Synteny and candidate gene prediction using an anchored linkage map of *Astyanax mexicanus*

Joshua B. Gross[a], Meredith Protas[a,1], Melissa Conrad[b], Paul E. Scheid[b], Oriol Vidal[b,c], William R. Jeffery[d], Richard Borowsky[b], and Clifford J. Tabin[a,2]

[a]Department of Genetics, Harvard Medical School, Boston, MA 02115; [b]Cave Biology Research Group, Department of Biology, New York University, New York, NY 10003; [c]Departament de Biologia, Universitat de Girona, Girona, Spain; and [d]Department of Biology, University of Maryland, College Park, MD 20742

Edited by Eric H. Davidson, California Institute of Technology, Pasadena, CA, and approved August 25, 2008 (received for review July 2, 2008)

The blind Mexican cave tetra, *Astyanax mexicanus*, is a unique model system for the study of parallelism and the evolution of cave-adapted traits. Understanding the genetic basis for these traits has recently become feasible thanks to production of a genome-wide linkage map and quantitative trait association analyses. The selection of suitable candidate genes controlling quantitative traits remains challenging, however, in the absence of a physical genome. Here, we describe the integration of multiple linkage maps generated in four separate crosses between surface, cave, and hybrid forms of *A. mexicanus*. We performed exhaustive BLAST analyses of genomic markers populating this integrated map against sequenced genomes of numerous taxa, ranging from yeast to amniotes. We found the largest number of identified sequences (228), with the most expect (E) values $<10^{-5}$ (95), in the zebrafish *Danio rerio*. The most significant hits were assembled into an "anchored" linkage map with *Danio*, revealing numerous regions of conserved synteny, many of which are shared across critical regions of identified quantitative trait loci (QTL). Using this anchored map, we predicted the positions of 21 test genes on the integrated linkage map and verified that 18 of these are found in locations homologous to their chromosomal positions in *D. rerio*. The anchored map allowed the identification of four candidate genes for QTL relating to rib number and eye size. The map we have generated will greatly accelerate the production of viable lists of additional candidate genes involved in the development and evolution of cave-specific traits in *A. mexicanus*.

sequence homology | physical genome | evolution | quantitative trait locus

The blind Mexican cavefish, *Astyanax mexicanus*, is a classic example of an organism evolving unique constructive and regressive traits in a novel environment and is an emerging model system for the genetic study of evolutionary and developmental biology (1). Since the Pleistocene, this species has invaded an extensive network of limestone caves located over a 140-km span of northeast Mexico giving rise to many cave populations (2). Strikingly, extant populations of the surface-dwelling ancestor of these cave forms persist in the rivers and streams surrounding the cave network. Cave forms of *Astyanax* were first formally described in 1936 (3) and gained increasing attention shortly thereafter, when it was shown that the surface and cave morphotypes were capable of producing viable hybrids. Cave × surface form crosses, and trait distribution analyses led to the first classical descriptions of the genetic architecture of several cave-specific traits (4).

Among the most conspicuous traits studied in this system include so-called "regressive" traits, i.e., phenotypic traits lost over the course of evolution in the cave ecosystem (4). Although loss of eyes and pigmentation is one of the most salient of cave-adapted characteristics, many additional morphological traits are affected in cave forms, including decreased size of the optic tectum, regression of the pineal gland, vertebrae, scales, and decreased metabolism and circadian activity (reviewed in ref. 5). Along with regressive traits, several "constructive" traits have been selected for in the cave environment. Included among these are increases in the size of the lower jaw, the telencephalon, the olfactory bulb, increased number of maxillary teeth (6), taste buds, cranial neuromasts, and increased egg yolk and adult fat content (reviewed in ref. 5).

Comparative studies between surface and cave morphotypes revealed an expansion of the cavefish taste system through increased number and position of peripheral taste buds from the labial region to the epithelium surrounding the maxilla, lower jaw and ventral head (reviewed in ref. 7). Similarly, cave forms have a more sensitive lateral line, a proprioceptive system consisting of canals throughout the body of the adult fish, populated with numerous mechanosensory neuromasts and ampullary organs (8). A number of behavioral differences have also been described between the cave and surface morphotypes, including modified feeding orientation (4), loss of schooling behavior (9), aggressive behavior (10), and sleep (11) in cave forms. Assuming these behaviors carry a genetic basis, each should yield a quantifiable phenotype that will reveal the genetic architecture of the trait and provide inroads to the gene(s) responsible.

