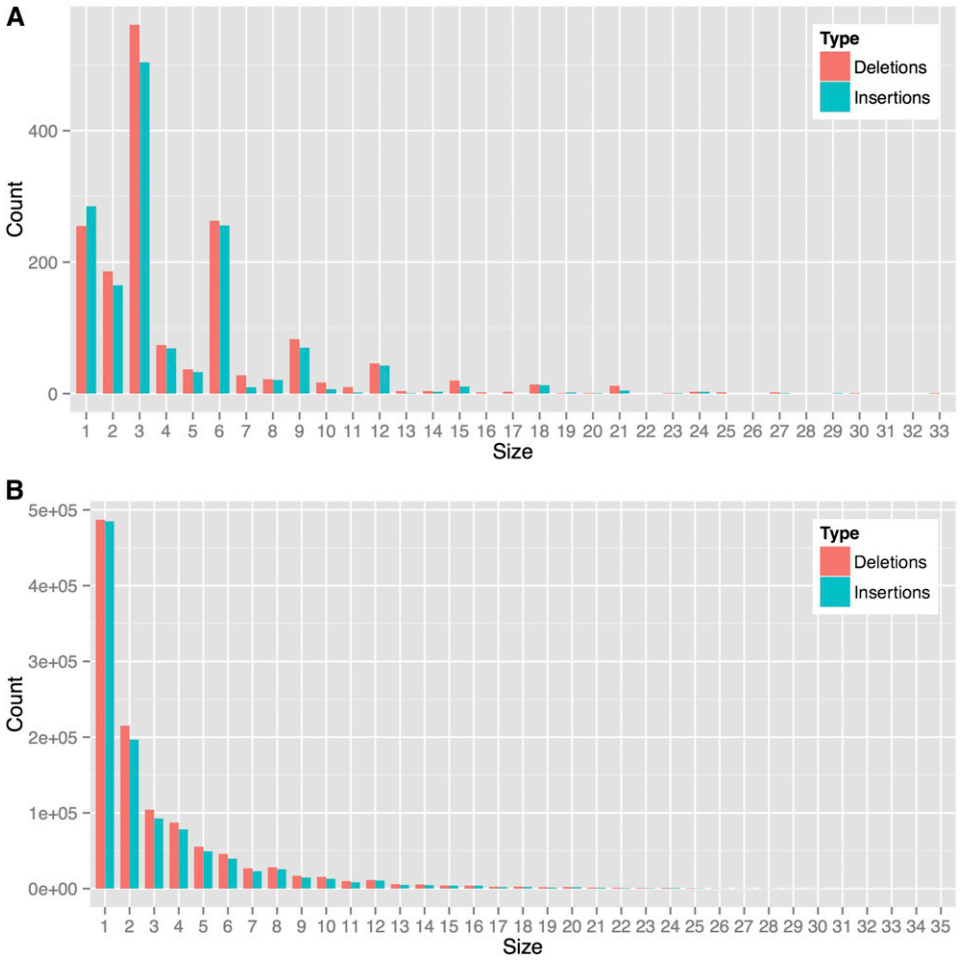

Figure 2 Deletion and insertion variant length distribution within exons. (A) The 3160 DIVs in exons. (B) The 2,210,080 DIVs detected genome-wide. Red bars indicate the number of deletions of a given length; blue bars represent insertions.

obtained from low-coverage sequencing of multiple zebrafish lines (Sherry *et al.* 2001; Bowen *et al.* 2012). For simplicity, we compared only biallelic SNVs for which the reference sequence is known (*i.e.*, no "N"s). The majority of NHGRI-1 SNVs had not been previously reported in either data set (Figure 3). We find that the rate of SNVs per sequenced base in NHGRI-1 is 0.01 or ~12.5–20× higher than the rate in humans (Kidd *et al.* 2008). It is important to note that, while the 0.01 number is relevant for NHGRI-1, the regions of homozygosity created by inbreeding mean it certainly underestimates the SNV load in zebrafish as a whole.

We also compared the mutational profile of NHGRI-1 to that reported for a zebrafish captured from the wild and sequenced at 39× coverage (Patowary *et al.* 2013). Different cutoffs had been applied for variant calling in said study, such as a minimum of 32 reads to call an SNV and 5 reads to call a DIV, but the ratios of variant types can be compared. The differences are statistically significant, but small. Among the SNVs in the wild zebrafish, 22.3% were reported as being homozygous, compared to 17.8% in NHGRI-1 (Fisher's exact test, $P < 2.2 \times 10^{-16}$). Deletions are more prevalent than insertions in both studies, with the wild zebrafish reported as having 53.9% deletions, compared to 51.6% in NHGRI-1 ($P < 2.2 \times 10^{-16}$).

