Deep Phenotyping Enhances Somatic Mutant Detection and Mosaic Pattern Recognition in the Zebrafish Skeleton

Running Title: Mosaic Detection Using Skeletal Phenomics

Claire J. Watson[1,2,7], Adrian T. Monstad-Rios[1,2], Rehaan M. Bhimani[1,2], Charlotte Gistelinck[1,4], Andy Willaert[4], Paul Coucke[4], Yi-Hsiang Hsu[5,6], Ronald Y. Kwon[1,2,3]

[1] Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, Washington, USA

[2] Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, Washington, USA

[3] Department of Mechanical Engineering, University of Washington, Seattle, Washington, USA

[4] Center for Medical Genetics Ghent, Ghent University, Ghent, Belgium

[5] Hebrew SeniorLife Institute for Aging Research, Boston, Massachusetts, USA

[6] Harvard Medical School, Boston, Massachusetts, USA

[7] Corresponding author: cwalk1@uw.edu

Key Words: CRISPR, crispant, zebrafish, bone, osteoblast, phenomics, Osteogenesis Imperfecta, brittle bone disease, osteoporosis, *bmp1a*, *plod2*, *wnt16*

ABSTRACT

Because CRISPR-based rapid-throughput biology often entails analysis of G0 somatic mutants, there is a need to improve our ability to discern and interpret variable phenotypes arising from mosaicism. Here, we examine mechanisms underlying mosaic patterns in "crispant" zebrafish, and assess the benefits of phenomics--i.e., in depth phenotyping at a large number of sites--in discerning variable phenotypes arising from mosaicism. We characterize spatial and size characteristics of loss-of-function cell clusters in the skeleton. Despite inter-animal differences in mutation efficiency, we identify a distinctive size distribution in bones of separate developmental lineages, replicating that predicted from clonal fragmentation and merger events. Using a microCT-based workflow enabling hundreds of measures per animal, we describe diverse phenotypic manifestations in mosaic models of brittle bone diseases (*plod2* and *bmp1a*). At the population level, we show convergence between somatic and germline mutant phenomes. We demonstrate that vertebral patterns confer heightened sensitivity in discriminating somatic mutant populations compared to single readouts. We also establish a framework to quantify spatial phenotypic variation, and show this is differentially altered in disease models. Finally, we detect a low bone mass phenotype in adult somatic mutants for *wnt16*, a gene implicated in heritable risk for osteoporosis, and validate this phenotype in germline mutants. Our studies demonstrate mosaic analysis and phenomic approaches to help detect and interpret spatially variable phenotypes in G0 animals. Applied to the skeleton, microCT-based phenomics, in concert with CRISPR-based G0 screens, is a promising direction for the identification of causal genes and targeted therapies for mono- and multigenic skeletal diseases.

Keywords: CRISPR, crispant, zebrafish, bone, osteoblast, phenomics, screen, microCT, Osteogenesis Imperfecta, brittle bone disease, osteoporosis, *bmp1a*, *plod2*, *wnt16*

INTRODUCTION

When complexed with a targeting guide RNA (gRNA), the bacterial Cas9 enzyme will create a double strand break at a genome-specific location determined by complementary sequence in the gRNA. Errors in the endogenous non-homologous end joining (NHEJ) repair mechanism lead to a high rate of insertions and deletions (indels) at the cut site, often leading to frameshift mutations and loss of gene product function. While the advent of CRISPR-based genome editing technologies have quickly changed the accessibility and ease by which the scientific community can approach rapid-throughput genetics using model organisms (1, 2), parallel advances in phenotyping are needed to increase the throughput and fidelity of such approaches.

A prime instance of a rapid-throughput biology which may benefit from advances in phenotyping is reverse genetic screening in vertebrates. Recently, several prototypes for rapid CRISPR-based reverse genetic screens have been developed in which phenotyping is performed in first generation (G0) somatic mutants (1, 2). G0 screens enable multiplexed approaches for rapid-throughput gene knockdown, the study of epistatic interactions of genes that are tightly linked, and knockdown of clusters of genes with functional redundancy (1, 3, 4). However, even with high mutation efficiencies, creating universal loss-of-function is challenging. For instance, when administered a single gRNA, 1/3 of indels are expected to be in-frame. Thus, less than half ($[2/3]^2 = 4/9$) of cells are expected to have bi-allelic out-of-frame mutations (1). While the use of multiple guides to redundantly target the same gene can increase the proportion of bi-allelic out-of-frame mutations, this also increases toxicity, and variable penetrance of null phenotypes is still prevalent (2). Taking advantage of microhomology-mediated end joining (MMEJ) is a promising direction to enrich somatic mutations for a predictable out-of-frame allele (5). However, imperfect editing efficiencies and some degree of mutated allelic mosaicism are still expected following MMEJ. Due to the lower fidelity in detecting somatic mutant phenotypes, prototypical screens have mostly focused on severely dysmorphic phenotypes (2), readily observable traits such as cutaneous phenotypes (2), or phenotypes whose spatial variations are easily observable (1). Since CRISPR-based rapid-throughput biology often entails analysis of G0 somatic mutants, there is a need to improve our ability to discern and interpret variable phenotypes arising from genetic mosaicism.

Because a phenotypic hallmark of mosaicism is site-to-site variability, a primary challenge to somatic mutant analysis is the need to phenotypically sample a large

number of anatomical sites. Phenotypic expression of mosaicism can be influenced by a number of factors, including timing of mutagenesis, the proliferation of mutant cells and their translocation to different sites, the function of the mutated genes in influencing phenotype (e.g., whether the gene acts cell autonomously or non-cell autonomously), as well as any pathological effects of mutations in the gene itself (6, 7). In humans, mosaicism can manifest as relatively uniform phenotypes reminiscent of a generalized condition, or alternating patterns of affected and unaffected body segments (8). Due to the challenges in phenotypic sampling at a large number of sites, much of our knowledge of the phenotypic consequences of mosaicism has been derived from easily observable traits where spatial variations are readily discernible (9). As such, our ability to discern phenotypic manifestations of mosaicism for complex traits remains relatively limited (9). Further, because most of our knowledge of the phenotypic consequences of mosaicism has been derived from qualitative observations of easily observable traits, our understanding of quantitative phenotypic variation arising from mosaicism is limited. Such a quantitative understanding is important to guide the development of analytical approaches which may serve to increase the fidelity of somatic mutant analysis in different tissues. Finally, the lack of systematic frameworks to analyze phenotypic variation arising from mosaicism hampers our understanding of biological factors that influence phenotypic expressivity in mosaic individuals. Such knowledge is important to better understand clinical manifestations of mosaicism in human diseases (6, 7).

Here, we examine mechanisms underlying mosaic pattern development in the skeletons of G0, "crispant", zebrafish, and use a microCT-based workflow enabling profiling of hundreds of measures per fish (10) to understand its benefits in discerning and interpreting phenotypic manifestations in mosaic bone disease models.

RESULTS

CRISPR-based gene editing results in clusters of cells with loss-of-function

To examine patterns of loss-of-function in the skeleton, *sp7:EGFP* (11) embryos were injected with Cas9:guide RNA (gRNA) ribonucleoprotein complexes (RNPs) targeting the fluorescent transgene. This enabled loss-of-function mutations in EGFP to be visualized as loss-of-fluorescence in *sp7*+ (*osterix*+) osteoblasts. Fish injected with 10μM RNPs were examined for functional EGFP loss at 10-12dpf, a stage when the larvae are still transparent and many skeletal elements are present. Regions of loss-of-fluorescence were observed in virtually all formed skeletal elements (Fig 1A). This included bones of the craniofacial skeleton (opercle, interopercle, subopercle, dentary,

pharyngeal arches, branchiostegal rays, cleithrum), median fin rays (ventral, caudal), hypurals, and the spine (neural and haemal arches, centra, Weberian apparatus). While penetrance of loss-of-fluorescence was high, expressivity was variable in regard to the composition of bony elements in each animal that exhibited regions of fluorescence loss, as well as the size and number of such regions within each bony element. Notably, some bony elements completely lost fluorescence, suggesting multi-allelic, functional knockdown in these structures.

Loss-of-fluorescence regions are composed of clusters of cells with loss-of-function mutations: cells comprising each cluster represent a single clone, or multiple clones that merged at an earlier point in development (we are unable to distinguish these two possibilities). Within individual vertebrae, we often observed multiple, contiguous centra with complete or partial loss-of-fluorescence, and which were flanked by at least one centra with no loss-of-fluorescence. This resulted in two distinguishable types of cell clusters: "microscale" clusters confined within single vertebrae, and "macroscale" clusters spanning contiguous vertebrae (Fig 1B). Inspection at higher magnification revealed that some centra exhibited loss-of-fluorescence in ventral, but not dorsal, regions (or vice versa). This dorso-ventral stratification could at times be seen across contiguous centra (Fig 1C), potentially due to these bodies' shared clonal partners. In regard to the neural arches, loss-of-fluorescence often appeared to be associated with loss-of-fluorescence in the centrum of the same vertebral body (Fig 1D). Because many bones retained partial or complete expression of the transgene, this suggested that, on the whole, individual bony elements are not explicitly derived from single clonal populations, and cannot be evaluated as independent functional or non-functional units.

A common distribution underlies the sizes of loss-of-function clusters in bones of distinct developmental lineages, and in animals with different mutation efficiencies

While some aspects of loss-of-fluorescence patterns appeared to be non-random, patterns from fish-to-fish were unpredictable, suggesting stochastic forces were an important etiological factor. Models of clonal population dynamics in fluorescence-based cell lineage tracing studies have demonstrated that while different factors can contribute to cluster size distributions during tissue growth, over time, contributions from random clonal merger and fragmentation (Fig 2A) become dominant over those from cell behaviors specified by developmental programs (e.g., cell division or loss)(12). As a consequence, cluster size distributions across diverse developmental processes often

exhibit the same characteristic distribution once cluster sizes in each individual are normalized by the average cluster size in that individual (12). This distribution has the form:

$$y=\exp(-x/\langle x\rangle), \qquad \text{(Eq 1)}$$

where x is cluster size, and $\langle x\rangle$ is the mean of x (12).

