Highly aneuploid zebrafish malignant peripheral nerve sheath tumors have genetic alterations similar to human cancers

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Highly aneuploid zebrafish malignant peripheral nerve sheath tumors have genetic alterations similar to human cancers

GuangJun Zhang*, Sebastian Hoersch**a,b, Adam Amsterdam*, Charles A. Whittaker*, Jacqueline A. Lees*, and Nancy Hopkins*a,1

*Department of Biology and David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139; and **Bioinformatics Group, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany

Contributed by Nancy Hopkins, August 17, 2010 (sent for review June 30, 2010)

Aneuploidy is a hallmark of human cancers, but most mouse cancer models lack the extensive aneuploidy seen in many human tumors. The zebrafish is becoming an increasingly popular model for studying cancer. Here we report that malignant peripheral nerve sheath tumors (MPNSTs) that arise in zebrafish as a result of mutations in either ribosomal protein (rp) genes or in p53 are highly aneuploid. Karyotyping reveals that these tumors frequently harbor near-triploid numbers of chromosomes, and they vary in chromosome number from cell to cell within a single tumor. Using array comparative genomic hybridization, we found that, as in human cancers, certain fish chromosomes are preferentially overrepresented, whereas others are underrepresented in many MPNSTs. In addition, we obtained evidence for recurrent subchromosomal amplifications and deletions that may contain genes involved in cancer initiation or progression. These focal amplifications encompassed several genes whose amplification is observed in human tumors, including met, cylindinD2, slo4S3a3, and cdk6. One focal amplification included fgf6a. Increasing fgf signaling via a mutation that overexpresses fgf8 accelerated the onset of MPNSTs in fish bearing a mutation in p53, suggesting that fgf6a itself may be a driver of MPNSTs. Our results suggest that the zebrafish is a useful model in which to study aneuploidy in human cancer and in which to identify candidate genes that may act as drivers in fish and potentially also in human tumors.

array comparative genomic hybridization | fgf8 | met | Illumina | copy number alteration

Chromosomal instability is a hallmark of human cancer (1, 2). It results in aneuploidy (a nondiploid number of chromosomes) and subchromosomal abnormalities, including inversions, translocations, deletions, and amplifications (3–5). Aneuploidy is particularly common in solid tumors (6). For example, in one study, 85% of colorectal cancers were aneuploid and possessed an average of 60–90 chromosomes (7). Such tumors are frequently heterogeneous, with the number of chromosomes varying from cell to cell.

Most of the chromosomal changes seen in complex cancer genomes are likely to be nonspecific by-products of chromosomal instability. Others, however, are clearly drivers of the cancer phenotype, including certain whole-chromosome amplifications and certain subchromosomal translocations, amplifications, and deletions. A major goal of cancer research is to distinguish pathogenetically relevant alterations from passive changes (2).

Cytogenetic technologies such as chromosome banding, fluorescence in situ hybridization (FISH), and comparative genomic hybridization (CGH) have been adapted for the purpose of characterizing numerical and structural chromosome abnormalities in cancers (8, 9). With CGH, chromosomal abnormalities are measured as DNA copy-number alterations (CNA). Recent large-scale CGH studies have found that human tumors possess preferred whole-chromosome, chromosome-arm, and focal CNAs (10, 11).

The conservation of gene function has made it possible to use animal models to study human cancer. The most powerful model system has been the mouse. However, most mouse cancer models do not display the extensive aneuploidy seen in many types of human tumors. For this reason, researchers have engineered mouse cancer models with chromosomal instability. Such mice develop tumors that are highly aneuploid and also possess subchromosomal alterations. These alterations can then be exploited to identify genetic alterations that drive the cancer phenotype (12).

Comparison of T-cell acute lymphoblastic leukemia (T-ALL) in this mouse model with human T-ALL using array CGH (aCGH) indicated important roles of FBXW7 and PTEN in this tumor type (12). In a manner similar to the use of the mouse model, comparison of CNAs in canine and human colorectal cancer has been used to extend the cross-species comparison strategy to identify human cancer genes (13).

The zebrafish is becoming a popular model organism for studying cancer, and a number of tumor models have been made by expression of oncogenes or the mutation of tumor suppressor genes (14). Zebrafish tumors have been shown to have similar gene expression signatures as human cancers (15). However, the nature of the zebrafish cancer genome, including numerical and structural changes, has been largely unexplored, although fluorescence-assisted cell sorting (FACS) analysis suggested that some tumors may be aneuploid (16–18) and low-resolution aCGH has indicated the presence of subchromosomal amplifications and deletions (19). Our laboratory previously reported that zebrafish heterozygous for mutations in any of 17 different rp genes develop malignant peripheral nerve sheath tumors (MPNSTs) (20, 21). This is an otherwise rare tumor type in our fish colony. Interestingly, fish homozygous for an inactivating mutation of p53 also develop MPNSTs (16). Here we report that MPNSTs that arise in either rp or p53 mutant zebrafish mimic human cancer in that they exhibit massive aneuploidy and heterogeneity within a single tumor. Furthermore, as in human cancers, custom-oligonucleotide aCGH and massively parallel synthetic sequencing reveal that despite their heterogeneity, fish MPNSTs display both preferred whole-chromosome copy-number alterations and significant focal copy-number alterations.

