An Interrogation of Shared and Unique Copy Number Variants Across Genetically Distinct Zebrafish Strains

Lindsay A. Holden,^{1,*} Charles Wilson,¹ Zachary Heineman,¹ Kimberly P. Dobrinski,² and Kim H. Brown¹

Abstract

Zebrafish (*Danio rerio*) are a widely utilized model system for human disorders, but common laboratory strains have distinct behavioral and physiological differences. Accompanying these known strain differences, commonly used “wildtype” zebrafish strains have both shared and unique suites of single nucleotide polymorphisms and copy number variants (CNVs). Despite this, genomic variation is often ignored in study design, and the actual strain used is often not adequately reported. The goal of this study was to assess CNVs across three common laboratory strains of zebrafish—AB, Tubingen (TU), and WIK—and provide these data as a tool for the zebrafish community. Herein we identified 1351 CNV regions within the most recent genome assembly (GRCz11) covering 1.9% of the zebrafish genome (31.7Mb). CNVs were found across all chromosomes, and 2200 genes (5121 transcripts) lie within ± 5 kb of identified CNVs, pointing to likely *cis* regulatory actions of CNVs on nearby gene neighbors. We have created a Public Session accessible on the UCSC Genome Browser to view CNVs from this study titled “danRer11 zebrafish CNV across strains” as a tool for the zebrafish community.

Keywords: array CGH, strain variation, reciprocal comparison

Introduction

ZEBRAFISH ARE AN IMPORTANT genetic model, but the incorporation of genomic variation across common lab strains into study design has been slow. It is well established that zebrafish strains contain many shared and unique single nucleotide polymorphisms (SNPs)^{1–4} and copy number variants (CNVs)^{5,6} with several groups diligently working to describe and characterize differences between strains at the genotypic and phenotypic levels. The phenotypic effects of these genetic differences are not well understood, but likely manifest as observable variation between strains. To this end, several studies have described behavioral and physiological differences between strains such as differences in sex determination,⁷ fear-related behavior,⁸ social preference,⁹ stress,¹⁰ susceptibility to toxicants,¹¹ locomotion,¹² and growth performance.¹³

CNVs, which cover 5%–10% of the human genome,¹⁴ can directly and indirectly affect gene expression through gene dosing¹⁵ and principally act through enhancers and other regulatory elements.¹⁶ CNVs can be positively or negatively correlated with gene expression¹⁷ and result in wide-ranging phenotypic effects such as testosterone metabolism,¹⁸ the ability to digest starch,¹⁹ or complex diseases including autism

and schizophrenia (reviewed in^{20,21}). As zebrafish continue to grow as an important model system for basic research, it is imperative to expand our knowledge of the genomic variation within the species. Previous CNV identification in zebrafish focused on within-strain (intrastrain) variation. This study aimed to describe zebrafish genomic copy number variation across commonly used laboratory strains (interstrain variation) by using a reciprocal comparison design. By providing these data to the global zebrafish community, we hope to highlight the important role that copy number variation has on phenotypes across strains in support of incorporating this information into study design and publication. This is of critical importance to the zebrafish community because clearly defining genomic variation will result in better replication and translation of our research into other model systems, for human health applications, and for application to ecological systems.

Materials and Methods

Animal care and husbandry

All zebrafish husbandry and experimental procedures were performed following protocols approved by Portland State University’s Institution Animal Care and Use Committee in

¹Department of Biology, Portland State University, Portland, Oregon.

²Department of Biology, The University of Tampa, Tampa, Florida.

*Current affiliation: Health Effects Division, Toxicology Directorate, U.S. Army Public Health Center, Aberdeen Proving Ground, Maryland.