Fluorescently-labeled cell clusters in cell lineage tracing studies and loss-of-fluorescence cell clusters in CRISPR-edited mutants share commonalities in their physical origins (postzygotic mutations) and interpretation (clusters may be comprised of a single clonal population, or multiple clones that merged earlier in development). As such, we hypothesized that loss-of-function clusters in G0 animals would also exhibit universality in their size distributions described by Eq 1. To test this, we manually traced regions of loss-of-fluorescence on the dorsal and ventral aspects in the centra in each animal (Fig 2B). Regions confined to a single vertebrae were annotated as "microscale clusters"; regions spanning contiguous vertebrae were annotated as "macroscale clusters". Individual fish exhibited variable numbers and sizes of loss-of-function clusters (Fig 2C), as well as different compositions of microscale versus macroscale clusters. This variability manifested as distinct distributions of loss-of-function cluster sizes in each fish (Fig 2D). However, when normalized by average cluster size in each animal, the cluster size distributions in each animal collapsed to the distribution in Eq 1 for both microscale and macroscale clusters (Fig 2E). We hypothesized that this distribution would also describe loss-of-function cluster sizes in bones of a different developmental lineage. We quantified loss-of-function cluster size distributions within the branchiostegal rays of the craniofacial skeleton, which unlike the somite-derived vertebral column, derives from neural crest (13). Consistent with our hypothesis, a similar data collapse was observed (Fig 2F). These studies demonstrate that loss-of-function cluster sizes in CRISPR-edited G0 animals in bones of distinct developmental lineages, and in animals with different mutation efficiencies, can be described by a single distribution (Eq 1), the origin of which derives from numerical convergence behaviors associated with clonal fragmentation and merger events.

Phenomic profiling in somatic mutants for plod2 *and* bmp1a

To understand how mosaicism is phenotypically manifested in the skeleton, we generated somatic mutants for *plod2* and *bmp1a*. In humans, mutations in *PLOD2* and

BMP1 are associated with Osteogenesis Imperfecta (OI), a heritable, heterogeneous group of connective tissue disorders commonly associated with skeletal deformities and bone fragility. OI is predominantly caused by dominant mutations in *COL1A1* or *COL1A2*, whereas mutations in non-collagenous genes mostly cause autosomal recessive forms of OI (14); the latter genes include *PLOD2* and *BMP1* (15). The enzyme encoded by *PLOD2*, lysyl hydroxylase 2, localizes to the endoplasmic reticulum, and catalyzes lysine residue hydroxylation in fibrillar collagen telopeptides (16). BMP1 is a secreted enzyme that functions in the cleavage of C-propeptides from procollagen precursors (17). Somatic mutants for *plod2* and *bmp1a* were generated by injection of RNP complexes into embryos, and a subset of larvae were individually screened for indels at 12dpf. Sanger sequencing and TIDE analysis (18) revealed intra-animal mutation efficiencies of 82.7-88.1% and 71.0-87.5% for *plod2* and *bmp1a*, respectively. Adult phenotypes were visible at 90dpf, adding to several recent reports (2, 7) examining the durability of crispant zebrafish phenotypes through the larval-to-adult transition.

We performed microCT-based phenomics in 90dpf adult animals (10). Recently, we developed a microCT-based workflow, FishCuT, which enables rapid (<5min/fish) quantification of 100s of measures in the axial skeleton of adult zebrafish (10, 14, 19). In this workflow, 25 different quantities are computed for each vertebra (Cent.TMD, Cent.Th, Cent.Vol, Cent.Le, Cent.SA, Cent.TMD.sd, Cent.Th.sd, Neur.TMD, Neur.Th, Neur.Vol, Neur.SA, Neur.TMD.sd, Neur.Th.sd, Haem.TMD, Haem.Th, Haem.Vol, Haem.SA, Haem.TMD.sd, Haem.Th.sd, Tot.TMD, Tot. Th, Tot.Vol, Tot.SA, Tot.TMD.sd, Tot.Th.sd; see (10) for description). Once calculated, these quantities are plotted as a function of vertebra number along the axial skeleton for each fish; we have termed such entities 'vertebral traces'. Next, the global test (20), a regression-based statistical test designed for data sets in which many features have been measured for the same subjects, is used to assess whether the collection of vertebral traces are significantly different between groups. For each combination of outcome/element, a standard score is computed as the difference between its value in each vertebral body and its mean value across all vertebrae in the control population, divided by the standard deviation across all vertebrae in the control population. These data are arranged into matrix constructs that we have termed ‘skeletal barcodes’. In our studies, 21 vertebrae (21*25=525 measures/animal) in 52 animals were analyzed, resulting in 52*525=27,300 data points, thus, representing high-dimensional phenotypic data across the majority of the axial skeleton for each fish. To facilitate comparisons with prior studies (10), we present data on ten combinatorial quantities (the nine possible combinations of (Cent, HA, NA) x (Vol,

TMD, and Th), plus Cent.Le) in the 16 anterior-most vertebrae in the main text, and have included all 25 quantities in the supplemental material.

Global tests of vertebral traces revealed that *plod2* somatic mutants (n=11 fish/group) exhibited significant differences in centrum, haemal arch, and neural arch tissue mineral density (Cent.TMD: p=0.000005, Haem.TMD: p=0.00008, Neur.TMD: p=0.00005); centrum volume (p=0.004), thickness (p=0.04), and length (p=0.00009); and neural arch thickness (p=0.00007) (Fig 3). Somatic mutants for *bmp1a* (n=15 fish/group) exhibited significant differences in centrum, haemal arch, and neural arch tissue mineral density (Cent.TMD: p=0.000004, Haem.TMD: p=0.00002, Neur.TMD: p=0.00005); centrum volume (p=0.004), thickness (p=0.02), and length (p=0.01); and haemal arch thickness (p=0.008) (Fig 4). Data for all 25 combinatorial measures for *plod2* and *bmp1a* are provided in Fig S1 and Fig S2, respectively. For both *plod2* and *bmp1a*, most significantly different features were associated with vertebral traces that were elevated or depressed across all vertebrae; an exception was neural arch thickness in *plod2* somatic mutants, which was lower in anterior vertebrae, but higher in posterior vertebrae. Standard length (S.L.) was significantly reduced compared to sham controls for both *plod2* (control: 22.8mm [21.1-24.3mm], mutant: 19.8mm [17.0-22.6mm]; p=0.002; n=11/group; median [range]) and *bmp1a* (control: 23.8mm [21.8-24.4mm], mutant: 22.7mm [20.8-25.3mm]; p=0.009; n=15/group).

Phenomic profiling provides a unique opportunity to study "outlier" phenotypes--e.g., those arising from variability in allelic spectrum--since the acquisition of a large number of phenotypic measures provides more opportunities to observe rare events or those associated with distribution tails. We characterized incidence of mild and extreme outlier phenotypic measures, which we defined as those having an absolute barcode value greater than 1.5 and 3.0 times the interquartile range (IQR) from the median, respectively (21). Across 1,760 and 2,400 measures in *plod2* and *bmp1a* somatic mutants, the proportion of mild outliers was 1.0% and 1.1%, respectively. Most of mild outliers could be explained by normal variation, as in a normal distribution, 0.74% of observations reside beyond 1.5*IQR from the median. We observed no incidence of extreme outlier values in either somatic mutant group.

Somatic mutants for plod2 and bmp1a *exhibit different degrees of spatial phenotypic variability*

We further characterized *plod2* and *bmp1a* somatic mutant phenotypes to better understand the extent of their phenotypic variability, as well as their resemblance to germline mutant counterparts. We and others previously showed that zebrafish germline loss-of-function mutants for *plod2* and *bmp1a* exhibit severe skeletal abnormalities as adults, reminiscent of OI phenotypes (10, 16, 22, 23). In the present studies, somatic mutants for both genes exhibited clear skeletal abnormalities similar to their adult germline mutant counterparts (Fig 5A,B). Somatic mutants for *plod2* exhibited severe vertebral malformations including compression of the vertebrae along the anteroposterior axis, kyphosis, and increased bone formation. Somatic mutants for *bmp1a* exhibited increased vertebral radiopacity and bone thickening.

Inter-animal variability in phenotypic expressivity was clearly evident. For *plod2*, such variability was perceptible by the variable number of dysmorphic vertebrae in each animal; in the 24 anterior-most precaudal and caudal vertebrae, *plod2* somatic mutants exhibited 12 [6-20] (median [range]) obviously thick or malformed vertebrae per fish. Further, 100% (11 out of 11) of animals were penetrant, as all individuals exhibited at least one severely malformed vertebra per animal. For *bmp1a* somatic mutants, most animals exhibited a qualitative increase in vertebral radiopacity and bone thickening. The extent of this increase was variable among individuals, ranging from mild to severe.

Spatial variability in phenotypic expressivity within each animal was also evident. More specifically, in some instances, *plod2* somatic mutants exhibited "patchy" expressivity characterized by contiguous spans of one or more dysmorphic vertebrae surrounded by vertebrae that appeared qualitatively normal. This pattern was reminiscent of spatial characteristics of macroscopic loss-of-function clusters in *sp7:EGFP* fish previously characterized. In *plod2* germline mutants, vertebrae are uniformly dysmorphic (16), suggesting that such "patchy" expressivity is not an inherent property of *plod2* loss of function in adult animals. In contrast to *plod2*, for *bmp1a* somatic mutants, intra-animal variability in phenotypic expressivity was less obvious: while radiopacity and thickening was variable from fish-to-fish, within each animal, these characteristics appeared to be relatively uniform. Spatial variability in phenotypic expressivity was also noticeable in vertebral traces of FishCuT measures; individual vertebral traces for Cent.Vol in *plod2* somatic mutants (Fig 5C) appeared to be jagged due to local variability in phenotypic expressivity. In contrast, individual vertebral traces for Cent.Vol in *bmp1a* somatic mutants were relatively smooth (Fig 5D). Taken together,

somatic mutants for *plod2* and *bmp1a* exhibit different degrees of intra-animal variability in phenotypic expressivity.

Somatic and germline mutant phenomes exhibit high correspondence

We assessed the extent to which somatic mutants for *plod2* and *bmp1a* could function as a faithful representation of phenotypes in their germline mutant counterparts. More specifically, we hypothesized that when averaged across a sufficient sample, phenotypic values would resemble those in counterpart germline mutants. For these analyses, we used measurements and FishCuT outputs for *plod2* and *bmp1a* germline mutants (n=3 for both germline groups) previously generated in (10). Phenome-wide correlations in somatic and germline mutants and the dependence of these correlations on sample size were examined. Average barcode values were computed for each somatic and germline mutant population, and linear regressions were used to compute R^2 values. Quantification of R^2 as a function of somatic mutant sample size revealed clear asymptotic behavior in R^2 values with increased sample size (Fig 5E). For small somatic mutant sample sizes, correlations were higher for *bmp1a* compared to *plod2*, which we attribute to the lower levels of spatial phenotypic variation in *bmp1a*.