Results

Zebrafish MPNSTs Are Highly Aneuploid. To determine whether, like human MPNSTs, zebrafish MPNSTs that arise in either rp or p53 mutant fish are aneuploid, we first investigated the DNA content


The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, http://www.ncbi.nlm.nih.gov/geo/ (accession no. GSE23666).

1To whom correspondence should be addressed. E-mail: nhopkins@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1011548107/-/DCSupplemental.
of fish MPNSTs by FACS. Primary cells were isolated from large externally visible tumors and immediately stained with propidium iodide (PI) and then analyzed. Cells from most zebrafish MPNSTs contain at least one peak at a location expected for aneuploid cells, usually near 3N in DNA content and varying between 2N and 4N. In addition, there was usually a peak representing cells with 2N or near-2N DNA content in the same tumor, and the relative number of cells with 2N versus apparently aneuploid DNA content varied among tumors (Fig. 1 A, D, G, and J). We do not know whether the 2N peak represents contaminating normal cells or diploid tumor cells.

To exclude the possibility that the peak with higher DNA content was either an artifact of cell aggregation or 2N cells in the G2 or M phase of the cell cycle, we double-labeled tumor cells with PI and antibody to phosphorylated histone H3 (pH3), which stains cells in late G2 and M, and analyzed them by FACS. Consistent with the results of pH3 staining, most of the BrdU-positive cells had DNA content greater than that of cells in the presumptive aneuploid peak (Fig. S1). Taken together, these data suggest that the ~3N peak that is not an artifact of cell aggregation but rather represents aneuploid cells in G1. Furthermore, the results of pH3 and BrdU labeling suggested that, even in samples with both 2N and ~3N peaks, most of the proliferating cells are derived from the ~3N population. Staining of tumor cells with antibody to γ-tubulin demonstrated that the aneuploid cells had centrosome abnormalities similar to multipolar spindles often found in human cancers (Fig. S1).

To further characterize the aneuploidy in fish MPNSTs, we prepared metaphase chromosome spreads from tumors of five fish heterozygous for mutations in rp genes. To prevent potential artifactual changes in chromosome number caused by cell culture, colchicine was injected intraperitoneally into tumor-bearing fish and both normal and tumor cells were harvested 4 h later and fixed. We counted the chromosome number in 19 cells from normal tissue and 100 cells from each of the five tumors. We consistently found 50 chromosomes (the 2N number) in metaphase spreads from normal tissue (Fig. 1 B and C). In contrast, the chromosome number in tumor cells averaged around 70 per cell in each of the five samples. Notably, the chromosome number varied widely from cell to cell within each tumor, from 48 to 124 chromosomes (Fig. 1 E, F, H, I, K, and L). Thus, we conclude that, like many human solid tumors, zebrafish MPNSTs are both highly aneuploid and heterogeneous in chromosome number.

**Preferential Whole-Chromosome Alterations in Zebrafish MPNSTs.**

Aneuploidy is characteristic of most human solid tumors, and...
preferential gain or loss of particular chromosomes has been observed for both specific tumor types and across many types of cancers. Based on this precedent, we sought to establish whether there are common chromosomal copy-number changes among zebrafish MPNSTs. To this end, we designed a custom DNA microarray for aCGH experiments. Our Agilent custom array, designed against the zebrafish Zv7/danRer5 genome assembly, comprised about 15,000 60-mer probes at an average separation of ~100 kb. After the release of the Zv8 assembly, we remapped the positions of all of the probe sequences and eliminated any for analysis that were not assigned to an assembled chromosome, resulting in 13,646 usable probes. For the most part, the changes in the assembly did not affect the chromosomal coverage on our array, except that a large part of chromosome 4 unique to the Zv8 assembly turned out to be poorly represented.

We analyzed DNA from 36 tumors, 5 of them from fish homozygous for a mutation in p53, and 31 of them obtained from 13 different lines of rp heterozygotes. To avoid artifacts arising from polymorphisms in the fish genome, we used DNA from the tail of the same fish as the reference DNA for each tumor sample. Tumor samples and their respective reference DNAs were differentially labeled with Cy3 or Cy5, and hybridized to the same array. Data were normalized across the entire probe set and subjected to a circular binary segmentation algorithm. Thus, if a segment corresponding to an entire chromosome was found to be either above or below the baseline, we considered that to be a whole-chromosome gain or loss, respectively. Importantly, as these tumors are ~3N, the baseline is likely three copies. Thus, losses can represent two or fewer copies, whereas gains should represent four or more copies.