TABLE 1. SUMMARY OF COPY NUMBER COUNTS AND TYPES ACROSS FOUR VERSIONS OF THE ZEBRAFISH REFERENCE GENOME

	Zv8	Zv9	GRCz10	GRCz11
High copy gain	432	375	355	350
Copy number gain	1036	731	631	626
Copy number loss	154	128	116	112
Homozygous copy loss	319	291	265	263
Partially deleted in new	—	5	4	3
Split in new	—	411	154	13

accordance with the National Institutes of Health Guidelines for Care and Use of Laboratory Animals, and the Public Health Service Policy on Humane Care and Use of Laboratory Animals. Zebrafish were housed on an Aquaneering semi-recirculating housing system at a density of five individuals per liter with 10% daily water changes. Water temperature was maintained at 27.5°C, and fish were kept on a 16-h light, 8-h dark photoperiod. pH and conductivity were maintained at approximately 7.4 and 1100 μ S, respectively. Zebrafish were fed commercial flake food twice daily and supplemented with live artemia and rotifers. This study used three strains of zebrafish: AB, Tuebingen (TU), and WIK.

aCGH data analysis

Microarray data were obtained from the National Center for Biotechnology Information (NCBI) BioProject portal (GEO sample IDs: GSM839719, GSM839720, GSM839721).⁵ Array data comprised of pooled DNA from 10 individuals each of AB, TU, and WIK strains run in a reciprocal design (AB vs. TU, WIK vs. TU, AB vs. WIK) on custom-designed Agilent Technologies SurePrint GS CGH microarrays. Arrays were designed against the zebrafish Zv8/danRer6 reference genome and had an average probe spacing of 1.4 kb. CNVs were called using normalized signal intensity files within Nexus Copy Number software (version 5.1; BioDiscovery) and reported as log2 ratios. Log2 ratios with a three-probe running average >2 were identified as CNVs.⁵ Microarray probe chromosomal locations were updated to the most recent zebrafish reference genome using the LiftOver tool in UCSC Genome Browser²² (Zv8 $\rightarrow$ Zv9 $\rightarrow$ GRCz10 $\rightarrow$ GRCz11).

Quantitative PCR and standard PCR

Quantitative PCR (qPCR) was performed across 10 regions identified as having a CNV using primers designed on Primer3Plus software.²³ DNA was isolated through standard phenol:chloroform extraction for 10 individuals from each strain and assayed across the 10 regions in triplicate on a 364-well plate format using Power SYBR Green Master Mix (Applied Biosystems). Fluorescence was measured on an Applied Biosystems 7900HT Real-Time PCR System. The qPCR cycling protocol included preliminary dissociation (10 min at 95°C) and 35 cycles of annealing and extension (95°C for 15 s, 60°C for 30 s) per manufacturer's protocol. A dissociation melt curve was also obtained to confirm single PCR products. Quantification of PCR product was performed using the $\Delta\Delta C_t$ method²⁴ with an ultra-conserved element

(UCE) as a standardized DNA copy number reference sequence⁵ and pooled DNA from AB, TU, or WIK strains as reference for each strain. Specifically, $\Delta C_t = \text{target} - \text{UCE}$ and $\Delta\Delta C_t = \Delta C_{t,\text{individual}} - \Delta C_{t,\text{pool}}$. Pairwise comparisons were made between each strain, and fold change was calculated as the inverse log2 of $\Delta\Delta C_t$ (or $2^{-\Delta\Delta C_t}$).

Standard PCR was performed across three CNV regions confirmed by qPCR as an additional confirmation technique and further resolution of the region using Hot Start Taq 2 $\times$ Master Mix (New England BioLabs) following the manufacturer's protocol. DNA from three individuals per strain was extracted using a DNeasy Blood and Tissue kit (Qiagen). PCR products from three or four regions across each CNV were run on 1% agarose gels with Gel Red nuclear binding stain in 0.5 $\times$ TBE at 110 V for 45 min and visualized on a digital gel imager. Bands were scored as present/absent and approximate size was noted. All qPCR and standard PCR primers, amplicon sizes, and locations are listed in Supplementary Table S1.