Next, we examined how differences in body size effect phenotypic correlations. We previously showed that by transforming WT sibling data to a 'virtual' phenome scaled to the mean standard length of age-matched mutants, allometric models provided a means to enable length-matched fish comparisons from an age-matched control group (10). Because there were differences in body size between somatic and germline mutant counterparts (Fig S3), we performed analyses both in the presence and absence of allometric normalization. R^2 values were computed using all fish in our samples ($n_{somatic}$=11 for *plod2*, and $n_{somatic}$=15 for *bmp1a*, $n_{germline}$=3 for both *plod2* and *bmp1a*). For *plod2*, the correlation between somatic and germline mutant barcode values was R^2=0.67 ($p<0.0001$) (Fig 5F,G); this increased to R^2=0.78 ($p<0.0001$) (Fig 5F,H) when both sets of mutants were subjected to allometric normalization. For *bmp1a*, the correlation between somatic and germline mutant barcode values was R^2=0.61 ($p<0.0001$). Correlations were not stronger (R^2=0.52) ($p<0.0001$) when both sets of *bmp1a* mutants were subjected to allometric normalization (Fig S4).

As a final method to compare somatic and germline mutant phenomes, we examined correspondence in FishCuT vertebral traces and global test results. For *bmp1a* somatic mutants, across the 10 primary measures, 6 out of 7 (86%) of the

measures exhibiting statistical significance in germline mutants were also significant in somatic mutants. Further, 2 out of 3 (and 67%) of the measures not exhibiting statistical significance in germline mutants were also not significant in somatic mutants. Across all 25 measures in FishCuT, *bmp1a* somatic mutants exhibited significant differences for 83.3% (15 out of 18) of the measures significantly different in *bmp1a* germline mutants. Further, 57.1% (4 out of 7) of the measures not significantly different in *bmp1a* germline mutants were also not different in *bmp1a* somatic mutants. There was similar correspondence in vertebral traces for somatic and germline mutants for *plod2*. The *plod2* somatic mutants exhibited significant differences for 80% (4 out of 5) of the combinatorial measures significantly altered in germline mutants subjected to allometric normalization. Moreover, 60% (3 out of 5) of the combinatorial measures not significantly different in *plod2* germline mutants were also not different in somatic mutants. Comparisons of results to the other 15 measures in FishCuT were not possible because allometric models have yet to be developed for them. Taken together, our studies suggest that somatic mutant phenomes can exhibit high correspondence with germline mutant counterparts when averaged over a sufficient sample.

Vertebral patterns enhance sensitivity in discriminating somatic mutant populations

We sought to better understand the virtues of vertebral patterns in discerning and interpreting somatic mutant phenotypes, as a general model of potential benefits of phenomic approaches to somatic mutant biology. For these studies, we performed Monte Carlo simulations. Multivariate distributions simulating mutant phenotypic distributions were constructed using Eq 1 (See SI Appendix). These distributions were used to generate samples of simulated WT and mutant phenotypic profiles (1,000 simulations per analysis), which enabled performance evaluations for different testing procedures.

We tested the hypothesis that assessing vertebral patterns with the global test would provide greater sensitivity in distinguishing somatic mutants with variable phenotypic expressivity compared to (a) Mann-Whitney (M.-W.) of individual vertebrae, and (b) M.-W. tests of quantities averaged across all vertebrae. We chose the M.-W. test as a reference univariate test because, like the global test, the M.-W. test is non-parametric. We initially analyzed Cent.Vol, a measure that exhibited high intra-animal variability in *plod2* somatic mutants. For microscopic loss-of-function clusters (Fig 6A), analyzing vertebrae 1:16 using the global test had higher sensitivity (fraction of times in which $p<0.05$ when comparing simulated mutant fish to WT fish) than using the M.-W.

test using vertebra 2, as well as the M.-W. test using quantities averaged across vertebrae 1:16 (Fig 6B). Relative differences between testing procedures increased with decreasing values of λ (extent of intra-animal variation in phenotypic expressivity; λ=0 is most variable, λ=1 is least variable). For λ=0, analyzing vertebrae 1:16 with the global test conferred a 1.41- and 1.15-fold increase in sensitivity compared to analyzing vertebra 2 and the mean of vertebrae 1:16 with the M.-W. test, respectively. Similar trends were observed for other combinatorial measures (Fig 6C). For macroscale loss-of-function clusters (Fig 6D,E,F), differences between testing procedures were exacerbated. For instance, for λ=0, analyzing Cent.Vol in vertebrae 1:16 with the global test conferred a 3.65- and 1.21-fold increase in sensitivity compared to analyzing vertebra 2 and the mean of vertebrae 1:16 with the M.-W. test, respectively. In these simulations, specificity (1 - fraction of times in which $p<0.05$ when comparing WT to WT fish) ranged between 0.94-0.97 (Fig 6G), close to the expected value of 0.95. Thus, vertebral patterns conferred increased sensitivity in discriminating somatic mutant populations, with similar levels of specificity, compared to single readouts.

We corroborated the above findings by performing non-parametric simulations using experimental data derived from *plod2* somatic mutants (Fig 6H,I). Specifically, the n=11 *plod2* somatic mutants (and sibling controls) from the previous section were used to perform bootstrap simulations. In these, control and somatic mutant phenotypic profiles were randomly sampled (with replacement) to form groups of biological replicates. Consistent with our analyses in parameterized simulations, the global test conferred higher sensitivity, with similar specificity, compared to M.-W. tests using averaged quantities as well as individual vertebrae. Specifically, for Cent.Vol (assuming n=10), the sensitivity of the global test, M.-W. test using a single vertebra, and M.-W. test using averaged quantities was 0.94, 0.56, and 0.88. Specificity was 0.96 in all cases (Fig 6H). Similar trends were observed for other phenotypic measures (Fig 6I). Thus, independent simulation approaches suggest the benefits of analyzing vertebral patterns in discriminating somatic mutant populations, compared to single readouts.

Spatial statistics can help decode relationships between mosaic patterns and spatial phenotypic variation

While spatial phenotypic variation in mosaic individuals encodes biological information, quantitative approaches to decode this information within phenotypic profiles are relatively lacking. Common measures of statistical dispersion (e.g., standard

deviation, coefficient of variation, etc.) do not account for spatial characteristics (e.g., whether vertebra with high phenotypic severity tend to cluster together), and thus do not facilitate the relation of phenotypic patterns to underlying mosaicism. We tested the utility of Moran's I, a measure of global spatial autocorrelation that measures the degree of similarity between nearby observations, for this purpose. Moran's I usually ranges from approximately -1 to 1, and can be interpreted as the extent to which values are spatially clustered (positive), dispersed (negative), or random (zero) (Fig 7A).

We performed Monte Carlo simulations to better understand the behavior of Moran's I in the presence of microscale and macroscale loss-of-function clusters (Fig 7B). While absolute changes in Moran's I increased with characteristic effect size *d*, the direction of the changes were dependent on loss-of-function cluster type. In the presence of macroscale loss-of-function clusters, Moran's I increased with *d*. Because, for wildtype *d*=0, this indicates that somatic mutants exhibited higher Moran's I than controls. In contrast, in the presence of microscale loss-of-function clusters, Moran's I decreased with *d*. Similar bidirectional changes were observed for other phenotypic measures (Fig 7C). Moran scatter plots revealed the basis for this bidirectional behavior (Fig 7D). These plots contain the original variable (e.g. Cent.Vol in vertebra 4) on the x-axis in individual fish and the spatially lagged variable (e.g. the mean of Cent.Vol in vertebrae 3 and 5) on the y-axis. For wildtype animals, there is natural phenotypic co-variation between vertebrae. For somatic mutants with microscale clusters, there was a greater prevalence of vertebrae with high phenotypic severity adjacent to those with low severity (and vice versa), resulting in loss of correlation in neighboring vertebrae. For somatic mutants with macroscale clusters, vertebrae within and outside of the cluster separate into the upper right and lower left quadrant, respectively, resulting in increased spatial autocorrelation in each animal. These demonstrate that Moran's I is sensitive to phenotypic spatial variation.

Finally, we assessed whether *plod2* and *bmp1a* somatic mutants exhibit distributional shifts in Moran's I compared to their age-matched controls. When the distribution of Moran's I was calculated across all 10 combinatorial measures, there was a marked shift in the center of the distribution toward I=0 for *plod2* somatic mutants compared to controls (Fig 7E). In contrast, no obvious shift in distribution center was observed for Moran's I in somatic mutants for *bmp1a* (Fig 7F).

Somatic mutant analysis identifies skeletal phenotype following loss-of-function in wnt16

While the above studies suggested the benefits of deep skeletal phenotyping in increasing the fidelity of CRISPR-based G0 screens, one limitation is that they were performed in zebrafish models of monogenic bone disorders. We sought to establish proof-of-concept that, by enhancing sensitivity in discerning somatic mutant phenotypes, G0 phenomic profiling may facilitate the detection of genes that are relevant to osteoporosis, a disease whose complex, multigenic nature has hampered the identification of causal variants and gene targets.

We investigated consequences of loss-of-function in *wnt16*, a candidate gene with strong ties to genetic risk for osteoporosis in humans (24). Somatic *wnt16* mutants were generated in the same manner as those for *bmp1a* and *plod2*. TIDE analysis revealed that *wnt16* mutation efficiencies for individual 12dpf larvae ranged from 67.4% to 93.6% with a median efficiency of 80.1%. Mutational efficiency at this site was confirmed using next generation sequencing (NGS) (Fig S5). At 90dpf, standard length (S.L.) was significantly reduced in *wnt16* somatic mutants compared to controls (control: 22.0mm [17.6-25.1mm], mutant: 18.8mm [16.0-22.6mm]; p=0.02; n=12/group; median [range]). MicroCT and FishCuT analysis revealed that *wnt16* somatic mutants exhibited significant changes in centrum volume (p=0.02) and centrum length (p=0.004) compared to sham injected clutchmates (n=12/group; Fig 8A). Evaluation of morphological defects (rib fracture calluses, neural arch non-unions, centrum compressions, and centrum fusions) revealed no significant differences in the occurrence of a single defect type in *wnt16* somatic mutants compared to controls. However, the proportion of animals exhibiting any such defect was significantly higher in *wnt16* somatic mutants (p=0.01, n=12/group) (Fig S6).

We examined whether the reduced centrum length and volume in *wnt16* crispants were solely attributable to developmental delay by determining whether these measures exhibit normal allometry with standard length. Following allometric normalization, significant differences were observed for centrum TMD (p=0.0001), haemal arch TMD (p=0.0001), neural arch TMD (p=0.0002), and neural arch thickness (p=0.03) (Fig S7). These data suggest that morphological differences in *wnt16* somatic mutants and sham-injected controls were not solely due to developmental delay, as if this were the case, no significant differences would be expected following allometric normalization.