The first genomic resources have recently been established for *Astyanax* in the form of a genome-wide linkage map (12). This linkage map, based on recombination frequencies of hundreds of microsatellite and candidate gene markers, has helped illuminate key aspects of *Astyanax* biology (13, 14). The first trait identified using quantitative trait locus (QTL) analyses in *Astyanax* was the genetic basis of albinism, shown to be a consequence of independent genetic lesions affecting the gene *ocular* and *cutaneous albinism type II* (*Oca2*) in two different cave populations (12). The underlying genetic basis of more complex phenotypes (e.g., supraoccipital bone morphology) involving the participation of several (potentially uncharacterized) genes is daunting in the absence of a sequenced physical genome, which is currently unavailable for *Astyanax*.

Herein, we report an integrated linkage map of *A. mexicanus*. We generated this map through comparative recombination frequency from pedigrees of four separate cave × surface crosses. The cave forms analyzed in this study include the Pachón (F_2 cross), Tinaja (two F_2 crosses) and Molino caves (backcross). We BLAST searched the genomic sequences flanking our microsatellites and candidate genes, and localized homologous sequences to specific

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Gene Networks in Animal Development and Evolution," held February 15-16, 2008, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS web site at: http://www.nasonline.org/SACKLER_Gene_Networks.

Author contributions: J.B.G. designed research; J.B.G., M.P., M.C., P.E.S., O.V., and R.B. performed research; J.B.G., M.P., O.V., W.R.J., and R.B. contributed new reagents/analytic tools; J.B.G., M.P., R.B., and C.J.T. analyzed data; and J.B.G., M.P., W.R.J., R.B., and C.J.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Department of Integrative Biology, University of California, Berkeley, CA 94720.

[2]To whom correspondence should be addressed. E-mail: tabin@genetics.med.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806238105/DCSupplemental.