Predicting effects of CNVs

Consequences of identified CNVs were predicted using the Ensembl Variant Effect Predictor.²⁵ Briefly, effects of CNVs were predicted for all RefSeq genes and transcripts falling within a very conservative zone of 5 kb upstream or downstream of the CNV location. CNV calls from the GRCz10 assembly were used for this analysis as this is the most recent genome version available for use with the tool. Additionally, gene ontology (GO) analysis was performed by assessing over-representation of genes that directly overlap CNV regions as compared to the genes represented by probes on the microarray platform. CNV loci and probeset loci were translated from Zv8 to GRCz11 as described above and annotated against GRCz11 RefSeq annotation database (last updated April 8, 2018) using the UCSC Genome Browser data integrator.²² Unique gene symbols for each dataset (background=probeset, target=CNVs) were entered into the Gene Ontology enRiChment anaLysis and visualiZation tool (GORilla)²⁶ and compared with two unranked lists of genes (target vs. background). Results were considered significant if they had a FDR-corrected q -value²⁷ ≤ 0.05 .

Results

Across the reciprocal comparisons of pooled DNA from three zebrafish strains, we identified 1941 CNV regions in the Zv8 genome (Supplementary Table S2). Stepwise LiftOver to GRCz11 resulted in the loss of CNV calls due to a split across the region or partial deletion in newer versions of the reference genome. The largest loss in CNV calls resulted from the Zv8 to Zv9 LiftOver because of a major genome update. Zv9 to GRCz10 and GRCv10 to GRCz11 LiftOvers also resulted in the loss of some, but fewer, CNV calls. The final count of GRCz11 CNV calls was 1351 (Table 1), and the identified CNV regions non-redundantly cover 1.9% of the zebrafish genome (31.7 Mb).

We chose 10 random CNV regions to confirm by qPCR and further interrogate the putative effects of those CNVs on the organism (Figs. 1 and 2). Of the 10 regions, 8 were fully confirmed by qPCR. Three regions showed a loss in WIK (CNV_0311, CNV_0437, CNV_0968), two regions showed a

FIG. 1. Log₂ ratios of microarray probes across chromosomal locations identified as CNVs with corresponding qPCR log₂-fold change values. CNV_0222=gain in TU; CNV_0302=disagreement between CNV (loss in AB) and qPCR (gain in TU); CNV_0311=loss in WIK; CNV_0437=loss in WIK; CNV_0559=loss in AB. Error bars represent standard deviation, and large error bars indicate variation within the strain (i.e., the loss or gain is not fully penetrant). CNV, copy number variant.

FIG. 2. Log₂ ratios of microarray probes across chromosomal locations identified as CNVs with corresponding qPCR log₂-fold change values. CNV_0572=not validated; CNV_0663=loss in TU; CNV_0900=gain in TU; CNV_0968=loss in WIK; CNV_1736=loss in AB. Error bars represent standard deviation, and large error bars indicate variation within the strain (i.e., the loss or gain is not fully penetrant).

FIG. 3. Standard PCR across subsets of three CNV regions. Each PCR amplicon was approximately 3.7 kb in length. CNV_0222 is only present in TU, while CNV_0311 and CNV_0900 show some variation across strains, but the array-based CNV call is confirmed by qPCR and PCR data across multiple individuals.

gain in TU (CNV_0222, CNV_0900), two regions showed a loss in AB (CNV_0559, CNV_1736), and one region showed a loss in TU (CNV_0663). The data for one region (CNV_0302) were in disagreement; the array data and CNV call identified a loss in AB, while the qPCR data showed a gain in TU. CNV_0572 was predicted to be a loss in WIK from array data, but qPCR failed to show any differences between strains. We further interrogated three of these regions by PCR (Fig. 3) and were able to confirm the CNVs across sub-regions approximately 3.7 kb in size for CNV_0222, CNV_0311, and CNV_0900. Not all CNVs are fully penetrant, as can be seen in CNV_0900.4 as a gain in WIK across this sub-region (in addition to the gain in TU).

CNVs were identified across all chromosomes, with the highest number of calls falling on chromosomes 3, 4, and 7 (Fig. 4) with the majority (72.2%) copy number gains. To understand the potential effects of these CNVs, we used the Ensembl Variant Effect Predictor to assess 1355 CNVs (GRCz10) against the RefSeq database and found 2200 genes and 5121 transcripts within ± 5 kb of identified CNVs (Fig. 5

and Supplementary Table S3). Of these genes and transcripts located overlapping with or proximal to CNV regions, 25% resulted in intron variants, 16% resulted in exon variants (coding sequence), and 21% impacted 3' or 5' UTRs. Furthermore, 10% of the genes and transcripts within ± 5 kb of identified CNVs are predicted to gain or lose sequence (elongation or truncation). Of the 8843 variant effects that were predicted, 2614 (29.6%) are expected to have highly disruptive impacts on the protein product, such as loss of function.