We examined the benefits of phenomic profiling in detecting *wnt16* somatic mutant phenotypes. We tested this by performing bootstrap simulations, similar to those performed in somatic mutants for *plod2*. For Cent.Vol (assuming n=12), the sensitivity of the global test, M.-W. test using a single vertebra, and M.-W. test using averaged quantities was 0.71, 0.48, and 0.58. Specificity was 0.95, 0.95, and 0.96, respectively. For Cent.Le (assuming n=12), the sensitivity of the global test, M.-W. test using a single vertebra, and M.-W. test using averaged quantities was 0.85, 0.76, and 0.71. Specificity was 0.94, 0.95, and 0.94, respectively. Consistent with our previous analyses, the global test conferred higher sensitivity, with similar specificity, compared to M.-W. tests using averaged quantities as well as individual vertebrae, for somatic mutants for *wnt16*. We also examined changes in global spatial autocorrelation in *wnt16* somatic mutants, as indicated by Moran's I. No obvious difference in distribution in *wnt16* somatic mutants compared to controls was observed (Fig S8).

Finally, to determine if the somatic phenotype is representative of germline dysfunction in *wnt16*, we generated *wnt16* F2 germline mutants harboring a frameshift allele (a 10 base pair insertion in exon 3) and subjected these fish to phenomic profiling (n=4 +/+, n=6 +/-, and n=6 -/-). Germline *wnt16*-/- mutants exhibited significant changes in centrum volume (p=0.002) and centrum length (p=7.1e-6) compared to +/+ clutchmates (Fig 8B). Heterozygous *wnt16* germline mutants showed no significant changes compared to controls (Fig S9). Standard length was also significantly reduced in homozygous mutants (+/+: 17.4mm [15.8-17.9mm], -/-: 20.3mm [18.9-20.8mm]; p=0.01; median [range]). No obvious differences in the frequency of centrum compressions and neural arch non-unions were observed between groups (while germline mutant studies were underpowered to detect morphological defects). Taken together, our studies identify impaired bone mass accrual in adult zebrafish arising from loss-of-function in *wnt16*, a gene implicated in heritable risk for osteoporosis, and provide evidence of the fidelity of phenotyping directly in CRISPR-edited G0 somatic mutants for genes implicated as risk factors in this disease.

DISCUSSION

In order to realize the full potential of CRISPR-based rapid throughput biology in animal models, parallel advances in phenotyping are needed. By examining spatial and size characteristics of loss-of-function cell clusters, and performing deep phenotyping in mosaic models of bone disease, we identified loss-of-function and phenomic approaches to help detect and interpret spatially variable phenotypes in CRISPR-edited G0 animals.

By visually depicting functional loss in *sp7*+ cells, we identified distinct characteristics in regard to how loss-of-function cell clusters in the skeleton distribute in space and size following CRISPR editing. We found that Eq 1, which was derived based on modeling of fluorescence-based lineage tracing studies, described loss-of-function cluster sizes in skeletal tissues derived from both the sclerotome (vertebrae), and neural crest (branchiostegal rays), and in animals with different levels of mutation efficiency. Because this relation arises via contributions from clonal fragmentation and merger events rather than cell fates specified by developmental programs, the data collapse in our study is unlikely to be unique to skeletal tissues, as clonal fragmentation and merger events initiate in the early developing embryo (12). While it remains to be seen whether a similar size distribution in loss-of-function clusters is observed in other tissues and organs, in lineage tracing studies, Eq 1 has been found to fit clonal cluster size distributions in diverse contexts including during development of the mouse and zebrafish heart, the mouse liver, and the mouse pancreas (12). Notably, this size distribution is not observed in developmental events where clonal fragmentation and merger events are suppressed (e.g., mouse acinar cells (12)), and we would expect similar conditionality to apply in regard to CRISPR-induced somatic mutations. Still, a general relationship describing loss-of-function clusters sizes in G0 animals may facilitate the interpretation of spatially varying phenotypes commonly observed in G0 screens. This may also have value in other instances where estimates of cluster size distributions may be needed, such as in therapeutic applications of CRISPR editing.

While our studies implicate clonal fragmentation and merger as an important influence in the development of mosaic patterns, they do not imply that developmental programs are unimportant in this process. Indeed, within individual fish or subgroups of individuals, we could observe spatial patterns that were clearly non-random, and which could be explained by models of teleost spine development. For instance, in zebrafish, the centrum first ossifies through direct mineralization of the notochord sheath, which is then encased by intramembranous bone produced by somite-derived osteoblasts (25). In medaka, *twist*+ sclerotomal cells migrate dorsally around the notochord, and concentrate in the intervertebral regions prior to centrum ossification (26). These intervertebral cells supply osteoblasts to the centra flanking either side of the intervertebral region, as well as the adjacent neural arches (26). In this intervertebral growth center (IGC) model, osteoblasts on adjacent centra, as well as the arches they flank, are all descendants of the same intervertebral cells. Thus, two predictions of the

IGC model are that loss-of-fluorescence should be related in the centra of adjacent vertebra, as well as in the neural arches and centra of the same vertebral bodies; our observation of such events is congruent with the IGC model. In this context, our studies suggest that developmental programs and clonal fragmentation and merger events may work in concert to influence mosaic patterns; while only the latter contributes to size distributions in loss-of-function clusters, both contribute to spatial distributions.

One limitation of our mosaic-GFP analyses is that we did not examine the evolution of mosaic patterns during embryonic and post-larval growth. Because bone tissue accumulates throughout life, it is possible that spatial variations in somatic mutant skeletal phenotypes arise from evolving spatial patterns in mosaicism during growth. This can potentially be tested by visualizing patterns of CRISPR-induced lesions in osteoblasts longitudinally at different developmental stages. However, this will require overcoming the fact that *in vivo* fluorescent imaging of the spine is difficult in adults, even in mutants that lack pigmentation, due to the scattering of light in deep tissue.

By performing phenomic profiling in mosaic models of bone disease, our studies expand our understanding of manifestations of mosaicism in the skeleton. We found that mutant phenotypes could be expressed relatively uniformly in the spine (*bmp1a*), or in alternating patterns of affected and unaffected vertebrae (*plod2*). The relatively uniform phenotypic expressivity in *bmp1a* somatic mutants was discordant with the patchy loss-of-fluorescence patterns observed in *sp7:EGFP* fish. Discordance between cell genotype and phenotype in mosaic animals is a hallmark of systemic/non-cell autonomous gene action. Consistent with this notion, the action of BMP1 is extracellular, as it is a secreted enzyme involved in processing of collagen and other extracellular matrix proteins (27). Indeed, inter-individual phenotypic variability in *bmp1a* somatic mutants resembled a dosage curve, which has been previously hypothesized to occur when mutating a gene encoding a secreted factor, with different numbers of cells carrying the relevant mutation (28). In contrast to the extracellular action of BMP1, in humans, the action of PLOD2 is intracellular as it modifies collagen in the endoplasmic reticulum. As CRISPR-based G0 screens are still nascent, an important question is whether somatic mutants are equally likely to yield phenotypes for genes associated with cell autonomous or non-cell autonomous function (28). Our studies suggest that both can yield an adult skeletal phenotype in G0 animals, though they may be associated with different degrees of spatial variation.

Mosaic models of bone disease have clinical importance. In dominant forms of OI affecting type I collagen, mosaic individuals are usually unaffected, or mildly affected, by the mutation that clinically manifests in their heterozygous offspring (29, 30). Because transplantation of mesenchymal stem cells is an approach to OI therapy, the study of mosaics in animals and in humans is clinically informative, as the presence of both wildtype and mutant cells in the same individual is a natural model of stem cell therapy (30). In humans, it has been estimated that a 40%–75% burden of osteoblasts heterozygous for a COL1A1 mutation is compatible with largely normal skeletal function (30). Our studies establish tools to acquire dose-response relationships between mutation burden and phenotypic penetrance and expressivity in zebrafish mosaic models of OI, as a route to understand engraftment rates necessary for effective stem cell therapy for recessive forms of OI. Intra-individual variability in phenotypic expressivity in OI is, in general, not rigorously examined. However, there are instances of localized skeletal deficiencies, such as proximal femoral focal deficiency (PFFD), in which ~85-90% of cases are unilateral (31). While PFFD is usually sporadic (31), familial cases have been described (32), suggesting mosaicism as a potential root cause.

Our studies suggest that the ability to discern spatially variable phenotypes in mosaic individuals is limited by the tendency for phenotypic assays to be limited to single readouts. In Monte Carlo simulations, analyzing vertebral patterns conferred up to 3.65-fold higher sensitivity in discerning phenotypic changes in somatic mutants compared to single readouts. Fold changes in sensitivity were higher than the ~1.1-fold increase in sensitivity previously reported in simulations of *bmp1a* germline mutants, in which the global test was compared to t-tests of individual vertebra and averaged quantities (10). This is consistent with our simulations demonstrating that analyzing vertebral patterns is of increasing benefit when phenotypes become increasingly spatially variable. The increased sensitivity in discerning somatic mutant populations conferred by our workflow can improve G0 screen productivity, and facilitate the study of genetic variants of smaller effect sizes.

Understanding quantitative phenotypic variation arising from mosaicism is important to facilitate the development of analytical approaches to discern spatial co-variation and its relation to biological function. We found that global spatial autocorrelation, as measured by Moran's I, was sensitive to spatial phenotypic variation arising from mosaicism. While there was a clear shift in the center of the distribution for *plod2* somatic mutants compared to controls, no obvious shift was observed for Moran's

I in somatic mutants for *bmp1a*. One explanation for this discrepancy is the non-cell autonomous gene action of *bmp1a*, which encodes a factor secreted into the extracellular environment, and which may "smooth" phenotypic variations between neighboring vertebrae. Spatial statistics such as Moran's I may help to decode relationships between mosaic patterns and spatial phenotypic variation.

Human genome-wide association studies have identified over 200 loci (33, 34) associated with BMD and fracture risk in humans, however, most of the causal genes responsible for these associations have yet to be definitively assigned. Large-scale screens of candidate genes provide a mechanism to attribute functional skeletal contributions of these genes to osteoporosis-related traits. We identified a bone mass phenotype in zebrafish somatic and germline mutants for *wnt16*. Because this gene is implicated in mediating inter-individual variation in BMD and fracture risk in humans (35, 36), this broadens our understanding of zebrafish mutant phenotypes that may be predictive of human BMD and fracture risk. Somatic *wnt16* mutants successfully recapitulated phenotypic signatures in germline *wnt16* mutants, providing further evidence of the fidelity of phenotyping directly in CRISPR-edited G0 somatic mutants. Heterozygous *wnt16* germline mutants showed no significant changes compared to controls, suggesting that the somatic *wnt16* phenotypes arises from mosaic bialleleic loss-of-function. In mice, loss of WNT16 affects both osteoclastogenesis and bone formation [6]. In zebrafish, the phenotype seen for *wnt16* mutants was concentrated in the centrum (somatic and germline). In teleosts, centrum growth through the larval-to-adult transition has been attributed to osteoblasts originating in the intervertebral region (26). While it remains to be seen if faithful correlation between somatic and germline mutant phenotypes exists for most skeletal genes, microCT-based phenomics, in concert with CRISPR-based G0 screens, is a promising direction for the rapid identification of candidate genes and development of targeted therapies for mono- and multigenic skeletal diseases.