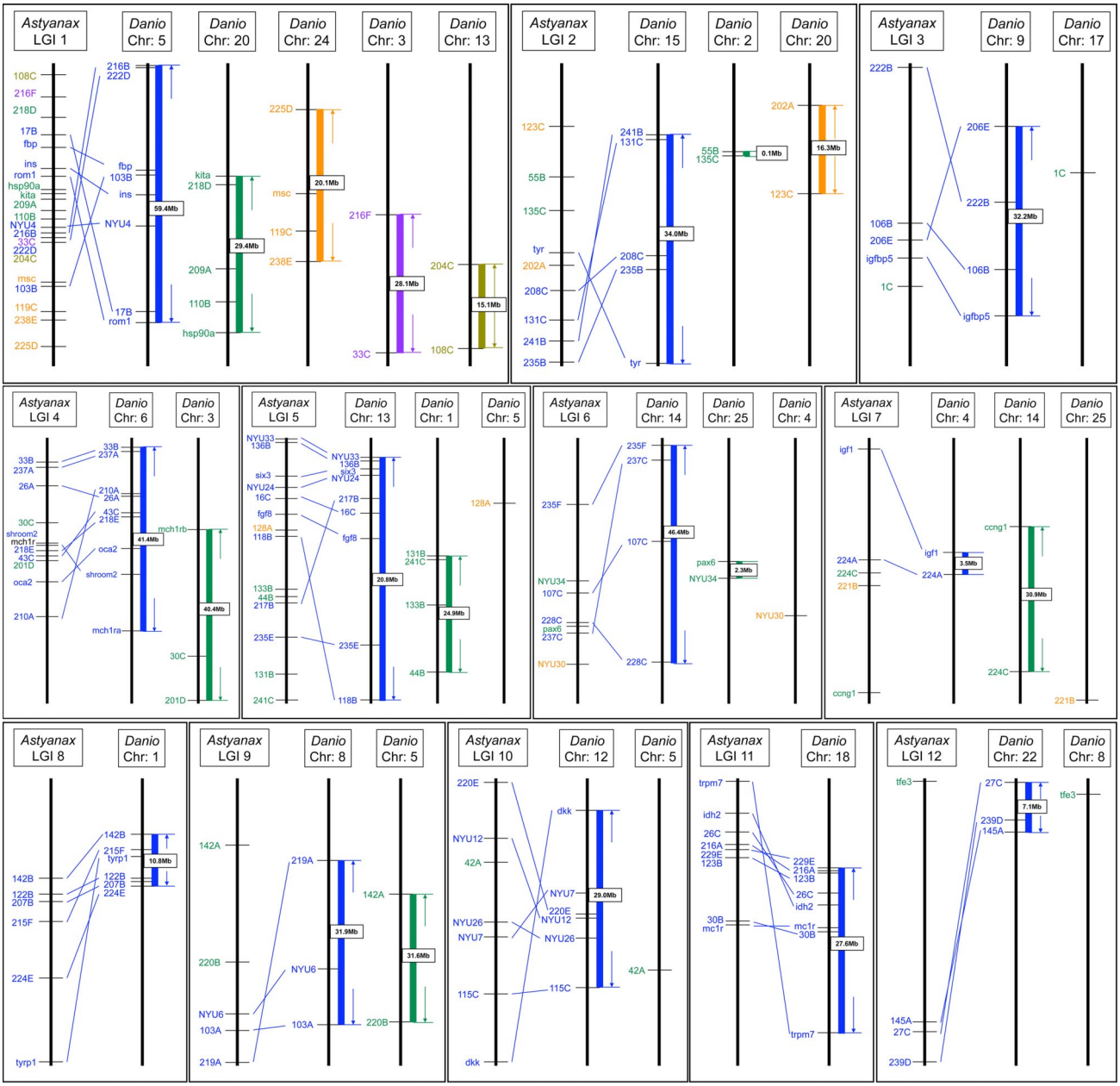

Fig. 1. Schematic diagram of the relative positions of anchored genomic sequences of *Astyanax* LGI 1–12 to chromosomes in *Danio*. *Astyanax* LGI 1 is anchored to *Danio* chromosomes 5, 20, 24, 3 and 13; LGI 2 is anchored to *Danio* chromosomes 15, 2, 20; LGI 3 is anchored to *Danio* chromosomes 9 and 17; LGI 4 is anchored to *Danio* chromosomes 6 and 3; LGI 5 is anchored to *Danio* chromosomes 13, 1, 5; LGI 6 is anchored to *Danio* chromosomes 14, 25, 4; LGI 7 is anchored to *Danio* chromosomes 4, 14, 25; LGI 8 anchors exclusively to *Danio* chromosome 1; LGI 9 is anchored to *Danio* chromosomes 8 and 5; LGI 10 is anchored to *Danio* chromosomes 12 and 5; *Astyanax* LGI 11 anchors exclusively to *Danio* chromosome 18; and LGI 12 is anchored to *Danio* chromosomes 22 and 8. Markers are color-coded, denoting their relative positions in each integrated *Astyanax* linkage group. In each image, chromosomes in *Danio* are shown with the highest number of hits first (blue), second highest number of hits (green), third highest (orange), fourth highest (purple) and fifth highest (brown). The largest predicted distance (Mb) spanned by respective anchored genomic markers in the *Danio* genome depicted within each color-coded region (white box). Note: schematic diagrams of chromosomes and linkage groups are not drawn to scale. Markers are indicated at their relative positions along each chromosome or linkage group.

chromosomes of numerous vertebrate and invertebrate taxa. The highest number of syntenic blocks are present in the most recent phylogenetic relative, *Danio rerio* [supporting information (SI) Figs. S1*A* and S2*A*; ref. 15], with gradually less identified sequences in more distantly related genomes (Fig. S1 *B–D*, Table S1). Using an "anchored" linkage map built from this data (Figs. 1 and 2), we predicted the positions of numerous candidate genes based on their relative positions in *D. rerio*. The predictive power afforded by this resource will greatly accelerate the search for candidate genes involved in the many intriguing traits that have expanded or regressed in our emerging model system, *A. mexicanus*.

Results and Discussion

An Anchored *Astyanax* Linkage Map. We created an integrated genetic map from the microsatellite analysis of crosses involving three different cave populations. This map contains 28 linkage groups. There are 25 chromosomes in *A. mexicanus* (12), suggesting that the map will further collapse as additional markers are added bridging currently distinct linkage groups. Nonetheless, this map provides a powerful genetic resource for the analysis of evolutionary traits in *Astyanax*. In the absence of a physical map of the *Astyanax* genome, we anchored the linkage map to the sequenced