This fish line will have utility in terms of automated design for targeted nucleases, as well as for studies such as ChIP-Seq or RNA-Seq where SNVs or DIVs might reduce the accuracy of mapping the raw sequence data. In addition,

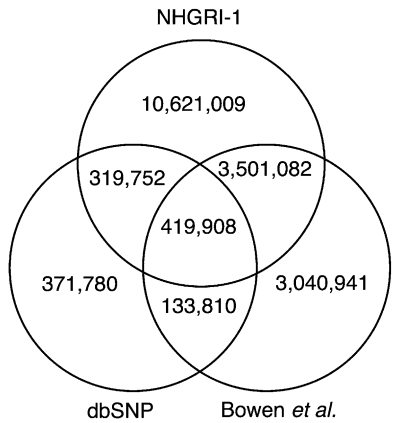

Figure 3 SNV overlap with publicly available data sets. This comparison incorporates only SNVs that were biallelic and for which the reference base was an unambiguous A, C, G, or T. The Bowen *et al.* (2012) SNVs were downloaded from http://fishbonelab.org/harris/Resources_files/parental_variants.tar; both data sets were downloaded on March 12th, 2014.

Table 2 Mutations introduced by variants in NHGRI-1

Annotations	Total
SNV annotation	
Nonsynonymous	77,791
Synonymous	149,378
Stop gain	640
Stop loss	90
Unknown	5,242
DIV annotation	
Frameshift deletion	638
Frameshift insertion	540
Nonframeshift insertion	944
Nonframeshift deletion	872
Stop gain	29
Stop loss	8
Unknown	129

techniques such as homologous recombination are very sensitive to variants (te Riele *et al.* 1992), and NHGRI-1 will allow researchers to target genomic regions that do not contain any variant nucleotides. Thus, NHGRI-1 will prove useful in a variety of circumstances where absolute knowledge of the possible sequence variation is needed. The line will be distributed by the Zebrafish International Resource Center (http://zebrafish.org) and the European Zebrafish Resource Center (http://www.ezrc.kit.edu).

Acknowledgments

This research was supported by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health.

Literature Cited

Bedell, V. M., Y. Wang, J. M. Campbell, T. L. Poshusta, C. G. Starker *et al.*, 2012 In vivo genome editing using a high-efficiency TALEN system. Nature 491: 114–118.

Bowen, M. E., K. Henke, K. R. Siegfried, M. L. Warman, and M. P. Harris, 2012 Efficient mapping and cloning of mutations in zebrafish by low-coverage whole-genome sequencing. Genetics 190: 1017–1024.

Chen, F. C., C. J. Chen, W. H. Li, and T. J. Chuang, 2007 Human-specific insertions and deletions inferred from mammalian genome sequences. Genome Res. 17: 16–22.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto *et al.*, 2013 Multiplex genome engineering using CRISPR/Cas systems. Science 339: 819–823.

Doyon, Y., J. M. McCammon, J. C. Miller, F. Faraji, C. Ngo *et al.*, 2008 Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26: 702–708.

Flicek, P., M. R. Amode, D. Barrell, K. Beal, K. Billis *et al.*, 2014 Ensembl 2014. Nucleic Acids Res. 42: D749–D755.

Guryev, V., M. J. Koudijs, E. Berezikov, S. L. Johnson, R. H. Plasterk *et al.*, 2006 Genetic variation in the zebrafish. Genome Res. 16: 491–497.

Haffter, P., M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt *et al.*, 1996 The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. Development 123: 1–36.

Howe, K., M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot *et al.*, 2013 The zebrafish reference genome sequence and its relationship to the human genome. Nature 496: 498–503.

Hruscha, A., P. Krawitz, A. Rechenberg, V. Heinrich, J. Hecht *et al.*, 2013 Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development 140: 4982–4987.

Hsu, P. D., D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann *et al.*, 2013 DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31: 827–832.

Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai *et al.*, 2013 Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31: 227–229.

Jao, L. E., S. R. Wente, and W. Chen, 2013 Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. USA 110: 13904–13909.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337: 816–821.

Kettleborough, R. N., E. M. Busch-Nentwich, S. A. Harvey, C. M. Dooley, E. de Bruijn *et al.*, 2013 A systematic genome-wide analysis of zebrafish protein-coding gene function. Nature 496: 494–497.

Kidd, J. M., G. M. Cooper, W. F. Donahue, H. S. Hayden, N. Sampas *et al.*, 2008 Mapping and sequencing of structural variation from eight human genomes. Nature 453: 56–64.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell *et al.*, 2013 RNA-guided human genome engineering via Cas9. Science 339: 823–826.

Moens, C. B., T. M. Donn, E. R. Wolf-Saxon, and T. P. Ma, 2008 Reverse genetics in zebrafish by TILLING. Brief. Funct. Genomics Proteomics 7: 454–459.

Patowary, A., R. Purkanti, M. Singh, R. Chauhan, A. R. Singh *et al.*, 2013 A sequence-based variation map of zebrafish. Zebrafish 10: 15–20.

Sherry, S. T., M. H. Ward, M. Kholodov, J. Baker, L. Phan *et al.*, 2001 dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 29: 308–311.

Stickney, H. L., J. Schmutz, I. G. Woods, C. C. Holtzer, M. C. Dickson *et al.*, 2002 Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. Genome Res. 12: 1929–1934.