MATERIALS AND METHODS

For more detailed methods, see SI Appendix.

Zebrafish rearing

All studies were performed on an approved protocol in accordance with the University of Washington Institutional Animal Care and Use Committee (IACUC). Studies

were conducted in mixed sex animals from either the wild-type AB strain or the transgenic *sp7:EGFP* (11) background. A description of *plod2* and *bmp1a* germline strains used for comparative analysis is described in (10).

CRISPR-induced mutagenesis

CRISPR mutagenesis was performed using the Alt-R™ CRISPR-Cas9 System from Integrated DNA Technologies (IDT). For each gene, gRNAs were generated by mixing the crRNA and tracrRNA in a 1:1 ratio, diluting to 20 µM in nuclease-free duplex buffer (IDT), incubating at 95°C for 5 minutes and cooling on ice. Cas9 protein (20 µM, NEB) was mixed in a 1:1 ratio with the gRNA complex and incubated for 5-10 minutes at room temperate to produce the Cas9:gRNA RNP complex at a final concentration of 10 µM. RNPs were loaded into pre-pulled microcapillary needles (Tritech Research), calibrated, and 2 nL RNP complexes were injected into the yolk of 1- to 4-cell stage embryos.

Sequencing and mutation efficiency analysis

Between 24 and 96 hpf, a few embryos from each injection group were pooled, DNA extracted, Sanger sequenced (GenScript), and screened for mutagenesis. efficiency using the TIDE webtool (37). Individual animals were also screened for mutagenesis using whole larvae at 12 dpf to predict mutation efficiencies using the TIDE webtool, and to check for clonal fitness effects. These data are reported as intra-animal efficiencies in the main text. For next generation sequencing, a single *wnt16* sample was prepared according to the recommended guidelines for CRISPR amplicon sequencing at the MGH CCIB DNA Core affiliated with Harvard University. High-quality reads were aligned, identical and similar sequences (those with isolated SNPs not affecting indels were considered similar) were collapsed, and counts of unique alleles with more than 10 paired reads are reported.

MicroCT scanning and image analysis

MicroCT scanning was performed using a vivaCT40 (Scanco Medical, Switzerland) with 21 µm isotropic voxel resolution, and analyzed using FishCuT software (10, 19).

Fluorescent imaging

Between 10-12dpf, zebrafish were anesthetized in MS-222 and mounted into borosilicate glass capillaries using 0.75% low melt-agarose (Bio-Rad) diluted in system

water containing 0.01% MS-222. Dual-channel (GFP, excitation 450-490, emission 500-550; DAPI, excitation 335-383, emission 420-470) images were collected on a high-content fluorescent microscopy system (Zeiss Axio Imager M2, constant exposure settings for all experiments) using a 2.5x objective (EC Plan-Neofluar 2.5x/0.075). Maximum intensity projections were generated from image stacks in Fiji (38) for analysis.

Monte Carlo simulations

We previously characterized multivariate distributions for select FishCuT measures that exhibit evidence of multivariate normality (10). Using parameter estimates previously derived for such measures, we constructed wildtype and mutant distributions using methods described in Supplemental Information. Mutant phenotypic distributions were parameterized by several variables: *d* (characteristic effect size), λ (extent of intra-animal variation in phenotypic expressivity; λ=0 is most variable, λ=1 is least variable), and p_i (a 'loss-of-function vector' whose values range from 0 to 1, and which encodes the spatial pattern of loss-of-function in each vertebra). We examined two classes of loss-of-function vectors. The first class, $p_i^{microscale}$, simulated phenotypes arising from microscale loss-of-function clusters. The second class, $p_i^{macroscale}$, simulated macroscale loss-of-function clusters. When λ=1, mutant distributions were identical for both microscale and macroscale clusters; because loss-of-function was uniform in this case, we refer to simulated mutants with λ=1 as "germline mutants". For simulations, unless otherwise noted, we assumed a characteristic effect size of d=2, and a sample size of n=10, or a characteristic effect size of d=2.5, and a sample size of n=15. Similar methods were used for Monte Carlo simulations for Moran's I.

Statistical analysis

In general, the Mann-Whitney test was used for univariate analyses between two groups. Dysmorphic phenotypes were analyzed using a test for equal proportions. Multivariate analysis was performed using the global test. $p<0.05$ was considered statistically significant in all cases.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health (NIH) under Award Number AR066061 and AR072199. The content is solely the

responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. ZIRC is supported by NIH grant RR12546. We thank Drs. Peter Byers and Cecilia Moens for helpful discussions.

FIGURES

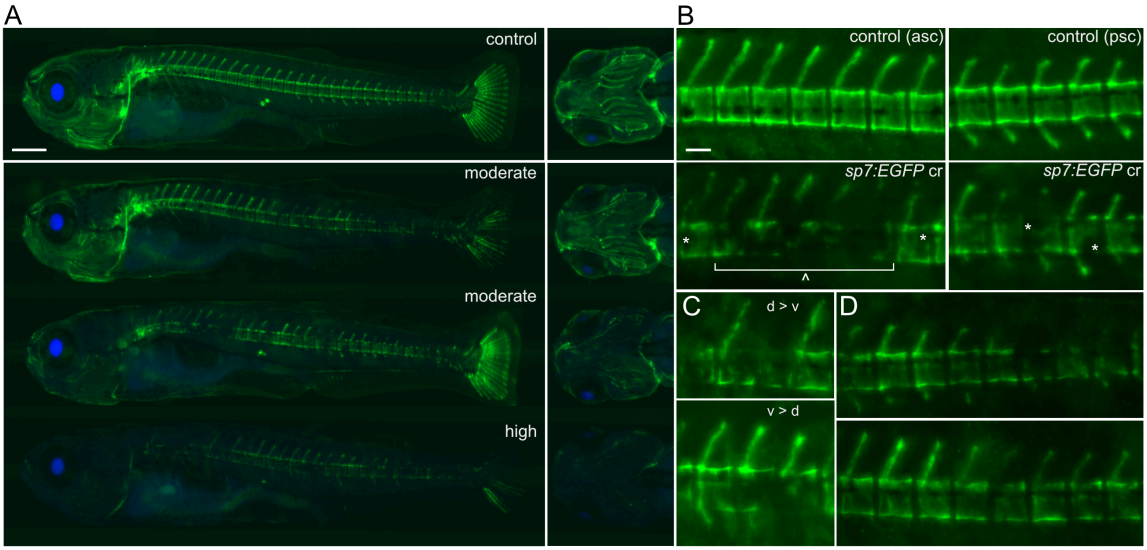

Fig 1. Somatic mutations in the *sp7:EGFP* transgene show mosaic loss of fluorescence. (A) Larval *sp7:EGFP* transgenic fish show relatively uniform EGFP expression in the skeleton of control fish (top). Somatic mutants with either a moderate (middle) or a high (bottom) degree of loss-of-fluorescence display a wide and varied range of loss-of-fluorescence cluster distributions, including in the craniofacial skeleton, the axial skeleton, and in the caudal fin rays. All fish presented are 10 dpf. Scale bar = 500 μm. (B) Loss-of-fluorescence occurs on both the macroscale (spanning multiple vertebral bodies, denoted by a caret) and the microscale (contained within a vertebral body, denoted by asterisks) compared to controls. Notice the distinction between opacity due to developing pigmentation in the controls (top) compared to loss of fluorescence in somatic mutants (bottom). asc, anterior spinal column; psc, posterior spinal column. (C) On multiple instances, loss-of-fluorescence in somatic mutants was stratified along the dorso-ventral axis of the centrum, with loss occurring preferentially on the dorsal side (top) or ventral side (bottom). (D) EGFP expression (or lack thereof) occurring on the dorsal side of the centrum in somatic mutants often corresponded to expression in the neural arch and spine in both the anterior (top) and posterior (bottom) spinal column. Scale bar for (B-D) = 100 μm.