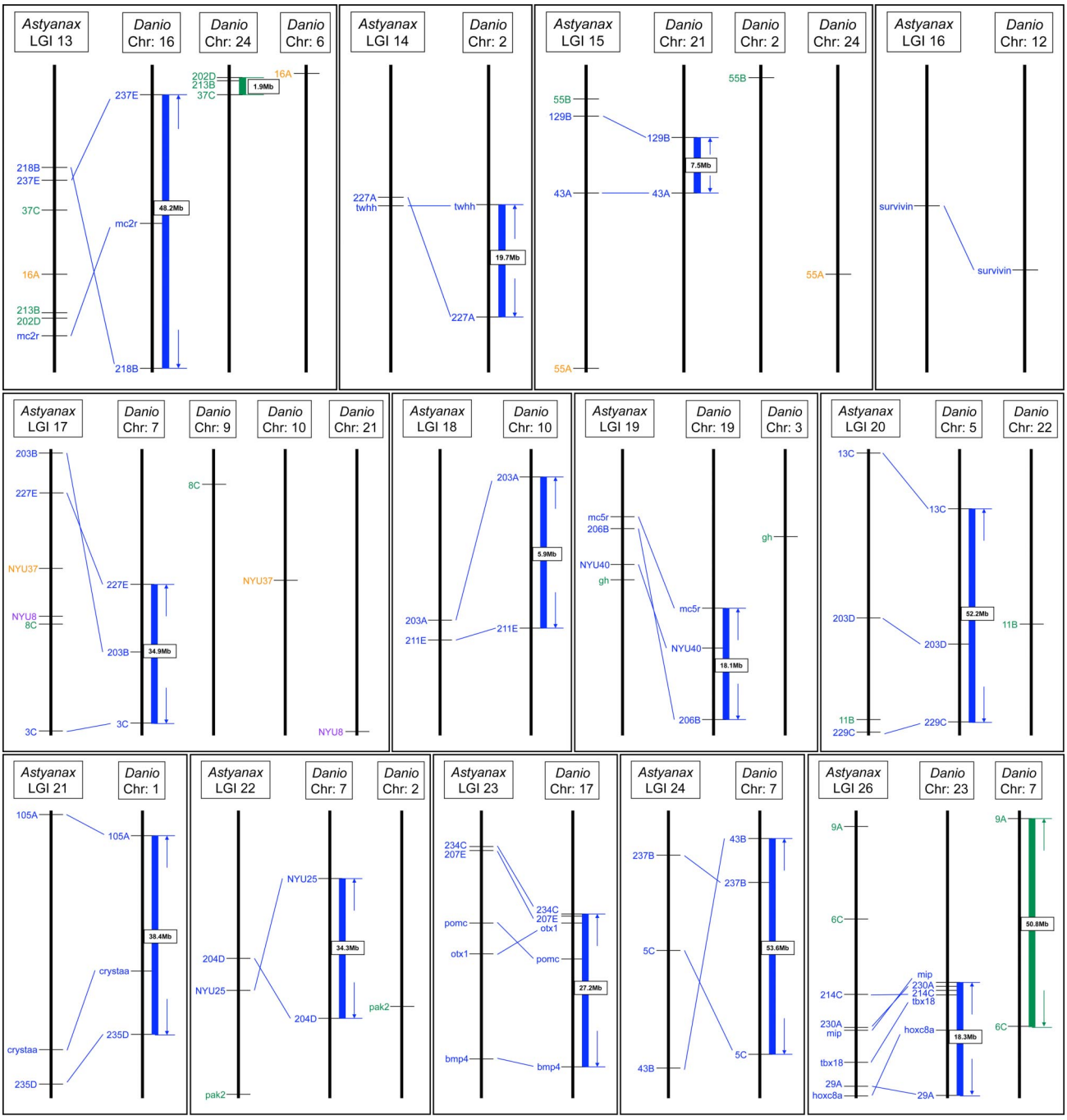

Fig. 2. Schematic diagram of the relative positions of anchored genomic sequences of *Astyanax* LGI 13–24 and 26 to chromosomes in *Danio*. *Astyanax* LGI 13 is anchored to *Danio* chromosomes 16, 24, 6; LGI 14 anchors exclusively to *Danio* chromosome 2; LGI 15 is anchored to *Danio* chromosomes 21, 2, 24; LGI 16 anchors exclusively to *Danio* chromosomes 12; LGI 17 is anchored to *Danio* chromosomes 7, 9, 10, 21; LGI 18 is anchored exclusively to *Danio* chromosomes 10; (*C*) LGI 19 is anchored to *Danio* chromosomes 19 and 3; LGI 20 is anchored to *Danio* chromosomes 5 and 22; LGI 21 anchors exclusively to *Danio* chromosomes 1; LGI 22 is anchored to *Danio* chromosomes 7 and 2; (*C*) LGI 23 anchors exclusively to *Danio* chromosomes 17; LGI 24 anchors exclusively to *Danio* chromosome 7; and LGI 26 is anchored to *Danio* chromosomes 23 and 7. Markers are color-coded, denoting their relative positions in each integrated *Astyanax* linkage group. In each panel, chromosomes in *Danio* are shown with the highest number of hits first (blue), second highest number of hits (green), third highest (orange) and fourth highest (purple). The largest predicted distance (Mb) spanned by respective anchored genomic markers in the *Danio* genome depicted within each color-coded region (white box). Note: schematic diagrams of chromosomes and linkage groups are not drawn to scale. Markers are indicated at their relative positions along each chromosome or linkage group.