Streisinger, G., C. Walker, N. Dower, D. Knauber, and F. Singer, 1981 Production of clones of homozygous diploid zebra fish (Brachydanio rerio). Nature 291: 293–296.

Teer, J. K., L. L. Bonnycastle, P. S. Chines, N. F. Hansen, N. Aoyama *et al.*, 2010 Systematic comparison of three genomic enrichment methods for massively parallel DNA sequencing. Genome Res. 20: 1420–1431.

te Riele, H., E. R. Maandag, and A. Berns, 1992 Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc. Natl. Acad. Sci. USA 89: 5128–5132.

Varshney, G. K., H. Huang, S. Zhang, J. Lu, D. E. Gildea *et al.*, 2013a The Zebrafish Insertion Collection (ZInC): a web based, searchable collection of zebrafish mutations generated by DNA insertion. Nucleic Acids Res. 41: D861–D864.

Varshney, G. K., J. Lu, D. E. Gildea, H. Huang, W. Pei *et al.*, 2013b A large-scale zebrafish gene knockout resource for the genome-wide study of gene function. Genome Res. 23: 727–735.

Wang, K., M. Li, and H. Hakonarson, 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38: e164.

Communicating editor: D. Parichy

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.166769/-/DC1

A Defined Zebrafish Line for High-Throughput Genetics and Genomics: NHGRI-1

Matthew C. LaFave, Gaurav K. Varshney, Meghana Vemulapalli, James C. Mullikin, and Shawn M. Burgess

DOI: 10.1534/genetics.114.166769

Table S1 Genes with premature stop codons in at least one transcript in NHGRI-1

Ensembl Gene ID	Associated Gene Name
ENSDARG00000022730	aasdh
ENSDARG00000016818	abcg2d
ENSDARG00000062156	ABI3BP (1 of 2)
ENSDARG00000035891	acana
ENSDARG00000062478	ACKR3 (2 of 2)
ENSDARG00000075906	ADAD2
ENSDARG00000025046	AL935186.1
ENSDARG00000089423	AL953879.1
ENSDARG00000005800	ampd3a
ENSDARG00000089810	AMZ2 (2 of 2)
ENSDARG00000044365	angptl3
ENSDARG00000089207	ANKRD27
ENSDARG00000057790	ankrd6a
ENSDARG00000014594	anxa1b
ENSDARG00000023185	APOLD1 (2 of 4)
ENSDARG00000076434	ARHGAP22 (2 of 2)
ENSDARG00000068008	arhgap44
ENSDARG00000089096	ASCL4
ENSDARG00000007823	atf3
ENSDARG00000088925	ATP9A
ENSDARG00000042236	atrx
ENSDARG00000001939	b3gat3
ENSDARG00000004396	b3gnt5b
ENSDARG00000094579	b4galnt2.1
ENSDARG00000080009	bahcc1
ENSDARG00000012699	BBX
ENSDARG00000079144	bcl2l11
ENSDARG00000011998	bfsp2
ENSDARG00000017661	braf
ENSDARG00000020076	brdt
ENSDARG00000079161	BSN (1 of 3)
ENSDARG00000086831	BX000701.1
ENSDARG00000090375	BX001033.2
ENSDARG00000086181	BX005052.1
ENSDARG00000077575	BX005223.1
ENSDARG00000089533	BX005484.1
ENSDARG00000074637	BX088696.1
ENSDARG00000079746	BX248122.1
ENSDARG00000090277	BX248410.3
ENSDARG00000079924	BX284672.1

ENSDARG00000087019	BX323444.2
ENSDARG00000091404	BX323836.1
ENSDARG00000077309	BX323861.1
ENSDARG00000077466	BX324215.2
ENSDARG00000057305	BX547930.1
ENSDARG00000088772	BX548000.3
ENSDARG00000086955	BX640584.1
ENSDARG00000088419	BX649431.1
ENSDARG00000036128	BX927234.1
ENSDARG00000071140	BX927308.1
ENSDARG00000077651	BX927317.3
ENSDARG00000078007	BX936298.1
ENSDARG00000070301	BX950203.1
ENSDARG00000078192	C12orf55
ENSDARG00000005179	C17H14orf159
ENSDARG00000092903	C18H15orf48
ENSDARG00000069781	C18H15orf60 (1 of 2)
ENSDARG00000075765	C20H8orf82
ENSDARG00000031749	C3AR1
ENSDARG00000068122	C5H9orf171
ENSDARG00000075022	CABZ01030119.1
ENSDARG00000074139	CABZ01031892.1
ENSDARG00000086681	CABZ01035647.1
ENSDARG00000076082	CABZ01041962.1
ENSDARG00000011463	CABZ01044048.1
ENSDARG00000090262	CABZ01044069.1
ENSDARG00000078289	CABZ01046425.1
ENSDARG00000055841	CABZ01046427.1
ENSDARG00000090511	CABZ01048092.1
ENSDARG00000078362	CABZ01048373.1
ENSDARG00000052350	CABZ01050249.1
ENSDARG00000089840	CABZ01053219.1
ENSDARG00000091217	CABZ01056821.2
ENSDARG00000087005	CABZ01059415.1
ENSDARG00000087345	CABZ01059415.2
ENSDARG00000078758	CABZ01063402.1
ENSDARG00000023916	CABZ01067098.1
ENSDARG00000090497	CABZ01067150.2
ENSDARG00000041086	CABZ01071177.1
ENSDARG00000087696	CABZ01073032.1
ENSDARG00000076587	CABZ01073417.1
ENSDARG00000091743	CABZ01074712.1
ENSDARG00000090396	CABZ01074946.1