Because the variant effect predictor identified a large number of potentially disruptive changes to nucleic acid sequence, we performed a GO analysis to explore potential cellular functions that may be impacted by CNVs. The major cellular component ontology hits for all CNVs identified in AB, TU, and WIK strains were terms linked to the cell membrane (Table 2 and Supplementary Fig. S1). One hundred thirty-nine gene isoforms related to immunity, ion/salt channels, receptors, and adhesion comprise this group (Supplementary Table S4). The major molecular function ontology hits were for ADP-ribosyltransferase and dimethylaniline monooxygenase across nine gene isoforms. These two molecular functions may serve to modify gene expression epigenetically and metabolize chemicals to excretable metabolites, respectively. The biological process ontology hits are dominated by 24 adhesion and neuron-related gene isoforms.

Discussion

In an effort to understand the effects of CNVs on common lab strains, we searched for available datasets reflecting structural variation in zebrafish. The current NCBI structural variation database (dbVar; www.ncbi.nlm.nih.gov/dbvar) contains only human and great ape datasets and wholly excludes zebrafish data. The majority of available datasets incorporating CNVs and phenotypes are from humans, where much work has been done to associate structural variants to disease phenotypes. The EMBL-EBI Database of Genomic Variants archive (DGVA; www.ebi.ac.uk/dgva) contains 200 genomic structural variant studies, but only a single study utilizes zebrafish (study ID nstd62). The design of the single zebrafish study in DGVA used within-population comparisons to identify strain-specific structural variation, and then used a subtractive model to infer between strain differences.⁵

FIG. 4. Copy number counts across all chromosomes in GRCz11 by type (homozygous copy loss, copy number loss, copy number gain, or high copy gain). Color images are available online.

FIG. 5. Predicted consequences of CNVs that occur within 5 kb $\pm$ of RefSeq transcripts (GRCz10) using the Ensembl Variant Effects Predictor. Color images are available online.

Therefore, the majority of structural variation presented in that study represents within-strain variation. We also found one previous study that looked at copy number variation in the wild zebrafish strain ASWT,²⁸ but the strain is not widely used, so the data are not applicable to most zebrafish investigators. Interestingly, we found no overlap in their copy number deletion or insertion loci with our dataset, which indicates that the CNVs found in that study are highly strain-specific.

The effects of these CNVs on phenotype are not fully known, but we can predict their transcriptional and transla-

tional consequences. CNVs can exert effects through direct interaction with genic regions (i.e., CNVs directly overlap a gene coding region) or through indirect interaction (i.e., CNVs fall upstream or downstream of genic regions, often in functional regulatory regions).¹⁷ To predict the functional effects of the CNVs described in this study, we applied the Ensembl Variant Effect Predictor²⁵ to a ± 5 kb genomic window upstream or downstream of all identified CNVs under the assumption that regulatory elements found within 5 kb of a gene are likely effectors of gene regulation. We found 2200 genes that comprise 5121 transcripts that fall