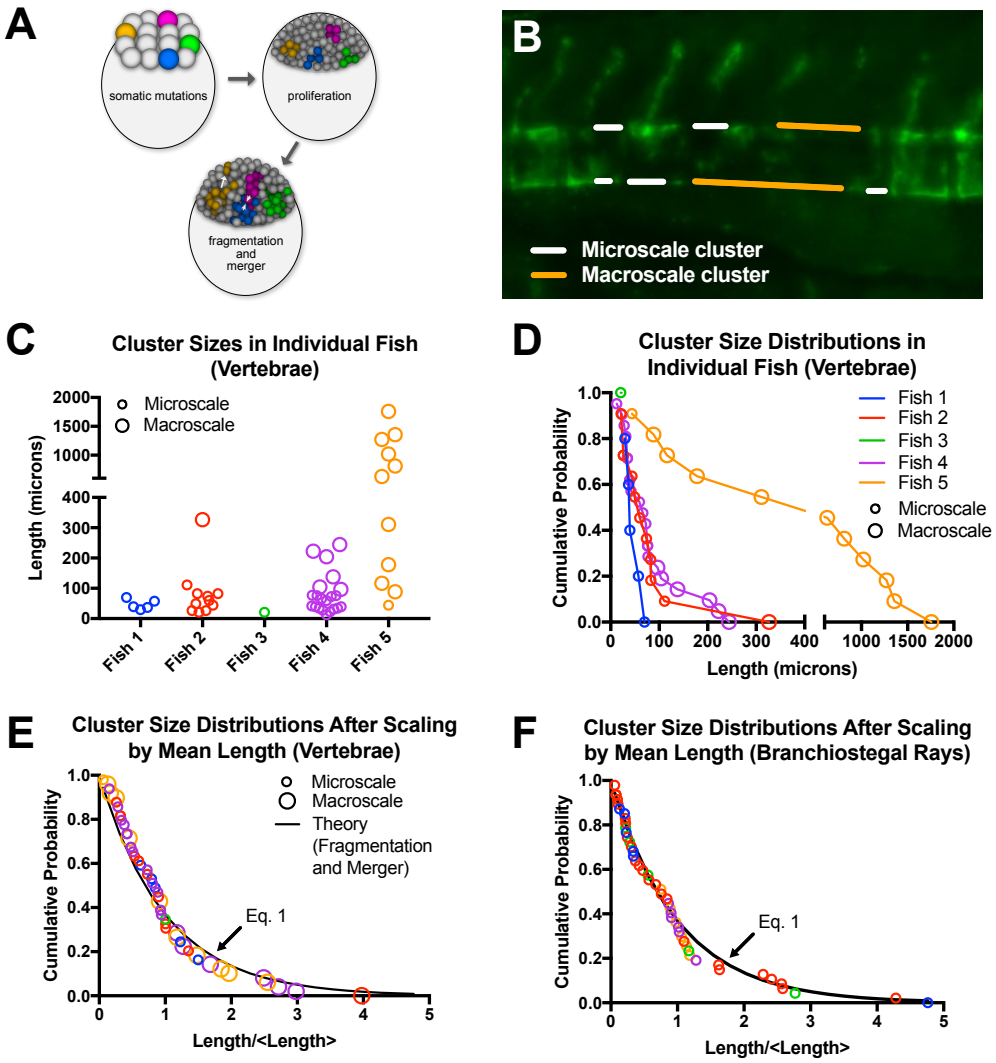

Fig 2. A common distribution underlies the sizes of loss-of-function clusters in bones of distinct developmental lineages, and in animals with different mutation efficiencies. (A) Schematic demonstrating clonal fragmentation and merger events. (B) Schematic demonstrating tracing of microscale (white) and macroscale (orange) loss-of-fluorescence clusters on the dorsal and ventral centrum surfaces. (C) Cluster sizes in vertebrae of individual fish. Note the differences in the number and sizes of clusters, indicative of differences in extent of loss-of-function in each fish. (D) Vertebral cluster size distributions in individual fish, using data from panel (C). (E) Vertebral cluster size distributions in individual fish, when normalized by the mean length in each fish. Color mapping is same as for panels C and D. Note the collapse of data onto a common distribution, which overlaps with Eq. 1. (F) Cluster size distributions for loss-of-function clusters in the branchiostegal rays. Color mapping is same as for panels C and D.

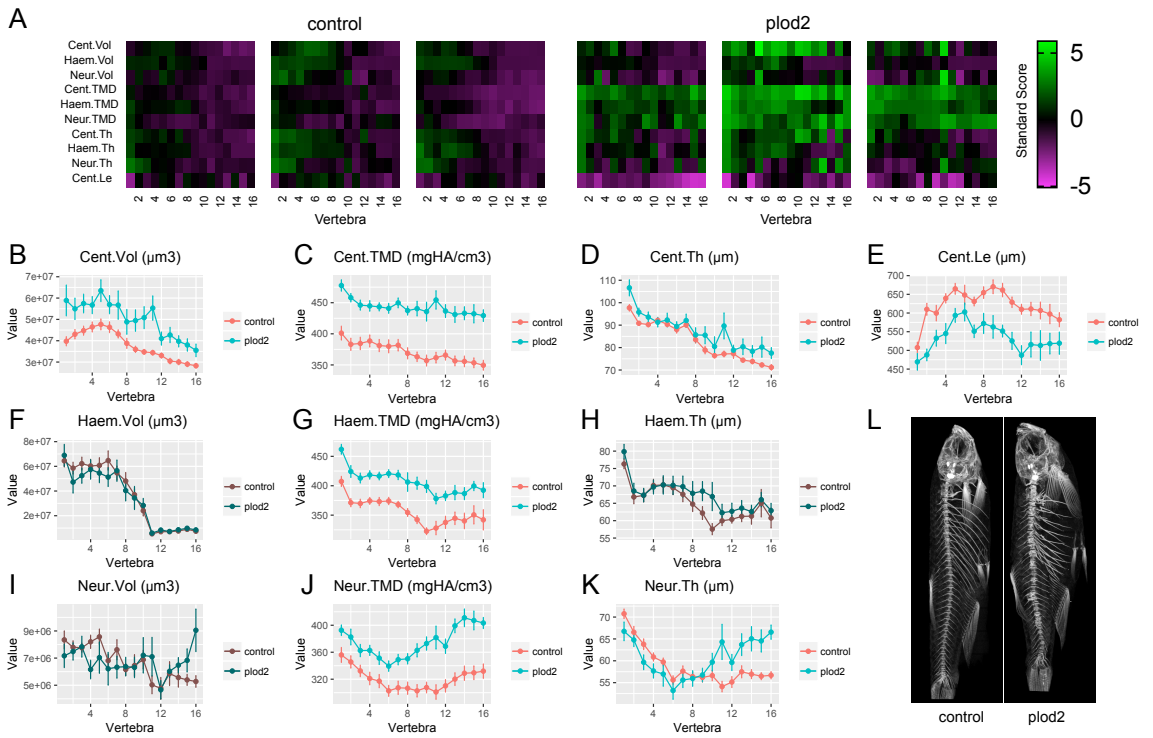

Fig 3. FishCuT analysis of *plod2* somatic mutants. (A) Skeletal barcodes for control and *plod2* somatic mutant fish. Each barcode represents a single fish (n = 3 fish per group shown). Standard scores are computed as the difference between the value of the feature in the individual and the mean value of the feature across all vertebrae in the control population, divided by the standard deviation of the feature across all vertebrae in the control population. (B–K) Phenotypic features plotted as a function of vertebra (mean ± SE, n = 11/group). Plots associated with $p<0.05$ in the global test are colored in a lighter coloring scheme (see text for p-values). (L) Maximum intensity projections of microCT scans.

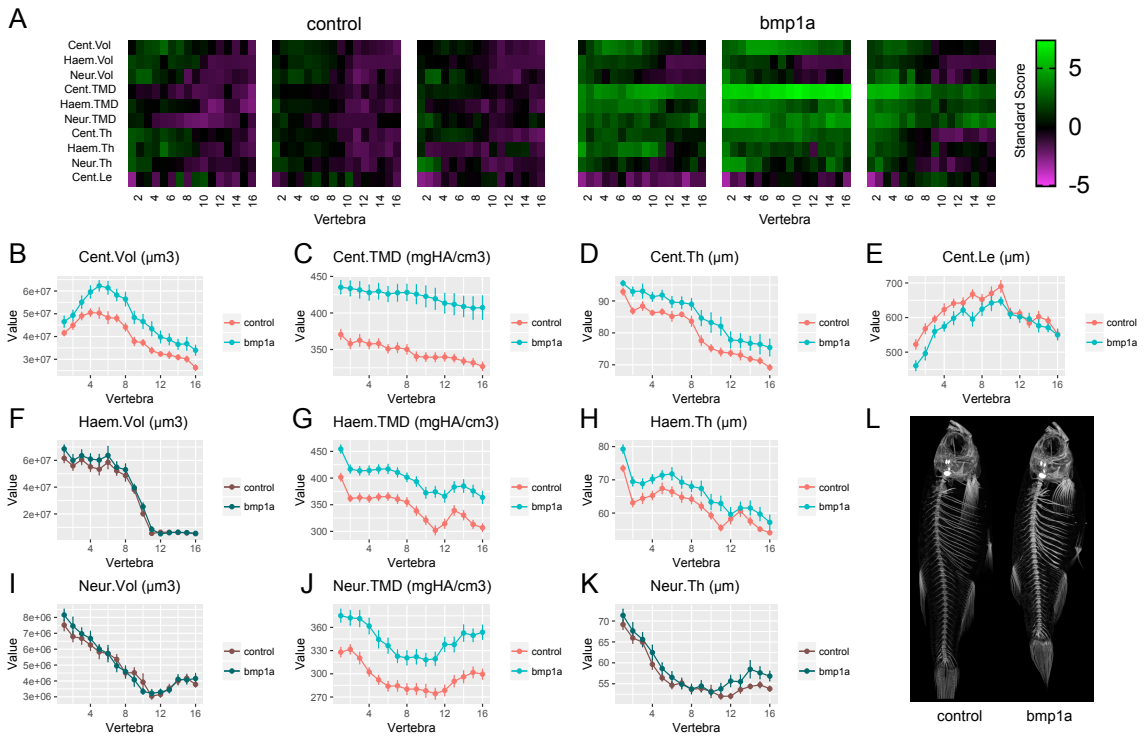

Fig 4. FishCuT analysis of *bmp1a* somatic mutants. (A) Skeletal barcodes for control and *bmp1a* somatic mutant fish. Each barcode represents a single fish (n = 3 fish per group shown). Standard scores are computed as the difference between the value of the feature in the individual and the mean value of the feature across all vertebrae in the control population, divided by the standard deviation of the feature across all vertebrae in the control population. (B–K) Phenotypic features plotted as a function of vertebra (mean ± SE, n = 15/group). Plots associated with $p<0.05$ in the global test are colored in a lighter coloring scheme (see text for p-values). (L) Maximum intensity projections of microCT scans.