ENSDARG00000088575	CABZ01077140.1
ENSDARG00000090419	CABZ01078499.2
ENSDARG00000075118	CABZ01079192.1
ENSDARG00000090860	CABZ01079541.1
ENSDARG00000089798	CABZ01080471.1
ENSDARG00000059047	CABZ01085069.1
ENSDARG00000069029	CABZ01087853.1
ENSDARG00000026651	CABZ01113978.1
ENSDARG00000095564	CABZ01118769.1
ENSDARG00000023683	cacna1f
ENSDARG00000026855	cacna2d4a
ENSDARG00000076290	calrl2
ENSDARG00000095701	card14
ENSDARG00000035103	caskb
ENSDARG00000054136	CCDC105
ENSDARG00000004794	ccdc146
ENSDARG00000054446	ccdc43
ENSDARG00000074386	CDH12 (1 of 2)
ENSDARG00000078404	CDH26
ENSDARG00000068586	CDHR2 (1 of 2)
ENSDARG00000041119	ceacam1
ENSDARG00000020777	cep85l
ENSDARG00000041569	ces2
ENSDARG00000075543	chd8
ENSDARG00000071298	chrm3b
ENSDARG00000061267	cidea
ENSDARG00000088825	CIT (1 of 2)
ENSDARG00000003145	CLMP
ENSDARG00000037865	cln3
ENSDARG00000078451	CMA1 (10 of 29)
ENSDARG00000095175	CMA1 (24 of 29)
ENSDARG00000069415	col17a1a
ENSDARG00000076163	col19a1
ENSDARG00000079330	COL26A1
ENSDARG00000059759	coq4
ENSDARG00000095275	CR356223.6
ENSDARG00000076235	CR376803.1
ENSDARG00000069625	CR382337.1
ENSDARG00000088623	CR388079.2
ENSDARG00000089959	CR388164.2
ENSDARG00000086243	CR391995.1
ENSDARG00000076396	CR751224.1
ENSDARG00000088911	CR774179.2

ENSDARG00000075179	CR848663.1
ENSDARG00000001712	CR855338.1
ENSDARG00000089113	CR927050.1
ENSDARG00000076076	crtc1b
ENSDARG00000058950	CSPP1
ENSDARG00000035665	CT027569.1
ENSDARG00000037253	CT573253.1
ENSDARG00000092366	CT573356.2
ENSDARG00000090579	CT583700.1
ENSDARG00000038568	CT955963.1
ENSDARG00000088893	CTRL (3 of 3)
ENSDARG00000052855	ctso
ENSDARG00000043081	ctsz
ENSDARG00000090267	CU062628.1
ENSDARG00000012535	CU184832.2
ENSDARG00000036575	CU424462.2
ENSDARG00000058564	CU457819.1
ENSDARG00000021678	CU463316.1
ENSDARG00000054955	CU570689.1
ENSDARG00000086084	CU571170.2
ENSDARG00000042583	CU596012.1
ENSDARG00000087227	CU856520.1
ENSDARG00000088542	CU896555.1
ENSDARG00000087448	CU929143.1
ENSDARG00000088946	CU929354.1
ENSDARG00000041799	cx43
ENSDARG00000069451	cx50.5
ENSDARG00000042978	cyp2p6
ENSDARG00000079653	cyp2x6
ENSDARG00000056753	dctn1b
ENSDARG00000022177	ddx59
ENSDARG00000007347	degs1
ENSDARG00000030250	dennd2db
ENSDARG00000068468	DNAH3
ENSDARG00000015566	dnmt3ab
ENSDARG00000036012	DOCK1 (1 of 3)
ENSDARG00000076213	DOCK4 (2 of 2)
ENSDARG00000078675	dock7
ENSDARG00000053548	dpf1
ENSDARG00000024748	DST
ENSDARG00000075504	DTX3L
ENSDARG00000059695	DUSP28
ENSDARG00000014717	DYNC1H1