As in human tumors, we found that chromosome copy number changes in zebrafish MPNSTs were not random. The most dramatic and common change was a relative gain of chromosome 25 (Fig. 2). Other chromosomes that were frequently overrepresented were 10 and 11. The most commonly underrepresented chromosomes were 8 and 15. By contrast, chromosomes 3, 12, 13, 14, and 16, for example, showed relatively little deviation from baseline across the 36 samples or had similar-sized subsets of tumor samples showing over- and underrepresentation (Fig. 2). To confirm these findings, we sequenced several samples using the Illumina Genome Analyzer platform. Based on a total of at least 900,000 alignable reads for both normal and tumor tissues in three fish, we obtained nearly identical postsegmentation results for copy-number inter- and intrachromosomal variations as with aCGH (Fig. 3A). To confirm the aCGH results by yet another method, we used Southern blots to investigate an additional 14 tumors arising in rp heterozygous fish. For this analysis, we used probes against the commonly overrepresented chromosome 25, against the commonly underrepresented chromosome 15, and against chromosomes 13, 14, and 16.
16, whose sequences show little tendency to be either over- or underrepresented. Quantification of the hybridization signal from these probes confirmed that chromosome 15 was frequently underrepresented and chromosome 25 was frequently overrepresented in this independent sample of 14 tumors (Fig. S2); the amount of chromosomes 13 and 14 did not consistently vary (Fig. S3). Overall, these data suggest that certain chromosomes are preferentially gained or lost in highly aneuploid MPNSTs of zebrafish. We found the same trends in p53 mutant tumors as in rp mutant tumors, although p53 mutant tumors appear to additionally show common loss of chromosomes 3 and 8 and gain of chromosome 24 (Fig. 2). We would need to analyze more p53 samples to be confident that these changes are consistently preferred in that genetic background.

**Subchromosomal Amplifications on Chromosome 25.** In addition to whole-chromosome changes, both aCGH and sequencing revealed subchromosomal areas of copy-number variation in fish MPNSTs (Figs. 2 and 3A). Striking areas of amplification occurred on the already overrepresented chromosome 25. To verify these changes by another method, we used Southern blot analysis of DNA from the 14 additional rp heterozygous tumors described above. As controls, we again used probes that hybridize to either chromosome 13, 14, or 16. We used three probes that hybridize to different regions of chromosome 25; two of the probes (Sc and M) hybridize to one or the other of the two regions found to be highly amplified on this chromosome by aCGH analysis, and one probe (T) hybridizes to an area that tracks with the rest of chromosome 25 in being overrepresented but not additionally amplified (see Fig. 3B for probe locations). In at least a third of the 14 tumors analyzed by Southern blot, sequences of chromosome 25 represented by the Sc and/or M probes were amplified relative to the T region on chromosome 25 (Fig. 3C and Fig. S4). As observed previously with aCGH on the initial panel of 36 tumors, some presented by the Sc and/or M probes were amplified in different regions of chromosome 25; two of the probes (Sc and M) hybridized to one or the other of the two regions found to be highly amplified on this chromosome by aCGH analysis, and one probe (T) hybridizes to an area that tracks with the rest of chromosome 25 in being overrepresented but not additionally amplified (see Fig. 3B for probe locations). In at least a third of the 14 tumors analyzed by Southern blot, sequences of chromosome 25 represented by the Sc and/or M probes were amplified relative to the T region on chromosome 25 (Fig. 3C and Fig. S4). As observed previously with aCGH on the initial panel of 36 tumors, sometimes both Sc and M regions were amplified, sometimes just one or the other was amplified.

**Regions of Common Focal Amplification or Deletion in Fish MPNSTs.** Identifying regions of focal amplification or deletion that are biologically relevant in cancers with complex genomes is not trivial, because of the large number of random alterations that occur in such tumors. Here we first identified regions that appeared to be either commonly gained or lost by visual inspection of heatmap displays of segmented aCGH data for each chromosome. To arrive at a final list, we then combined our findings with the results of a permutation-based test (22) to determine which regions met the criterion of statistical significance in terms of ruling out the null hypothesis that overlaps between segments in different samples are merely random. We used this two-pronged approach because, whereas statistical significance is obviously impossible to judge by eye, the algorithmic results themselves are subject to certain limitations and artifacts requiring human input. For example, a key parameter in such analysis is a fixed cutoff level that is required for segment means to be considered for candidate focal regions. Although we generally used a cutoff of ±0.2 (segment mean), we found that for chromosome 25, a higher cutoff level (0.6) was needed in addition to adequately take into account the large dynamic range (up to a segment mean of 3.47) encountered here.