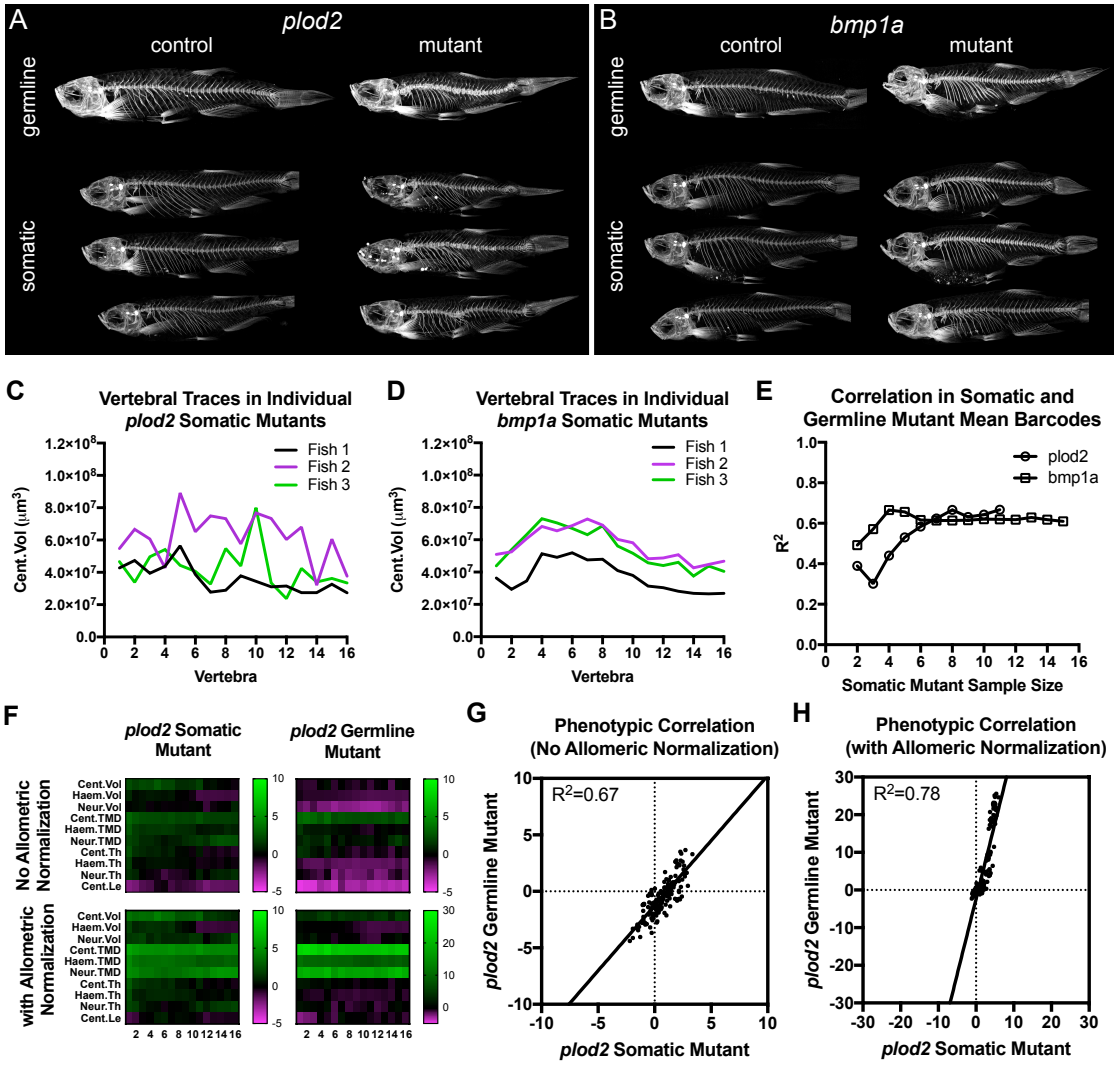

Fig 5: Somatic mutants for bmp1a and plod2 exhibit inter- and intra-animal phenotypic variability at the individual level, and phenotypic correspondence with germline mutants at the population level. (A-B) Representative maximum intensity projections of microCT scans for *plod2* (A) and *bmp1a* (B) mutants. (C-D) Vertebral traces for individual *plod2* (C) and *bmp1a* (D) somatic mutants depicted in (A-B). (E) Correlation in somatic and germline mutant mean barcodes as a function of somatic mutant sample size. Asymptotic behavior demonstrates convergence as sample size is increased. (F) Mean barcodes for *plod2* somatic and germline mutants in the absence and presence of allometric normalization. (G-H) Phenotypic correlation in *plod2* somatic and germline mutant barcodes in the absence (G) and presence (H) of allometric normalization. Each point represents a single phenotypic measure.

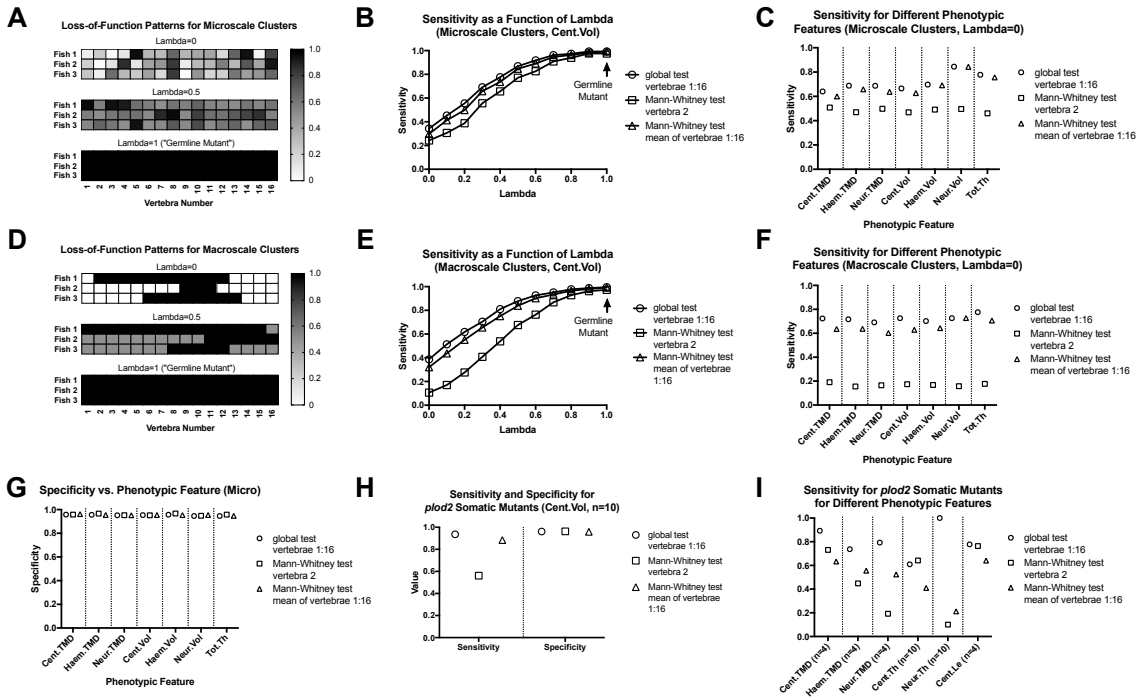

Fig 6. Analyzing vertebral patterns increases sensitivity in discriminating somatic mutants compared to analyzing single readouts. Figures A-C, and G show results for Monte Carlo simulations using microscale loss-of-function clusters. (A) Representative loss-of-function vectors for different values of λ (extent of intra-animal variation; λ=0: most spatial variation, λ=1: no spatial variation). (B) Sensitivity (fraction of times in which $p<0.05$ when comparing simulated mutant vs. WT fish) in discriminating increase in Cent.Vol as a function of λ. (C) Sensitivity for different phenotypic features. Figures D-F show results for Monte Carlo simulations using macroscale loss-of-function clusters. (D) Representative loss-of-function vectors for different values of λ. (E) Sensitivity in discriminating increase in Cent.Vol as a function of λ. (F) Sensitivity for different phenotypic features. (G) Specificity for different phenotypic features in Monte Carlo simulations using microscale loss-of-function clusters. Figure H-I show results for bootstrap simulations using phenotypic profiles from *plod2* somatic mutants. (H) Sensitivity and specificity for Cent.Vol. (I) Sensitivity for different phenotypic features.

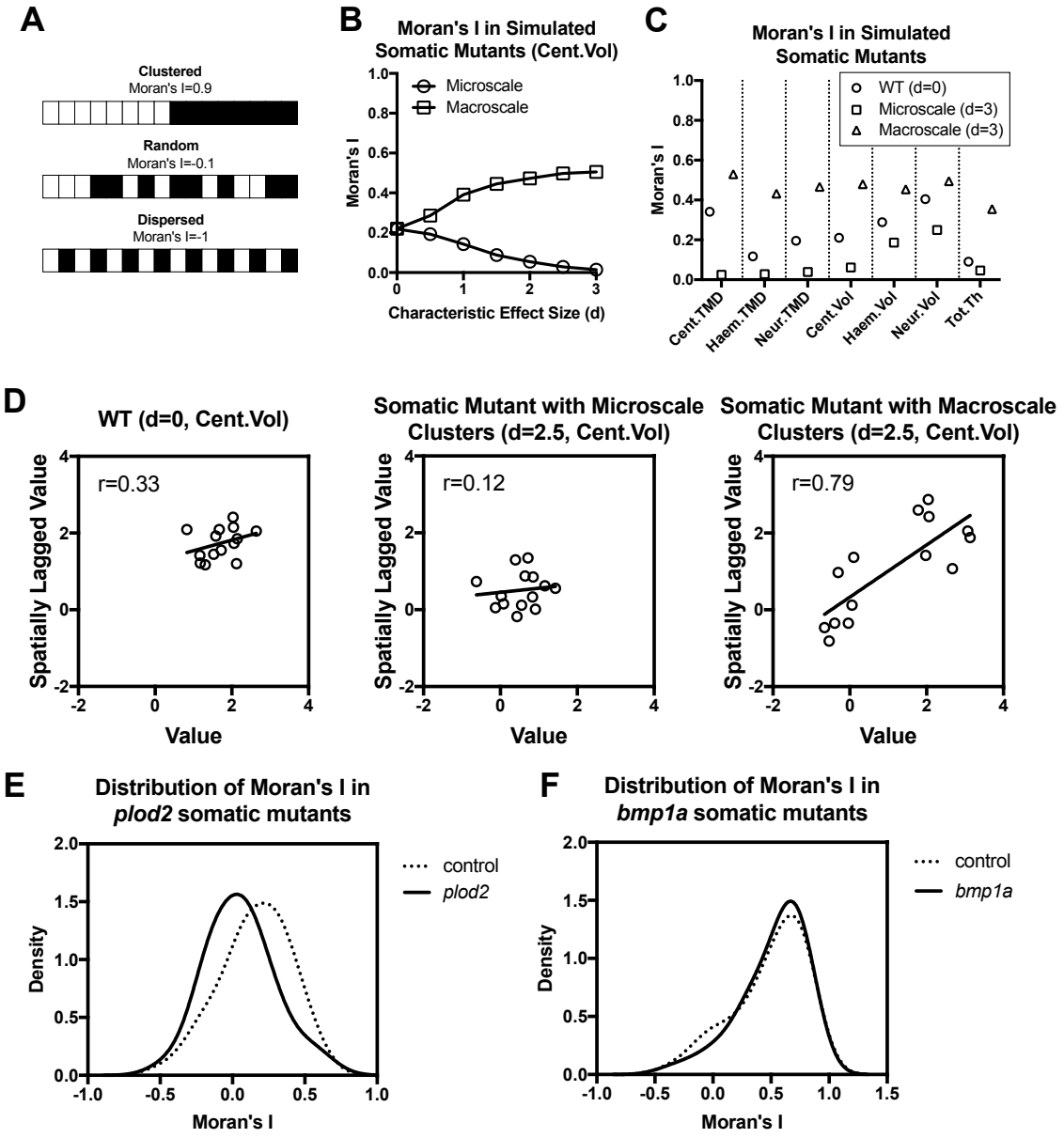

Fig 7. Phenomic profiling enables quantification of spatial phenotypic variation in somatic mutants. Figures A-D show results from Monte Carlo simulations. (A) Schematic demonstrating Moran's I for different phenotypic expressivity vectors (black=0, white=1). (B) Changes in Moran's I with characteristic effect size *d* are dependent on loss-of-function cluster type. (C) Moran's I for other phenotypic measures. (D) Moran scatter plots (i.e., plots with the spatially lagged variable on the y-axis and the original variable on the x-axis in individual fish). For somatic mutants with microscale clusters, there was a greater prevalence of vertebrae with high phenotypic severity adjacent to those with low severity (and vice versa), resulting in loss of correlation that arises from natural phenotypic co-variation between vertebrae. For somatic mutants with macroscale clusters, vertebrae within and outside of the cluster separate into the upper

right and lower left quadrant, respectively, resulting in increased spatial autocorrelation in each animal. (E,F) Moran's I for *plod2* and *bmp1a* mutants for Cent.Vol.