ENSDARG00000057353	EHBP1L1 (1 of 2)
ENSDARG00000062056	elmod1
ENSDARG00000063613	ENDOD1 (3 of 15)
ENSDARG00000039007	eno3
ENSDARG00000019917	epb41l3b
ENSDARG00000021859	ERAP1 (2 of 2)
ENSDARG00000021013	f3a
ENSDARG00000015247	f8
ENSDARG00000077696	FAM188B
ENSDARG00000070216	FANCA
ENSDARG00000019250	fancf
ENSDARG00000061631	fbxo11a
ENSDARG00000059115	fbxo7
ENSDARG00000004782	fgfr3
ENSDARG00000075519	FMNL1 (2 of 2)
ENSDARG00000070389	foxf2b
ENSDARG00000091355	FP074874.2
ENSDARG00000086257	FP102463.1
ENSDARG00000074247	FP102786.1
ENSDARG00000090551	FP102786.2
ENSDARG00000086323	FP243385.2
ENSDARG00000076856	frem2a
ENSDARG00000060622	FRRS1 (1 of 5)
ENSDARG00000003909	ftr01
ENSDARG00000077412	ftr08
ENSDARG00000079537	ftr13
ENSDARG00000092745	ftr29
ENSDARG00000093203	ftr32
ENSDARG00000073880	ftr41
ENSDARG00000076106	ftr43
ENSDARG00000069508	ftr46
ENSDARG00000095609	ftr60p
ENSDARG00000016725	GADD45G (1 of 3)
ENSDARG00000069543	galt
ENSDARG00000036239	gatm
ENSDARG00000003244	gbp3
ENSDARG00000018329	gc2
ENSDARG00000055086	ggnbp2
ENSDARG00000075757	gig2e
ENSDARG00000075647	GIGYF1 (1 of 3)
ENSDARG00000053644	got1l1
ENSDARG00000045481	gsap
ENSDARG00000036494	gstcd

ENSDARG00000042983	has1
ENSDARG00000019156	haus5
ENSDARG00000042877	heca
ENSDARG00000059231	hephl1
ENSDARG00000063594	HIPK1 (1 of 2)
ENSDARG00000037154	hivep3b
ENSDARG00000078472	hsf5
ENSDARG00000076564	hspg2
ENSDARG00000092625	HTRA2 (15 of 31)
ENSDARG00000094236	HTRA2 (21 of 31)
ENSDARG00000016405	hug
ENSDARG00000074424	ibtk
ENSDARG00000096387	ighv1-1
ENSDARG00000096350	ighv1-2
ENSDARG00000096355	ighv1-4
ENSDARG00000096280	ighz
ENSDARG00000052417	im:6904045
ENSDARG00000087779	im:7148292
ENSDARG00000004925	immp2l
ENSDARG00000024964	inadl
ENSDARG00000059631	ints1
ENSDARG00000013025	ints6
ENSDARG00000092483	iqcc
ENSDARG00000069844	irg1
ENSDARG00000068657	irgq2
ENSDARG00000057787	ITGAE (1 of 2)
ENSDARG00000006891	KCNS3 (1 of 2)
ENSDARG00000091613	KCNU1
ENSDARG00000059653	kdm2aa
ENSDARG00000092488	KHNYN
ENSDARG00000015016	kif26ab
ENSDARG00000056133	kitb
ENSDARG00000032802	ktn1
ENSDARG00000071462	lama3
ENSDARG00000076464	lamtor1
ENSDARG00000057075	lamtor3
ENSDARG00000010137	ldb1a
ENSDARG00000096334	LECT2 (4 of 4)
ENSDARG00000071212	lepre1
ENSDARG00000009693	llgl1
ENSDARG00000076053	lrp6
ENSDARG00000078283	LRRC8E
ENSDARG00000010400	lrrfip2

ENSDARG00000006052	malt1a
ENSDARG00000034396	mars
ENSDARG00000052978	mbnl1
ENSDARG00000032326	mecr
ENSDARG00000089236	METTL15
ENSDARG00000078878	METTL21C (2 of 2)
ENSDARG00000088745	MFAP4 (4 of 14)
ENSDARG00000086287	MGA (2 of 2)
ENSDARG00000002377	mief2
ENSDARG00000089107	MLXIPL
ENSDARG00000087446	mmd2a
ENSDARG00000026325	mmp11a
ENSDARG00000045887	mmp30
ENSDARG00000063694	mon2
ENSDARG00000031136	moxd1
ENSDARG00000012495	mphosph10
ENSDARG00000053453	mpp2a
ENSDARG00000053561	ms4a17a.11
ENSDARG00000034105	MTUS1 (1 of 2)
ENSDARG00000038658	MURC (1 of 2)
ENSDARG00000021688	mxa
ENSDARG00000087845	mxb
ENSDARG00000041776	ndst3
ENSDARG00000025335	NEK10
ENSDARG00000040008	neurod6a
ENSDARG00000075707	nid2a
ENSDARG00000022761	nitr1m
ENSDARG00000036308	nod1
ENSDARG00000059711	nol6
ENSDARG00000027538	NOXRED1
ENSDARG00000010712	npffr1l3
ENSDARG00000009046	nphp1
ENSDARG00000017007	nr4a2a
ENSDARG00000062372	nrd1
ENSDARG00000074726	nrde2
ENSDARG00000043746	nrxn3a
ENSDARG00000034852	nt5c2l1
ENSDARG00000016212	nup214
ENSDARG00000076110	nwd1
ENSDARG00000003934	nxnl2
ENSDARG00000024815	ogfrl2
ENSDARG00000096264	or124-4
ENSDARG00000091929	or126-1