genomes of other teleost species based on the conserved positions of numerous genomic sequences surrounding microsatellites and genomic polymorphisms associated with a select number of cloned genes (Fig. S2*B*). The highest number of syntenic hits was returned from sequence analyses of the *D. rerio* genome, the species with a physical map that is phylogenetically closest to *A. mexicanus* (Fig. S2*A*). Each of the 155 hits is charted in an Oxford grid (Fig. S1*A*) depicting the relative position of each hit in *Astyanax* and *Danio*. The majority of syntenic hits from a given *Astyanax* integrated linkage group (LGI) are distributed in a mosaic pattern among several different chromosomes (Chr) in *Danio* (Figs. 1 and 2). This would be expected given the extensive genomic rearrangements

that have occurred among even closely related species over the course of genomic evolution. Nonetheless, for several linkage groups (*Astyanax* LGI 8, 11, 14, 18, 21, 23, 24), we found that every syntenic hit resides on a single *Danio* chromosome (*Danio* chromosomes 1, 18, 2, 10, 1, 17, 7, respectively).

None of the nine microsatellite sequences from *Astyanax* LGI 25, 27, and 28 returned any hits from any search algorithms of the *Danio* genome (Fig. S1*A*; gray). Two of these linkage groups (27 and 28) were generated from only one of the four crosses (surface × Molino cave cross) and may represent unintegrated portions of one of the other linkage groups in *Astyanax*. Thus, they may appear as independent linkage groups but are not collapsed into their appropriate group. Linkage group 25 is also a small linkage group (14.1 cM) and therefore may not carry any markers with strongly enough conserved sequences to be identified by current sequence search algorithms.

A single *Danio* chromosome, 11, is not represented by any current *Astyanax* sequences in this analysis (Fig. S1*A*, red). Because the strongest and most reliable sequence "hits" between *Astyanax* and *Danio* appear to be markers near candidate genes, one would expect that the addition of candidate genes residing on *Danio* chromosome 11 would allow positions within this chromosome to be identified and anchored to the *Astyanax* linkage map.

Significant Regions of Comparative Synteny Are Observed in Three Additional Teleost Species: *Gasterosteus, Oryzias,* and *Tetraodon*. We extended our genomic sequence analyses to additional teleost species, including *Gasterosteus* (sticklebacks), *Oryzias* (medaka), and *Tetraodon* (pufferfish). We did not evaluate *Takifugu rubripes* (another pufferfish species), despite extensive sequence information, because at present the genomic data for this species has not been assembled into chromosomes. Each of the three species analyzed using an Oxford grid (Fig. S1*B–D*) demonstrated a lower number of syntenic hits than *D. rerio*. The number of syntenic hits identified in our Oxford grid analysis for *Gasterosteus* were 106, compared with 113 in *Oryzias* and 94 in *Tetraodon*. The decrease in synteny demonstrated in this analysis is not surprising given the increased phylogenetic distance between these organisms and *Astyanax* (Fig. S2*A*).

Similar to *Danio*, *Astyanax* linkage groups 25, 27, and 28 returned no hits in any of these genomes (Fig. S1*B–D*). Additionally, the microsatellites residing on linkage groups 9 and 14 returned no hits in *Gasterosteus* (Fig. S1*B*, gray). Every chromosome in this species, however, was recognized in our analysis. In contrast, the genomic markers in four linkage groups were unrecognized by the current draft of the *Oryzias* genome (Fig. S1*C*, gray). Chromosome 11 in this species was unrepresented by any hits from *Astyanax* markers (Fig. S1*C*, red). Finally, our sequence analysis was least successful in the current draft of the *Tetraodon* genome with five linkage groups (15, 17, 25, 27, 28) returning no hits from this species (Fig. S1*D*, gray). Three *Tetraodon* chromosomes (4, 20, 21) remained uncharacterized in our sequence analysis (Fig. S1*D*, red).