We identified two types of apparent focal changes. The first type was similar to what has been observed in human tumors: amplifications or deletions with varying boundaries from tumor to tumor, among which areas of overlap were evident. We found eight such regions on five different chromosomes among the 36 tumor samples analyzed. Together, they contain ≈220 genes (Dataset S1). Seven of these regions were gains, one was a loss. Two of the amplified areas identified lie on chromosome 25, one near the Sc probe and one near the M probe described above. The Sc region includes five to seven genes: ccnd2a, tigara, fgfb6a, slc45a3, c1orf4, and possibly dyrk4 and ndufl9 (these last two lie between the last probe in the segment and the first probe outside of the segment). The M region includes 14 genes, one of which is the proto-oncogene met, which has been previously identified as a gene frequently amplified in human MPNSTs and which is activated by mutation or overexpressed in perhaps 50% of human tumors (23, 24). The other genes in this region are amtl2, loki12r1, cep290, tmtc3, kifla, duasp6, pocc1b, tsga14, cpal1, cpal5, tes, cav2, and cav1.

In the second type of focal change, two to five fish had the same few (two to six) probes representing an amplification or deletion. We found eight cases of this type. Although we cannot rule out the possibility that these are genuine copy-number changes, it seems highly possible that these are cases where, even in the most recent assembly of the zebrafish genome (Zv8), the genomic positions of these probes are misassigned. Should these probes in reality be on chromosomes that are (for that tumor) up or down at the whole-chromosome level relative to the chromosome to which they have been misassigned, they would appear to represent a focal change. In initially working with the aCGH data with the probes mapped to the previous assembly (Zv7), we had observed multiple instances of this type of focal event, and nearly all of them disappeared after probes were reassigned based on the Zv8 assembly. Thus, we conclude that this type of “narrow” focal change cannot be called with any reliability under the present circumstances.

To determine whether genes in a focally amplified region may contribute to the initiating process, we also carried out an additional search that introduces a biological assay. As described next, we obtained preliminary evidence that fgf6 might be a driver of zebrafish MPNSTs.

**Preliminary Biological Validation of a Candidate Gene from a Focal Amplification: fgf8 Overexpression Accelerates Tumorigenesis in a p53 Mutant Background.** The fgf gene family contains many members, but in mammals they can all bind to four FGF receptors, members of the receptor tyrosine kinase superfamily, suggesting that they might have similar ability to signal through common MAP-kinase signaling pathways. Although we do not have a mutant line of fish that overexpresses fgf6a, a mutant line designated Hаг31 overexpresses fgf8a as a result of a retroviral insertion (25). This line develops large neuroblastomas at a low frequency and rarely before 1 yr of age (25). To determine whether fgf8a may be a driver gene for MPNSTs (as well as neuroblastomas), we introduced the Hаг31 mutation into fish heterozygous or homozygous for a mutation in p53. fgf8a overexpression accelerated the onset of MPNSTs in both p53 heterozygous and homozygous backgrounds (Fig. 4). This is consistent with the possibility that, by analogy, fgf6a may also be a driver in MPNSTs, and it demonstrates the feasibility of testing candidate genes identified by genomic approaches.

**Discussion**

Our results suggest that the zebrafish may be a useful model in which to study aneuploidy in human cancer. As with many human cancers, we showed that zebrafish MPNSTs that arise in fish heterozygous for rp mutations or homozygous for a p53 mutation are highly aneuploid, frequently possessing pseudotriplloid genomes. Furthermore, as in human cancers, the number of chromosomes per cell is extremely heterogeneous within a single zebrafish MPNST. Finally, aCGH, massively parallel sequencing, and Southern blotting revealed that despite the heterogeneity seen in fish tumors, some copy-number changes dominate any given tumor, presumably because cells with those changes have been selected for better growth. These copy-number alterations include both whole-chromosome and subchromosomal regions.

The similarities in genomic changes between fish and human MPNSTs extend beyond single tumors to properties shared by multiple tumors. We observed that among the 36 zebrafish MPNSTs analyzed by aCGH, whole-chromosome copy number changes are far from random: Several chromosomes are frequently gained and seldom lost in many tumors, whereas others show frequent loss and
rare gain. Recent analysis of large numbers of human tumors by CGH and sequencing shows that many chromosomes or chromosome arms are preferentially gained or lost in particular tumor types, and some preferences are shared across many human cancers (10, 11). It is postulated that the preferential over- or underrepresentation of certain chromosomes or chromosome arms in both human and zebrafish tumors may reflect a growth advantage conferred by genes that lie on these chromosomes.