ENSDARG00000095222	or128-7
ENSDARG00000005865	or131-1
ENSDARG00000094819	or133-10
ENSDARG00000054319	oxct1b
ENSDARG00000042440	p2rx7
ENSDARG00000062811	p2ry10
ENSDARG00000055875	pag1
ENSDARG00000018476	pak2b
ENSDARG00000007818	palm1b
ENSDARG00000059544	pappaa
ENSDARG00000024702	parn
ENSDARG00000016704	PBLD (1 of 2)
ENSDARG00000077038	pbrm1l
ENSDARG00000074738	pcdh2g9
ENSDARG00000033012	pcnt
ENSDARG00000067683	PDE1C (2 of 2)
ENSDARG00000090631	PDZD7 (2 of 2)
ENSDARG00000076974	pdzd7a
ENSDARG00000041511	pex10
ENSDARG00000030687	PHKA2
ENSDARG00000029905	phyhd1
ENSDARG00000062823	pi4kab
ENSDARG00000060841	pik3c2a
ENSDARG00000041100	PITPNM2
ENSDARG00000033029	pkd1b
ENSDARG00000045982	PLA2G4C (4 of 5)
ENSDARG00000070953	PLA2G4F (2 of 2)
ENSDARG00000058835	PLEKHG3 (1 of 2)
ENSDARG00000040851	plekhh2
ENSDARG00000079892	PLXNB2 (4 of 5)
ENSDARG00000015278	plxnc1
ENSDARG00000076542	pogzb
ENSDARG00000067798	ppapdc1a
ENSDARG00000053592	prdm1b
ENSDARG00000043511	prdx6
ENSDARG00000089259	prf1.8
ENSDARG00000010482	prg4a
ENSDARG00000010246	prmt1
ENSDARG00000069823	proca1
ENSDARG00000021154	PRODH2
ENSDARG00000039756	prop1
ENSDARG00000015524	prps1a
ENSDARG00000075849	prr12b

ENSDARG00000059269	PRTG (2 of 2)
ENSDARG00000006242	ptp4a1
ENSDARG00000073862	ptpn13
ENSDARG00000075459	ptpn20
ENSDARG00000051814	ptprz1a
ENSDARG00000002021	pygb
ENSDARG00000052454	rhpn1
ENSDARG00000004218	rnd1l
ENSDARG00000078802	RNF111
ENSDARG00000087197	ros1
ENSDARG00000056617	rpgra
ENSDARG00000076712	RPH3A (2 of 2)
ENSDARG00000057556	rpl17
ENSDARG00000057927	rps6ka3b
ENSDARG00000008947	rtf1
ENSDARG00000033586	ryroxd1
ENSDARG00000075891	sall1b
ENSDARG00000070502	sc:d0661
ENSDARG00000030265	scdb
ENSDARG00000017266	scml4
ENSDARG00000041873	SDR16C5
ENSDARG00000074156	sgca
ENSDARG00000041378	sh2d4a
ENSDARG00000023600	sh3gl2
ENSDARG00000052816	shmt1
ENSDARG00000053739	si:busm1-160c18.1
ENSDARG00000051791	si:busm1-169i8.11
ENSDARG00000092224	si:ch1073-181h11.1
ENSDARG00000077336	si:ch1073-181h11.2
ENSDARG00000076023	si:ch1073-209h18.1
ENSDARG00000096589	si:ch211-112b1.3
ENSDARG00000075379	si:ch211-113a14.19
ENSDARG00000078620	si:ch211-114l13.10
ENSDARG00000086498	si:ch211-116o3.3
ENSDARG00000077068	si:ch211-11p18.6
ENSDARG00000079459	si:ch211-122l14.3
ENSDARG00000060911	si:ch211-126i22.5
ENSDARG00000089477	si:ch211-132g1.3
ENSDARG00000093102	si:ch211-132g1.6
ENSDARG00000076809	si:ch211-141h20.6
ENSDARG00000042063	si:ch211-145c1.1
ENSDARG00000076534	si:ch211-14a17.10
ENSDARG00000095885	si:ch211-158m24.12