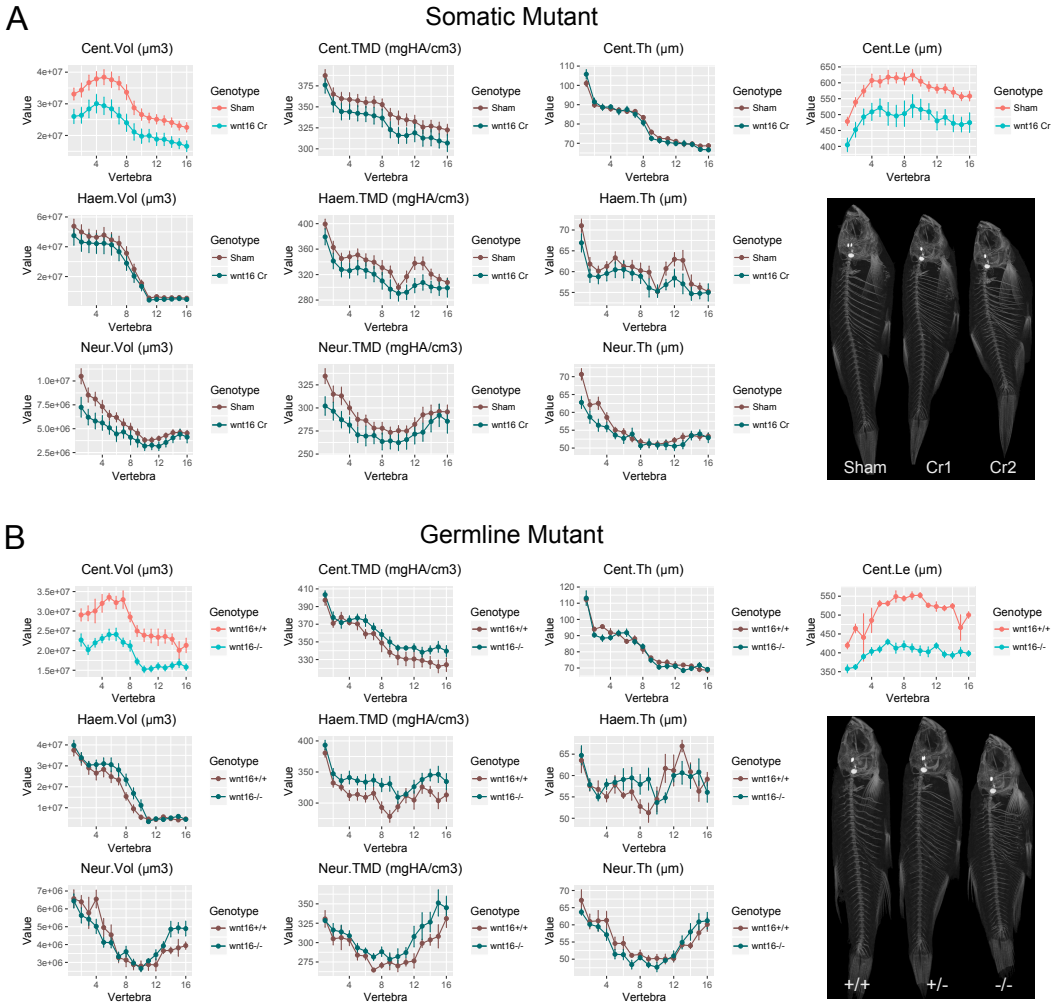

Fig 8. FishCuT analyses of *wnt16* somatic and germline mutants. In Figures A-B, phenotypic features are plotted as a function of vertebra (mean ± SE). Plots associated with $p<0.05$ in the global test are colored in a lighter coloring scheme (see text for p-values). (A) Analysis of *wnt16* somatic mutants. Somatic *wnt16* mutants show decreased centrum volume and centrum length compared to sham injected controls (n=12/group). Maximum intensity projections of microCT scans of representative sham fish (sham) and two somatic mutants displaying different expressivity (Cr1, Cr2) are shown. (B) Analysis of *wnt16* germline mutants. Germline *wnt16*-/- mutants show decreased centrum volume and centrum length compared to *wnt16*+/+ controls (+/+: n=4, -/-: n=6). Maximum intensity projections of microCT scans of representative fish of each genotype are shown.

REFERENCES

1. Shah AN, Davey CF, Whitebirch AC, Miller AC, & Moens CB (2015) Rapid reverse genetic screening using CRISPR in zebrafish. *Nat Methods* 12(6):535-540.

2. Wu RS*, et al.* (2018) A Rapid Method for Directed Gene Knockout for Screening in G0 Zebrafish. *Dev Cell* 46(1):112-125 e114.

3. Shalem O*, et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343(6166):84-87.

4. Sanjana NE (2017) Genome-scale CRISPR pooled screens. *Anal Biochem* 532:95-99.

5. Ata H*, et al.* (2018) Robust activation of microhomology-mediated end joining for precision gene editing applications. *PLoS Genet* 14(9):e1007652.

6. Freed D, Stevens EL, & Pevsner J (2014) Somatic mosaicism in the human genome. *Genes (Basel)* 5(4):1064-1094.

7. D'Agati G*, et al.* (2017) A defect in the mitochondrial protein Mpv17 underlies the transparent casper zebrafish. *Dev Biol* 430(1):11-17.

8. Bernards A & Gusella JF (1994) The importance of genetic mosaicism in human disease. *N Engl J Med* 331(21):1447-1449.

9. Biesecker LG & Spinner NB (2013) A genomic view of mosaicism and human disease. *Nat Rev Genet* 14(5):307-320.

10. Hur M*, et al.* (2017) MicroCT-based phenomics in the zebrafish skeleton reveals virtues of deep phenotyping in a distributed organ system. *Elife* 6.

11. DeLaurier A*, et al.* (2010) Zebrafish sp7:EGFP: a transgenic for studying otic vesicle formation, skeletogenesis, and bone regeneration. *Genesis* 48(8):505-511.

12. Rulands S*, et al.* (2018) Universality of clone dynamics during tissue development. *Nat Phys* 14(5):469-474.

13. Kague E*, et al.* (2012) Skeletogenic fate of zebrafish cranial and trunk neural crest. *Plos One* 7(11):e47394.

14. Gistelinck C*, et al.* (2018) Zebrafish type I collagen mutants faithfully recapitulate human type I collagenopathies. *Proc Natl Acad Sci U S A* 115(34):E8037-E8046.

15. Forlino A & Marini JC (2016) Osteogenesis imperfecta. *Lancet* 387(10028):1657-1671.

16. Gistelinck C*, et al.* (2016) Loss of Type I Collagen Telopeptide Lysyl Hydroxylation Causes Musculoskeletal Abnormalities in a Zebrafish Model of Bruck Syndrome. *J Bone Miner Res* 31(11):1930-1942.

17. Muir AM*, et al.* (2014) Induced ablation of Bmp1 and Tll1 produces osteogenesis imperfecta in mice. *Hum Mol Genet* 23(12):3085-3101.

18. Brinkman EK*, et al.* (2018) Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Res* 46(10):e58.

19. Hur M*, et al.* (2018) MicroCT-Based Phenomics in the Zebrafish Skeleton Reveals Virtues of Deep Phenotyping in a Distributed Organ System. *Zebrafish* 15(1):77-78.

20. Goeman JJ, van de Geer SA, de Kort F, & van Houwelingen HC (2004) A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 20(1):93-99.

21. Mcgill R, Tukey JW, & Larsen WA (1978) Variations of Box Plots. *Am Stat* 32(1):12-16.

22. Asharani PV*, et al.* (2012) Attenuated BMP1 function compromises osteogenesis, leading to bone fragility in humans and zebrafish. *Am J Hum Genet* 90(4):661-674.

23. Charles JF*, et al.* (2017) Utility of quantitative micro-computed tomographic analysis in zebrafish to define gene function during skeletogenesis. *Bone* 101:162-171.

24. Moverare-Skrtic S*, et al.* (2014) Osteoblast-derived WNT16 represses osteoclastogenesis and prevents cortical bone fragility fractures. *Nat Med* 20(11):1279-1288.

25. Fleming A, Keynes R, & Tannahill D (2004) A central role for the notochord in vertebral patterning. *Development* 131(4):873-880.

26. Inohaya K, Takano Y, & Kudo A (2007) The teleost intervertebral region acts as a growth center of the centrum: in vivo visualization of osteoblasts and their progenitors in transgenic fish. *Dev Dyn* 236(11):3031-3046.

27. Amano S*, et al.* (2000) Bone morphogenetic protein 1 is an extracellular processing enzyme of the laminin 5 gamma 2 chain. *J Biol Chem* 275(30):22728-22735.

28. Teboul L, Murray SA, & Nolan PM (2017) Phenotyping first-generation genome editing mutants: a new standard? *Mamm Genome* 28(7-8):377-382.

29. Pyott SM*, et al.* (2011) Recurrence of perinatal lethal osteogenesis imperfecta in sibships: parsing the risk between parental mosaicism for dominant mutations and autosomal recessive inheritance. *Genet Med* 13(2):125-130.

30. Cabral WA & Marini JC (2004) High proportion of mutant osteoblasts is compatible with normal skeletal function in mosaic carriers of osteogenesis imperfecta. *Am J Hum Genet* 74(4):752-760.

31. Doger E, Kopuk SY, Cakiroglu Y, Cakir O, & Yucesoy G (2013) Unilateral isolated proximal femoral focal deficiency. *Case Rep Obstet Gynecol* 2013:637904.

32. Makino Y*, et al.* (1998) A case of congenital familial short femur diagnosed prenatally. *Fetal Diagn Ther* 13(4):206-208.

33. Estrada K*, et al.* (2012) Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 44(5):491-501.

34. Kemp JP*, et al.* (2017) Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat Genet* 49(10):1468-1475.

35. Zheng HF*, et al.* (2012) WNT16 influences bone mineral density, cortical bone thickness, bone strength, and osteoporotic fracture risk. *PLoS Genet* 8(7):e1002745.

36. Gori F, Lerner U, Ohlsson C, & Baron R (2015) A new WNT on the bone: WNT16, cortical bone thickness, porosity and fractures. *Bonekey Rep* 4:669.

37. Brinkman EK, Chen T, Amendola M, & van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42(22):e168.

38. Schindelin J*, et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676-682.