In addition to whole-chromosome changes, we also found subchromosomal segments of gain or loss in fish MPNSTs. Using statistical methods to identify common focal areas of copy-number variation in human tumors has been a powerful tool for distinguishing focal changes that contain driver genes from those that are merely a harmless consequence of genomic instability. We used similar methods here and identified eight statistically significant regions of gain or loss manifest in several tumors. An analysis of more tumors, as well as use of a higher-resolution platform (both the aCGH and sequencing approaches used here had windows of about 0.1 Mb) will likely identify additional focal areas as well as narrow their size. This analysis in zebrafish is currently hampered by the unfinished nature of the zebrafish genome assembly, as well as incomplete gene annotation, but both will surely improve. In addition, adapting sophisticated tools such as GISTIC (26) for use with zebrafish, especially in combination with an increased sample size, could improve the statistical robustness of our findings. Meanwhile, the focal changes we have found thus far, although certainly not complete, clearly contain genes whose amplification is observed in human tumors, such as met, cyclinD2, slc45a3, and cyclin-dependent kinase 6. met is of particular interest because it has been identified as a gene that is often overexpressed in human MPNSTs (23). Furthermore, an increase in copy number of an activated met oncogene is thought to underlie the common gain of an extra chromosome 7 in papillary renal carcinomas (27), and increased expression of met may in part explain the common overrepresentation of chromosome 7 in many human cancers.

The high degree of aneuploidy observed in fish MPNSTs could prove to be extremely beneficial in helping to identify important drivers in human cancer. Most murine cancer models do not show extensive aneuploidy. As a result, mouse cancer models have been engineered specifically to generate a greater degree of genomic instability (e.g., 12). Using such models, a comparison of syntenic regions between mouse and human chromosomes that show copy number variation in tumors in both organisms has been helpful in further identifying biologically relevant focal changes in human tumors (12, 28). A similar approach using shared copy-number alterations between fish and human cancers could be even more powerful. This is because mouse and human are relatively close evolutionarily, so focal changes shared by these two organisms tend to share nearly all of the same genes, making it difficult to separate driver genes from passengers that merely cosegregate due to proximity. In contrast, zebrafish and human are far more evolutionarily distant, and syntenic blocks between them tend to be much shorter (29). Thus, focal regions in each organism containing the same driver gene are unlikely to share many additional genes, making it much easier to narrow down the list of candidate driver genes within the shared focal region.

Even more difficult to analyze in human tumors has been the significance of preferred chromosome or chromosome arm-level changes. As with focal regions, because of the nature of the synteny between fish and human genomes, a comparison of the genes on chromosomes preferentially gained or lost in human and zebrafish tumors could severely limit the number of candidate genes to examine. Although some kind of biological assay would be needed, such as the effects upon growth and survival properties in tumor cell lines, it is far easier to study the effects of knockdown or overexpression of a few dozen genes on a chromosome than of several hundred. Furthermore, as the example here with overexpression of fgf8a in zebrafish with p53 mutations shows, it is also possible to test the effects of copy-number alterations of specific genes on tumorigenesis in the fish itself, including in the context of different tumor-promoting mutations or transgenes.

The goal of much cancer research today is to identify the combination of driver genes that allow each and every human tumor to proliferate and evolve. The ability to type many highly aneuploid tumor genomes by next-generation sequencing and aCGH in zebrafish should help to parse tumors into groups with specific combinations of drivers generated by copy-number alterations and then test combinations of inhibitors to treat the different subclasses of cancer.

Materials and Methods

Zebrafish Lines. Tumor-prone lines of zebrafish used in these studies have been described previously. They include fish heterozygous or homozygous for a point mutation in p53 (16), fish heterozygous for an insertional mutation in any one of 15 different ribosomal protein genes (21), and fish with the insertion allele Hag225, which overexpresses fgf8a (25). Stocks were maintained as described previously and genotypes were determined by PCR at 8–18 wk of age as described (30).

Genomic DNA Isolation and Southern Blot Analysis. Genomic DNA from tail tissue and tumors was prepared as in ref. 31. Approximately 5 μg of each sample was cut with HindIII and Southern blots were conducted as in ref. 32. Details of probe sequences and quantitation are provided in SI Materials and Methods.

Flow Cytometry. Tumors were dissected out after euthanization with 500 mg/L tricaine, and tail fins were clipped to obtain normal cells from the same fish. After euthanization, tumors were dissected out of p53 homozygous or rp heterozygous fish between the ages of 8 and 24 mo. Single-cell suspensions were made by digestion at room temperature for 1 h in 2 U/mL dispase (InVitrogen) in PBS, and were strained through a 40-μm filter. Cell suspensions were then pelleted and washed one time with PBS before being fixed with 70% ethanol and stored at −20 °C. For BrdU labeling, 10 mM BrdU was injected into the peritoneal cavity of tumor-bearing fish, and the tumor was dissected out about 45 min later. BrdU, pH3, and propidium iodide (PI) double staining were performed as has been described for embryonic cells (33). FACS analysis was conducted by FACScan (Becton-Dickinson). DNA content, pH3, and BrdU data were analyzed by FlowJo (Tree Star).