ENSDARG00000045752	si:ch211-161n3.4
ENSDARG00000044937	si:ch211-163m16.6
ENSDARG00000096024	si:ch211-178j18.4
ENSDARG00000096545	si:ch211-191a16.5
ENSDARG00000090532	si:ch211-197e7.2
ENSDARG00000094747	si:ch211-197g15.6
ENSDARG00000054655	si:ch211-1f22.5
ENSDARG00000039468	si:ch211-1n9.6
ENSDARG00000095055	si:ch211-205a14.1
ENSDARG00000091287	si:ch211-207n2.6
ENSDARG00000071618	si:ch211-213a13.5
ENSDARG00000092364	si:ch211-218c6.8
ENSDARG00000091963	si:ch211-223g7.6
ENSDARG00000095101	si:ch211-229g14.3
ENSDARG00000087700	si:ch211-229l10.12
ENSDARG00000019686	si:ch211-236c15.2
ENSDARG00000086256	si:ch211-236p5.2
ENSDARG00000069038	si:ch211-242e8.1
ENSDARG00000091735	si:ch211-244h4.1
ENSDARG00000079603	si:ch211-256e16.6
ENSDARG00000091695	si:ch211-261i17.1
ENSDARG00000094098	si:ch211-278p9.3
ENSDARG00000070850	si:ch211-282j17.10
ENSDARG00000071050	si:ch211-8c17.4
ENSDARG00000040902	si:ch73-113g13.1
ENSDARG00000079050	si:ch73-268p17.1
ENSDARG00000088908	si:ch73-281f12.3
ENSDARG00000092164	si:ch73-299h12.1
ENSDARG00000039752	si:ch73-308m11.1
ENSDARG00000057826	si:ch73-61d6.3
ENSDARG00000070015	si:dkey-108d22.5
ENSDARG00000052332	si:dkey-117k10.8
ENSDARG00000089640	si:dkey-117n7.5
ENSDARG00000076224	si:dkey-11f4.16
ENSDARG00000074881	si:dkey-11n14.1
ENSDARG00000078101	si:dkey-121n8.3
ENSDARG00000070203	si:dkey-13e3.1
ENSDARG00000071029	si:dkey-147f3.4
ENSDARG00000090252	si:dkey-14o6.7
ENSDARG00000094384	si:dkey-15h8.12
ENSDARG00000094598	si:dkey-15h8.16
ENSDARG00000089841	si:dkey-16p6.1
ENSDARG00000079530	si:dkey-17m8.1

ENSDARG00000094797	si:dkey-181m19.11
ENSDARG00000094012	si:dkey-192g7.3
ENSDARG00000090734	si:dkey-196n19.2
ENSDARG00000096488	si:dkey-19c16.12
ENSDARG00000073965	si:dkey-202b22.5
ENSDARG00000058685	si:dkey-204a24.11
ENSDARG00000024708	si:dkey-204f11.59
ENSDARG00000071714	si:dkey-20i20.5
ENSDARG00000036728	si:dkey-211g8.1
ENSDARG00000092259	si:dkey-211g8.5
ENSDARG00000089602	si:dkey-217f16.1
ENSDARG00000096187	si:dkey-21h14.10
ENSDARG00000094603	si:dkey-21h14.12
ENSDARG00000096221	si:dkey-222o15.5
ENSDARG00000074519	si:dkey-224j12.3
ENSDARG00000059410	si:dkey-225n22.5
ENSDARG00000074062	si:dkey-234i14.15
ENSDARG00000079843	si:dkey-234i14.9
ENSDARG00000093588	si:dkey-234l24.8
ENSDARG00000057253	si:dkey-236a14.13
ENSDARG00000075315	si:dkey-237g15.2
ENSDARG00000031588	si:dkey-239b22.1
ENSDARG00000043518	si:dkey-239i20.2
ENSDARG00000011581	si:dkey-23k10.5
ENSDARG00000034227	si:dkey-243i1.1
ENSDARG00000096186	si:dkey-247i3.5
ENSDARG00000079762	si:dkey-253d23.4
ENSDARG00000074820	si:dkey-253d23.9
ENSDARG00000077260	si:dkey-28d5.13
ENSDARG00000077136	si:dkey-28d5.14
ENSDARG00000087639	si:dkey-33o22.1
ENSDARG00000094563	si:dkey-4c15.6
ENSDARG00000091753	si:dkey-56m15.3
ENSDARG00000094284	si:dkey-58f10.10
ENSDARG00000095532	si:dkey-58f10.13
ENSDARG00000092759	si:dkey-61p9.9
ENSDARG00000075561	si:dkey-76k16.5
ENSDARG00000094399	si:dkey-78k11.4
ENSDARG00000095126	si:dkey-82i20.2
ENSDARG00000093612	si:dkey-92i17.2
ENSDARG00000096577	si:dkey-9c18.1
ENSDARG00000087176	si:dkeyp-118e10.5
ENSDARG00000079526	si:dkeyp-120g4.1