To determine whether the same regions of synteny between *Astyanax* and *Danio* could be observed in other teleosts, we compared the genomic region surrounding the gene *Oca2*, in all five species (Fig. S3 and Table S2). We found the highest number of genomic markers was recognized in *Danio*, with eight markers falling within a ≈19.8-mB interval (Fig. S3, *Danio* Chr:6). Several of the same markers, however, fell within similar-sized genomic regions of *Gasterosteus* (Chr:VIII), *Oryzias* (Chr:4), and *Tetraodon* (Chr:1) chromosomes (Fig. S3, red). Further, three of the microsatellites not anchored in our analysis of the *Danio* genome colocalized with other hits from linkage group 4 in other species (Fig. S3, green boxes). This would suggest that these microsatellites (NYU2, 213E, 132B) have diverged or become lost in the *Danio* genome but retain their identities and syntenic positions in other teleost lineages. The position of microsatellite marker 43C falls on the same chromosome as *Oca2* in *Danio*, *Oryzias*, and *Tetraodon* implying a primitive syntenic relationship between this genomic marker and *Oca2*. The position of the gene *Oca2* has not yet been annotated in *Gasterosteus*, however, the positions of flanking genomic markers (including 43C and *Shroom2*) would suggest that this gene resides on chromosome VIII in this fish genome.

Gene Position in *Astyanax* Is Reliably Predicted by Homologous Gene Position in *D. rerio*. Although several genomic markers were identified exclusively in *D. rerio*, our analyses in additional teleost species, amphibian, amniotes and other genomes indicated that certain markers were widely recognized across distant taxa (Table S1). Therefore, we selected the top nine hits that were identified in every teleost genome with an "Expect" cutoff value of 10^{-10} (see *SI Methods*). Each of these markers was identified with a very strong match and high level of sequence similarity (Table S3). We examined these sequences to determine their identity. In every case, these clones were in gene sequences, including (with the exception of 230A) partial coding sequence flanking the intronic $CA_{(N)}$ microsatellite. This analysis demonstrates that coding regions are optimal genomic markers for defining syntenic positions across widely different taxa.

This was further confirmed by an analysis of the position of test genes taken from the anchored linkage map. Based on the anchored map, we chose 21 genes from the *Danio* genetic map, predicted to be linked to specific QTL on the *Astyanax* map. We cloned intronic or neighboring regions of these genes (including genes accessioned to public databases) to identify informative size-length polymorphisms or SNPs. Using primers designed to amplify these polymorphic regions (Table S4), we placed numerous candidate genes on our linkage map.

In 18 of 21 cases, we found the predicted position of the test gene was supported by at least one neighboring genomic marker. In some cases, the genomic marker was another gene (e.g., Fig. 2, LGI 26). In other cases, the neighboring markers were microsatellites (e.g., Fig. 1, LGI 10). Many of these genomic markers span the critical region(s) of QTL identified in our phenotypic analyses (14). Thus, the addition of genes to the linkage map can be used directly to seek associations between a gene of interest and a particular trait. Further, in the absence of a sequenced physical genome, the addition of candidate genes to our linkage map is an essential step toward increasing the resolution of our syntenic analysis with the genome of *D. rerio* and the search for genes that define evolutionarily relevant traits.

The Relative Position of *Oca2* in *Astyanax* and *Danio* Offers a Proof-of-Principle for Candidate Gene Placement Predictions. At present, the only gene responsible for a QTL so far identified in *Astyanax* is the trait of albinism (12). The gene *Oca2* was shown to control this trait in multiple cave populations via different lesions. This gene has also been shown to cause albinism in humans and mice. To test the reliability of our strategy of defining syntenic regions using *Astyanax* genomic markers, we anchored the microsatellites and candidate genes surrounding the critical region of this QTL to the *Danio* genome (Fig. 1, LGI 4). Indeed, we found that eight of these markers colocalized to chromosome 6 in *Danio rerio*, near the homologous position of *Oca2*. Additionally, four genomic markers colocalize to chromosome 4 in *Oryzias* and two genomic markers colocalize to chromosome 1 in *Tetraodon*, where *Oca2* resides in these species (Fig. S3). This result implies the presence of a primitive syntenic block inclusive of the gene *Oca2* and surrounding genomic markers that remains conserved in the genomes of *Astyanax*, *Danio*, *Oryzias*, and *Tetraodon*.