TABLE 2. GENE ONTOLOGY CLASSIFICATIONS OF GENES THAT DIRECTLY OVERLAP WITH COPY NUMBER VARIANTS

	Description	p-value	FDR q-value	Enrichment	Gene count
Component	Integral component of plasma membrane	1.86E-08	2.14E-05	2.78	37
	Intrinsic component of plasma membrane	7.23E-08	4.15E-05	2.64	37
	Integral component of membrane	4.38E-06	1.68E-03	1.4	134
	Intrinsic component of membrane	6.57E-06	1.89E-03	1.39	134
	Cell surface	8.08E-06	1.86E-03	6.57	9
	Membrane part	1.85E-04	3.55E-02	1.29	138
	Plasma membrane part	3.55E-04	5.83E-02	1.74	41
Function	NAD+ ADP-ribosyltransferase activity	5.00E-06	1.39E-02	12.73	6
	N,N-dimethylaniline monooxygenase activity	4.41E-05	6.12E-02	33.41	3
Process	Homophilic cell adhesion via plasma membrane adhesion molecules	5.62E-10	4.29E-06	6.7	17
	Cell-cell adhesion via plasma-membrane adhesion molecules	1.46E-09	5.58E-06	6.31	17
	Cell-cell adhesion	8.89E-08	2.26E-04	4.82	17
	Biological adhesion	1.07E-05	2.04E-02	2.94	21
	Cell adhesion	1.07E-05	1.63E-02	2.94	21
	Nervous system development	7.15E-05	9.10E-02	3.39	14
	Cell-cell signaling	7.54E-04	8.22E-01	2.82	13

Classifications fall into three main categories: cellular component, molecular function, or biological process. Each statistically significant GO annotation (q -value < 0.05) is indicated in **bold**. Enrichment indicates the proportion of genes in the target list (CNVs) annotated to the descriptive GO term versus the proportion of genes in the background list (probeset) annotated to the same term. For a complete list of all GO IDs and identified genes, see Supplementary Table S4. GO, gene ontology.

FIG. 6. UCSC Genome Browser view of chromosome 4 (GRCz11) with CNV locations in teal; RefSeq genes in blue; and GenBank mRNAs in black. Freely available at https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=holdenl&hgS_otherUserSessionName=danRer11%20zebrafish%20CNV%20across%20strains. Color images are available online.

within ± 5 kb of identified CNVs. CNVs that directly overlap or are nearby neighbors of these genes and their transcripts are likely to have differential gene expression relative to their copy number status. Because of the design of this study (interstrain variation), CNVs identified in this study differ across commonly used laboratory strains of zebrafish. GO analysis suggests that these CNVs likely impact cellular adhesion, neuronal development, epigenetic processes, metabolism, immune function, and ion transfer, specifically across cell membranes. Therefore, it is imperative that we use CNV status as a modifier when comparing results from studies performed on different strains or populations of zebrafish.

This study performed reciprocal comparisons between three common laboratory strains (AB, TU, and WIK) and found 1351 CNV regions covering 1.9% (31.7 Mb) of the current zebrafish genome assembly (GRCz11). Our analysis used pooled sample arrays ($n = 10$ individuals per strain), not individuals, which likely reduces the level of variation detected as only the most common variants are likely to be identified, but provide strong evidence for CNVs across strains. These CNV loci are available to the zebrafish community for further exploration as Supplementary Table S2. To facilitate utilization of this dataset, we created a Public Session accessible on the UCSC Genome Browser for the zebrafish community titled “danRer11 zebrafish CNV across strains” (as an example, see Fig. 6). This study highlights the potential impact that CNVs have on phenotype, and we argue that incorporating this type of genomic variation is imperative in today’s research environment where replication of results is often difficult. Incorporating genomic variation, such as SNPs² and CNVs, will allow the zebrafish community to more efficiently translate findings from zebrafish to human health, ecological systems, and basic genomic variation research.

Acknowledgments

This study was funded by NIEHS K99/R00 ES018892 and start-up funds awarded to K.H.B. Additional support was provided by the Science, Mathematics, and Research for Transformation (SMART) scholarship, a National Defense Education Program, to L.A.H.

Disclosure Statement

No competing financial interests exist.