Cytogenetics and Chromosome Counting. Colchicine (0.025%; 10 μL/g weight) was injected peritoneally into fish 4 h before dissection. After fish had been killed, tumors were dissected out and cell suspensions were prepared in the same way as for FACS. Cells were treated with 0.5% KCl and incubated at 35 °C for 30 min. Following hypotonic treatment, the cells were fixed in Carnoy’s fixative (75% methanol, 25% acetic acid) and stored at −20 °C. Chromosome spreads were made as described (34). For chromosome counting, slides were stained with either DAPI or Giemsa. Pictures were taken at 1,000x magnification using a Zeiss Axioplan II upright microscope, and chromosome numbers were counted using imageJ (National Institutes of Health).
Array CGH and Data Processing. Genomic DNA was isolated from tumors and from tail tissue of the same fish (used as reference) as in ref. 35. Samples were prepared and hybridized to the array containing 13,848 probes mapped to as- semblized chromosomes in Zv8/danRer6 according to the Agilent standard proto- col with some modifications. The arrays were scanned with Agilent scanner G2505B, and Agilent Feature Extraction software v9.1.3.1 was used to obtain normalized data (column logRatio in output file) that were subsequently con- verted from log2 to log2 scale. Normalized log2-ratio data were submitted to the circular binary segmentation algorithm (26) as implemented in the Bio- Conductor package DNAcopy (v1.16.0) and processed with default parameters. The segments and segment mean values obtained were the basis for subsequent analyses. For two sample pairs, an ancillary dye swap was identified by analysis and comparison of single-channel data, and segment means for these samples were multiplied by –1 to correct this mistake. The STAC algorithm (22) in a Java implementation provided on the authors’ Web site (http://cbil.upenn.edu/STAC v1.2) was applied to assess whether the overlap of two or more subchromosomal segments was likely to be a chance event. Full datasets are publicly available at http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE23666. Ad- ditional details on the design of the array, labeling and hybridization conditions, and STAC analysis are provided in SI Materials and Methods.

Massively Parallel Synthetic Sequencing. Samples for sequencing with the illumina Genome Analyzer lix system were prepared from the same batches of genomic DNA used for aCGH, according to published methods (27–39).

Sequences were processed with either Bustard.py (OLB 1.6.0) or GERALD.pl (CASAVA 1.6.0), downloaded from the illumina Web site (http://www.illumina. com/software.GenomeAnalyzer_analyser.software.htm), with a produced read length of 41 nucleotides. Further details on the analysis of the sequencing data are provided in SI Materials and Methods. Sequencing data are available at http:// www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE23666.

Note Added in Proof. In further support of the concern that the very small focal CNAs that were observed in several fish might have been an artifact of misassignment of the underlying probe sequences in the genome assembly, we found that probe sequences for 5 of the 8 cases map to a different chromosome in the pre-Ensembl release of the Zv9 assembly in July 2010.

ACKNOWLEDGMENTS. We thank Tim Angelini and Kate Anderson for maintenance of the zebrafish colony. We thank Eduardo Torres and Angelika Amon for assistance with aCGH experiments. We thank members of the Lees laboratory for helpful discussion, especially Christopher Sansam, and for help with the FACS analysis. We thank Kim Brown, Lynda Lee, Charles Lee, and members of their laboratories for helpful discussions. We thank Agilent Technologies for making available to us their zebrafish CGH/CHIP optimized probe database. This work was supported by a grant from Arthur C. Merrill to N.H., a National Institutes of Health–National Cancer Institute grant (CA106416) to J.A.L. and N.H., and a core grant (CA14051) to the Koch In- stitute for Integrative Cancer Research. J.A.L. is a Ludwig Scholar at the Mas- sachusetts Institute of Technology.
Supporting Information

Zhang et al. 10.10173/pnas.1011548107

SI Materials and Methods

Immunohistochemistry and Immunofluorescence. Tissues were fixed in 4% paraformaldehyde at 4°C for up to 2 d. The fish were dehydrated in alcohol, cleared in xylene, and infiltrated with paraffin. Cell suspensions were made before 4% paraformaldehyde fixation. Primary antibody against phosphorylated histone H3 (pH3) (Santa Cruz Biotechnology) was diluted 1:100 and the antibody against γ-tubulin (Sigma) was used at 1:100. Nuclei were counterstained with DAPI for visualization using a Zeiss microscope.