This gene, however, is only one of many potential genes capable of causing an albinism phenotype. Therefore, we asked whether this analysis would have predicted the participation of any other known genes involved in albinism. In humans, four forms of ocular and cutaneous albinism have been described and characterized, including: OCA1 (*tyrosinase, tyr*; ref. 16), OCA2 (*Oca2*; ref. 17), OCA3

(*tyrosinase related protein I*; *tyrp1*; ref. 18) and OCA4 (*slc45a2*; ref. 19). None of the three other gene candidates, *tyr* (Fig. 1, LGI 2; *Danio* chr:15), *tyrp1* (Fig. 1, LGI 8; *Danio* chr:1) or *slc45a2* (predicted on *Danio* chr:21), were found to reside on either chromosome 3 or 6 in *D. rerio*. This analysis further supports *Oca2* as the gene controlling albinism as reflected in the comparative syntenic data provided by our flanking marker analyses. Further, this analysis would have narrowed our list of candidate genes controlling albinism, had these data been available at an earlier stage of the project.

Application of an Anchored Linkage Map to the Study of Cave-Specific Traits in *A. mexicanus*. The blind Mexican cave tetra, *A. mexicanus*, is typical of troglobitic forms in exhibiting numerous traits related to cave life. Many different cave animals from around the globe seem to converge on similar phenotypes as a result of colonizing the darkness of the cave microenvironment. For example, a Bahamian cave crustacean, *Speleonectes lucayensis*, completely lacks any remnant of an eye or body pigmentation (20). The European blind cave salamander, *Proteus anguinus*, has evolved a decreased basal rate of metabolism and activity and loss of circadian rhythms (21). Subterranean water beetles demonstrate an increased rate of mutation in the eye pigment gene, *cinnabar* (22). Thus, despite distant phylogenetic ancestry of these taxa, the same or similar phenotypes appear as a recurring theme for obligate cave-dwelling animals.

Of the various cave-dwelling species, *Astyanax* provides the best system for the study of the genetic basis for cave-adapted traits, in particular because the ancestral surface-dwelling morph is known and is still interfertile with the cave morph. The variable presence of particular traits (e.g., pigmentation, eye size) among multiple independently derived populations of *Astyanax* cave morphs makes it an excellent model for the evolution of these traits in parallel. Additionally, once particular genes have been established for traits of interest (e.g., *Oca2* and albinism), we can extend our analyses globally to other obligate cave-dwelling populations sharing these traits that are not themselves directly accessible to genetic investigation. Thus, we can determine whether the same or different genes are involved in convergent phenotypes across much more distantly related taxa. Ultimately, this will inform whether particular genes are optimal targets for mutation, or rather whether convergent phenotypes are selected for in the cave environment by different genes. The anchored linkage map described here will be an invaluable tool to aid in narrowing the prediction of candidate genes for cave-adapted traits in *Astyanax* until a physical map is available.

Identification of Candidate Genes for Rib Number and Reduction in Eye Size Through Analysis of an Anchored Linkage Map of *Astyanax*. Based on success this approach would have had in identifying a candidate gene for albinism, we turned our attention to two other traits for which no candidate genes were previously known. Previous analyses have discovered numerous QTL for trait differences segregating between surface and cave populations of *Astyanax*, including eye size and rib number (14). We therefore asked whether the anchored map would provide potential candidate genes for these QTL.

Among the morphological differences between surface and cavefish, the latter exhibit a decrease in rib number. A single major QTL was identified for this trait in the Pachón cave. Several microsatellites flanking this locus on LGI 26 (Fig. 2) anchor strongly to chromosome 23 in *Danio*. Strikingly, the *Hoxca* locus lies on the syntenic region of chromosome 23. We therefore added the gene *Hoxc8a* to the *Astyanax* linkage map and found that it indeed maps near the rib number QTL. This is a promising candidate gene for the evolution of rib axial variability in *Astyanax* given the role described for paralogous group 8 *Hox* genes in mouse rib patterning (23). Further, the expression of the anterior limit of expression of *Hoxc8a* in *Danio* corresponds with the thoracic-lumbar transition (24), the region of variable morphology (i.e., rib number) found in cave forms of *Astyanax*. Determining whether mutations in this, or a linked, *Hox* gene are indeed responsible for the variation in rib number requires further analyses. It is, however, an exciting possibility. Although mutations in *Hox* genes are known to affect segment identity and vertebral type, including between thoracic and lumbar, and although changes in domains of *Hox* gene expression are known to correlate with evolutionary shifts in vertebral identity, specific mutations in *Hox* genes have not yet been linked to evolutionary shifts in vertebral formulae.