References

- Butler MG, Iben JR, Marsden KC, Epstein JA, Granato M, Weinstein BM. SNPfisher: tools for probing genetic variation in laboratory-reared zebrafish. *Development* 2015; 142:1–11.
- Balik-Meisner M, Truong L, Scholl EH, Tanguay RL, Reif DM. Population genetic diversity in zebrafish lines. *Mamm Genome* 2018;0:1–11.
- Guryev V, Koudijs MJ, Berezikov E, Johnson SL, Plasterk RHA, van Eeden FJM, *et al.* Genetic variation in the zebrafish. *Genome Res* 2006;16:491–497.
- Coe TS, Hamilton PB, Griffiths a M, Hodgson DJ, Wahab M a, Tyler CR. Genetic variation in strains of zebrafish (*Danio rerio*) and the implications for ecotoxicology studies. *Ecotoxicology* 2009;18:144–150.
- Brown KH, Dobrinski KP, Lee AS, Gokcumen O, Mills RE, Shi X, *et al.* Extensive genetic diversity and substructuring among zebrafish strains revealed through copy number variant analysis. *Proc Natl Acad Sci U S A* 2012;109:529–534.
- Faber-Hammond JJ, Brown KH. Pseudo-*de novo* assembly and analysis of unmapped genome sequence reads in wild zebrafish reveal novel gene content. *Zebrafish* 2016;13: 95–102.
- Wilson CA, High SKS, McCluskey BMB, Amores A, Yan Y, Titus TA, *et al.* Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* 2014; 198:1291–1308.
- Drew RE, Settles ML, Churchill EJ, Williams SM, Balli S, Robison BD. Brain transcriptome variation among behaviorally distinct strains of zebrafish (*Danio rerio*). *BMC Genomics* 2012;13:323.
- Scerbina T, Chatterjee D, Gerlai R. Dopamine receptor antagonism disrupts social preference in zebrafish: a strain comparison study. *Amino Acids* 2012;43:2059–2072.
- Wong RY, Perrin F, Oxendine SE, Kezios ZD, Sawyer S, Zhou L, *et al.* Comparing behavioral responses across multiple assays of stress and anxiety in zebrafish (*Danio rerio*). *Behaviour* 2012;149:1205–1240.
- Waits ER, Nebert DW. Genetic architecture of susceptibility to PCB126-induced developmental cardiotoxicity in zebrafish. *Toxicol Sci* 2011;122:466–475.
- Lange M, Neuzeret F, Fabreges B, Froc C, Bedu S, Bally-Cuif L, *et al.* Inter-individual and inter-strain variations in zebrafish locomotor ontogeny. *PLoS One* 2013;8: e70172.
- Meyer BM, Froehlich JM, Galt NJ, Biga PR. Inbred strains of zebrafish exhibit variation in growth performance and myostatin expression following fasting. *Comp Biochem Physiol A Mol Integr Physiol* 2013;164:1–9.
- Zarrei M, MacDonald JR, Merico D, Scherer SW. A copy number variation map of the human genome. *Nat Rev Genet* 2015;16:172–183.
- Handsaker RE, Van Doren V, Berman JR, Genovese G, Kashin S, Boettger LM, *et al.* Large multiallelic copy number variations in humans. *Nat Genet* 2015;47:296–303.

16. Chiang C, Scott AJ, Davis JR, Tsang EK, Li X, Kim Y, *et al.* The impact of structural variation on human gene expression. *Nat Genet* 2017;49:692–699.
17. Stranger B, Forrest M, Dunning M. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* (80-) 2007;315:848–853.
18. Schulze JJ, Lundmark J, Garle M, Skilving I, Ekström L, Rane A. Doping test results dependent on genotype of uridine diphospho-glucuronosyl transferase 2B17, the major enzyme for testosterone glucuronidation. *J Clin Endocrinol Metab* 2008;93:2500–2506.
19. Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, *et al.* Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 2007;39:1256–1260.
20. Buchanan JA, Scherer SW. Contemplating effects of genomic structural variation. *Genet Med* 2008;10:639–647.
21. Wain L V., Armour J AL, Tobin MD. Genomic copy number variation, human health, and disease. *Lancet* 2009;374:340–350.
22. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, *et al.* The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res* 2014;42:D764–D770.
23. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, *et al.* Primer3-new capabilities and interfaces. *Nucleic Acids Res* 2012;40:1–12.
24. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25:402–408.
25. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 2016;17:1–14.
26. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 2009; 10:1–7.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995;57:289–300.
28. Patowary A, Purkanti R, Singh M, Chauhan R, Singh AR, Swarnkar M, *et al.* A Sequence-Based Variation Map of Zebrafish. *Zebrafish* 2013;10:15–20.

Address correspondence to:

Kim H. Brown, PhD
Department of Biology
Portland State University
1719 SW 10th Avenue
Portland, OR 97201

E-mail: kibr2@pdx.edu