Array CGH and Data Processing. Array probes (15,104) against zebrafish chromosomes 1–25 were selected from a pool of more than 7 million zebrafish probes kindly provided by Agilent. Preprocessing was performed excluding probes containing AluI and RsaI restriction sites to ensure compatibility with an enzymatic comparative genomic hybridization (CGH) protocol for genomic DNA digestion for labeling purposes, applying a cutoff to Agilent-provided probe quality scores, and excluding probes exhibiting perfect secondary matches down to 21 nucleotides as assessed by a BLAT (1) search against zebrafish genomic sequence (Zv7/danRer5 assembly). Among the resulting set of 60,000 probes, the target number of about 15,000 probes for chromosomes 1–25 was reached by an iterative heuristic for selecting approximately equidistant probes. After the Zv8/danRer6 genome assembly was released, we remapped the probes by BLAT search against the improved assembly Zv8/danRer6. Probes were flagged as “unreadable” if they mapped with high quality to more than one assembled chromosome (1–25) in Zv8/danRer6, resulting in 13,648 available probes. With the exception of a novel and highly repetitive portion on chromosome 4, probe coverage, although not as uniform as for the Zv7/danRer5 assembly underlying the original design, was satisfactory under the new assembly.

Genomic DNA was isolated from tumors and from tail tissue of the same fish (used as reference) as in ref. 2. Samples were prepared and hybridized to the arrays according to the Agilent standard protocol with some modifications. Briefly, after fragmentation of 5 μg of tumor or reference DNA, DNA samples were labeled with Cy3-dCTP or Cy5-dCTP by random priming. Unincorporated nucleotides were removed using a QIAquick PCR purification kit (QIAGEN). After hybridization the probes were measured by NanoDrop 1000 (Thermo Scientific), and were denatured by incubation at 95°C for 2 min followed by cooling to 30°C in 100% ethanol. Array hybridization was carried out at 65°C with about 200 ng probes/array in Agilent HI-RPM hybridization buffer. After washing, the arrays were scanned with Agilent scanner G2505B, and Agilent Feature Extraction software v9.1.3.1 was used to obtain normalized data (column logRatio in output file) that was subsequently converted from log2 to log10 scale. Normalized log2 ratio data were submitted to the circular binary segmentation algorithm (3) as implemented in the BioConductor package DNAcopy (v1.16.0), and processed with default parameters. The segments and segment mean values obtained were the basis for subsequent analyses. For two sample pairs, an accidental dye swap was identified by analysis and comparison of single-channel data, and segment means for these samples were multiplied by −1 to correct this mistake. Full datasets are publicly available at http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE23666.

Assessment of the Significance of Focal Changes with STAC. The STAC algorithm (4) in a Java implementation provided on the authors’ Web site (http://cbil.upenn.edu/STAC; v1.2) was applied to assess whether the overlap of two or more subchromosomal segments was likely to be a chance event, given the chromosome-specific incidence of observed segments (STAC “footprint P value”). Segments indicative of relative gain and relative loss were processed separately. Only segments smaller than 20 MB and with a mean value above (below) 0.2 (−0.2) were considered for analysis, with the exception of gains on chromosome 25, where a mean value cutoff of 0.6 was used to better account for the high density of segments with mean values above 0.2 and their unusually broad mean value range. STAC was run with a resolution of 50,000 nucleotide positions and 5,000 permutations.

Massively Parallel Synthetic Sequencing. Samples for sequencing with the Illumina Genome Analyzer IIX system were prepared from the same batches of genomic DNA used for array CGH (aCGH) according to published methods (5–7). All of the enzymes used in the preparation were from New England Biolabs, and the oligonucleotides were synthesized by Eurofins MWG Operon. Oligonucleotide sequences (ATTGCG; GATCTG; TCAAGT; GTAGCC; TACAAG; CACCTG; GCC-TAA; TGGTCA; CACTGT) were added to the 3’ end of the Illumina adaptors used for the paired-end library preparation to serve as barcodes so that multiple samples could be sequenced in the same reaction. Sequencing was performed on an Illumina GAIIx Sequencer using Sequence Control software v2.6.26. Sequences were processed with either Bustard.py (OLB 1.6.0) or GERALD.pl (CASAVA 1.6.0), downloaded from the Illumina Web site (http://www.illumina.com/software/genome_analyzer_software.ilmn), with a produced read length of 41 nucleotides.