ENSDARG00000041835	si:dkeyp-59a8.2
ENSDARG00000094442	si:dkeyp-73g8.5
ENSDARG00000036479	si:dkeyp-80c12.4
ENSDARG00000096059	si:dkeyp-85d8.3
ENSDARG00000022101	si:rp71-18a8.2
ENSDARG00000071045	si:rp71-1k13.6
ENSDARG00000093164	si:rp71-23d18.8
ENSDARG00000071662	si:rp71-36a1.3
ENSDARG00000090946	si:zfos-364h11.1
ENSDARG00000078157	SLC18A1
ENSDARG00000055445	SLC22A15
ENSDARG00000007449	slc25a39
ENSDARG00000077875	slc2a5
ENSDARG00000007180	slc30a4
ENSDARG00000036864	slc34a2b
ENSDARG00000021820	SLC45A4 (1 of 2)
ENSDARG00000063133	SLC4A10 (2 of 2)
ENSDARG00000067784	SLC9A1
ENSDARG00000037605	slco1f1
ENSDARG00000045469	slco1f4
ENSDARG00000039528	slit1b
ENSDARG00000012574	slkb
ENSDARG00000017744	smc2
ENSDARG00000038518	snap29
ENSDARG00000076283	snrpd3l
ENSDARG00000062745	socs5b
ENSDARG00000037476	sorbs3
ENSDARG00000004017	spag1a
ENSDARG00000039256	speg
ENSDARG00000061211	spg20b
ENSDARG00000079827	srbd1
ENSDARG00000060592	SRCRB4D
ENSDARG00000036584	st8sia5
ENSDARG00000091548	stard9
ENSDARG00000061060	STPG2
ENSDARG00000009499	syne1a
ENSDARG00000061956	SYTL2 (2 of 3)
ENSDARG00000062135	sytl5
ENSDARG00000092603	TAAR1 (42 of 92)
ENSDARG00000095395	TAAR1 (74 of 92)
ENSDARG00000041454	taar12b
ENSDARG00000054252	taar18d
ENSDARG00000093266	taar19f

ENSDARG00000076345	taar19p
ENSDARG00000093526	taar20a1
ENSDARG00000054261	taar20c
ENSDARG00000092138	taar20q
ENSDARG00000096106	taar20r
ENSDARG00000095156	taar20w
ENSDARG00000093829	taar20x
ENSDARG00000089190	TANC1 (2 of 2)
ENSDARG00000079055	TARBP1 (1 of 2)
ENSDARG00000025314	tcerg1a
ENSDARG00000036481	tcn2l
ENSDARG00000035633	tctn2
ENSDARG00000075230	tet1
ENSDARG00000061479	thsd7a
ENSDARG00000090496	THSD7A (2 of 2)
ENSDARG00000022048	tlr4bb
ENSDARG00000052723	TMC3
ENSDARG00000078744	tmem63bb
ENSDARG00000029841	tmprss9
ENSDARG00000041869	tnfrsf19
ENSDARG00000030270	tnnt3a
ENSDARG00000026895	tor1l3
ENSDARG00000042548	tpd52l1
ENSDARG00000073872	tpst1
ENSDARG00000034707	TRIM29 (14 of 87)
ENSDARG00000093263	TRIM29 (71 of 87)
ENSDARG00000075637	trim35-19
ENSDARG00000071225	trim35-22
ENSDARG00000052310	trim35-35
ENSDARG00000032169	trim35-36
ENSDARG00000029907	trim54
ENSDARG00000000370	triob
ENSDARG00000036678	TRMT10A
ENSDARG00000060837	trpc5b
ENSDARG00000095696	trpm2
ENSDARG00000036232	trpm7
ENSDARG00000053183	TSPAN17
ENSDARG00000074126	ttc39a
ENSDARG00000044405	ttc4
ENSDARG00000028213	ttna
ENSDARG00000069250	uaca
ENSDARG00000008200	ugp2b
ENSDARG00000062518	ulk1a

ENSDARG00000060308	URB1 (1 of 2)
ENSDARG00000028249	utp23
ENSDARG00000078680	vcana
ENSDARG00000089647	VDAC2
ENSDARG00000060477	vps8
ENSDARG00000053386	wdr27
ENSDARG00000095879	wdr46
ENSDARG00000076934	WDR60
ENSDARG00000036245	WIPF3
ENSDARG00000021611	WNK3
ENSDARG00000092091	wu:fd49b10
ENSDARG00000054847	xcr1a
ENSDARG00000038872	ZBED1 (1 of 11)
ENSDARG00000090671	ZBED1 (9 of 11)
ENSDARG00000044038	zbtb2a
ENSDARG00000091762	ZBTB40
ENSDARG00000074645	ZCCHC6
ENSDARG00000070800	zgc:109744
ENSDARG00000032263	zgc:110224
ENSDARG00000070948	zgc:123284
ENSDARG00000091320	zgc:136791
ENSDARG00000069058	zgc:136924
ENSDARG00000017038	zgc:152670
ENSDARG00000071044	zgc:153722
ENSDARG00000094736	zgc:162358
ENSDARG00000078093	zgc:172065
ENSDARG00000078728	zgc:173710
ENSDARG00000073760	zgc:174708
ENSDARG00000077176	zgc:174888
ENSDARG00000088686	zgc:195633
ENSDARG00000076981	zgc:198329
ENSDARG00000071155	zgc:64065
ENSDARG00000011376	zgc:77816
ENSDARG00000076287	ZMYM1
ENSDARG00000057726	znf185
ENSDARG00000016763	znf292a
ENSDARG00000076988	znf839