FASTQ sequence files were split into sample sets according to a barcode strategy and then barcodes were trimmed from the sequences using the applications fastx_barcode_splitter.pl and fastx_trimmer from the FASTX-Toolkit-0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit) package. Illumina sequence qualities were converted to Sanger sequence qualities using the MAQ application fq_all2std.pl (http://maq.sourceforge.net/qual.shtml) (8). Sequences were then aligned to chromosomes in the zebrafish Zv8/danRer6 assembly using bowtie 0.5,7 (9). Unassembled “scaffold” and “NA” fragments were not included in the alignment target. Alignments with quality scores of 10 or greater were extracted from the resulting Sequence Alignment/Map files and counts were obtained for the number of reads aligning to consecutive, adjacent 100-kb physical windows along the genome. Per-window counts per sample were rescaled to a total of 1 million reads, and log2 ratios were calculated for tumor/normal pairs for all 100-kb windows (or set to 0, if the read count for one or both samples was 0). The log2 ratios were then submitted to the circular binary segmentation algorithm in the same manner as described for the aCGH data (see above). Sequencing data are available at http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE23666.

Southern Blot Analysis. Probes for Southern blots were ≈1 kb in length; fragments were amplified by PCR and subcloned into pBlueScriptIIXS+ (Stratagene) or SP72 (Promega) before labeling. Sequences of the primers used to amplify the probe fragments are as follows:

Chr 13: AATGGCATCATGTTCAAGAC and TCTAGCCC- TGCCACTGAATA
Chr 14: GGGAGAGAGGTCTGACAAGT and ACTTCTCT- TGTGCGCATGT
Chr 15: CTCGCAATGCTACCCACAT and CGTTTCA- TGTGTGTCATGT

Zhang et al. www.pnas.org/cgi/content/short/1011548107
Chr 16: GGGCAAGATCACCACTCATA and ATCACCAA-
CAGCCACATCTC
Chr 25S: GGGACAAAGGAGAAGAAC and GCCCA-
TTTCTAGGTTTCC
Chr 25M: CTTCTGCAATGCAAGGAAAC and AGCTTCA-
CGAACCAAGACAA
Chr 25T: TCAGCAAACGCTTGTATT and GATTGT-
GTATGCCAGAAT.

Hybridization signals were quantitated on a phosphoimager using Imagequant software (Molecular Dynamics). For each probe, a ratio was determined between the signal in the tumor and the signal in the tail. The average of this ratio for the “neutral” probes (Chr 13, Chr 14, and Chr 16 probes) was then used to normalize the ratio for all of the probes for each tumor/tail pair.


Fig. S1. Cell-cycle properties of zebrafish malignant peripheral nerve sheath tumor (MPNST) cells. Results are shown for cells from three different MPNSTs from different rp heterozygotes: rpl7 (A–C), rps8 (D–F), and rpl36 (G–I). (A, D, and G) DNA content distribution by FACS analysis of propidium iodide (PI)-labeled cells. On the x axis, 200 represents the normal 2N position, which was calibrated by running a mixture of normal and tumor cells. (B, E, and H) FACS analysis of cells double-stained with PI (x axis) and antibody to pH3 (y axis). The mitotic cells are pH3-positive, and are circled in red. The percentages of the cell populations are indicated beside the circles. (C, F, and I) S-phase labeling with a 30-min BrdU pulse. The positions of major horseshoe shapes indicate that the majority of proliferating cells are aneuploid. (J) Histological sections from the rpl35 tumor were stained with antibodies to pH3. Mitotically active tumor cells were stained brown. (K) Tumor cells from rpl35 were double-stained with DAPI and anti-γ-tubulin. The multiple centrosomes are indicated by red arrowheads.
Fig. S2. Southern blot analysis confirms that chromosome 15 is often underrepresented and chromosome 25 is often overrepresented in zebrafish MPNSTs. Southern blots were performed with DNA from tail and tumor tissue from 14 rp heterozygotes. Signals for probes to chromosomes 13, 14, 15, 16, and 25 (T; Fig. 3B) were quantified, and a tumor/tail ratio was determined for each probe. The values for the chromosome 13, 14, and 16 probes were used to normalize the values for the chromosome 15 and 25 probes.

Fig. S3. Southern blot analysis confirms that chromosomes 13 and 14 do not vary in copy number consistently across zebrafish MPNSTs. Southern blots were performed with DNA from tail and tumor tissue from 14 rp heterozygotes. Quantitation of the normalized tumor/tail ratios for chromosome 13 and 14 probes was performed from the same Southern blots used in Fig. S2.

Fig. S4. Southern blot analysis confirms subchromosomal amplifications on chromosome 25 in zebrafish MPNSTs. The same Southern blots described in S2 were additionally probed with sequences from different positions on chromosome 25 (Fig. 3B) and analyzed as in Fig. S2. For each tumor, the ratio of the value for probe Sc or probe M to probe T is shown. Seven of 14 tumors show greater amplification for the Sc probe and/or the M probe than the T probe.
Other Supporting Information Files

Dataset S1 (XLSX)