Eye loss remains one of the most intriguing aspects of cave adaptation and the F_2 pedigrees of surface $\times$ Pachón crosses reveal that at least six QTL are responsible for the development of this trait. One significant QTL for this trait resides on LGI 21 that anchors strongly to chromosome 1 in *Danio* (Fig. 2). Examination of the *Danio* genome indicates that the antiapoptotic chaperone encoded by the gene, *Crystaa* is predicted to be tightly linked to this QTL. This is notable because *Crystaa* has been shown to be down-regulated during eye degeneration in *Astyanax* (25). Previously, it was unclear whether this was a direct or indirect effect. Our analysis suggests that regulatory mutations affecting the expression *Crystaa* are likely one of the six genetic factors affecting eye development in blind cave tetras.

A second QTL for eye size resides on LGI 1 in *Astyanax* (Fig. 1). Using the anchored linkage map, we find that in *Danio* the gene *Rom1* is predicted to be tightly linked to this trait. *Rom1* is a structural protein expressed specifically in the rod outer segment, maintaining the flattened shape of outer segment disks (26). Mutations in this gene are responsible for inherited forms of macular degeneration and retinitis pigmentosa in humans, and the classical "retinal degeneration slow" (RDS) mouse mutant (26). Thus, although at present *Rom1* can be viewed only as a candidate for this trait, similar coding mutations may cause parallel eye size alterations in populations of *Astyanax*.

Finally, a third QTL for eye size maps very close to the *Oca2* gene and the albinism trait in *Astyanax*. The gene *Shroom2* is linked to *Oca2* in other teleosts (Fig. S3). This was intriguing, because *Shroom2* plays an essential role in regulating the biogenesis and localization of melanin in the developing retina (27). Further, the *Shroom2* locus lies within a crucial genomic region of two forms of ocular albinism in humans, possibly contributing to visual system disorders. We therefore placed *Shroom2* on the *Asytanax* linkage map and found that it indeed has conserved tight linkage to *Oca2*, as predicted on the basis of synteny to other teleosts. Thus, alterations to *Shroom2* function in *Astyanax* may similarly affect normal melanin localization, mediating one of the eye development QTL (LGI 4; Fig. 1) in cave populations of *Astyanax*.

Putative Duplication Event Revealed Through Syntenic Analyses of *Astyanax* and *Danio*. One linkage group in *Astyanax*, LGI 4, anchors very strongly to two chromosomes in *Danio*, chromosomes 6 and 3 (Fig. 1). Although the majority of genomic sequences flanking the microsatellites populating this linkage group are colocalized to chromosome 6, two microsatellites anchored strongly to chromosome 3 (30C and 201D). A candidate gene later added to our map, *melanin concentrating hormone receptor 1* (Fig. S2*B*, *mch1r*), was positioned within an overlapping region of the markers on this linkage group. In *Danio*, this gene has two paralogues: *mch1ra* and *mch1rb*; whereas in *Astyanax*, we have so far identified only a single member. Interestingly, *mch1ra* is localized in *Danio* to chromosome 6, whereas *mch1rb* is localized to chromosome 3. This observation suggests that the markers present on linkage group 4 represent a "primitive" chromosome (or segment therein) that was duplicated in the *Danio* lineage but not in *Astyanax*.

Although the majority of the markers retain their identity on chromosome 6, some of the duplicate paralogous members have possibly been lost or accumulated mutations and therefore are unrecognized in the Ensembl search algorithm. Presumably their nonmutated paralogues are retained (30C and 201D) and recognized on chromosome 3, along with the duplicated gene *mch1r*. This

example of duplication is interesting in light of the hypothesis that a genome duplication event occurred in ray-finned fish before the divergence of *Danio* and *Astyanax* (6). Our data may suggest that the putative duplication of *mch1r* in *Danio* occurred after the divergence of these two species. Alternatively, the gene duplication of *mch1r* may have occurred before the divergence of the two species, and a paralogue was subsequently lost or become unrecognized in the lineage leading to *Astyanax*.

This example, although informative regarding the chromosomal evolution in the species we are examining, also highlights a potential problem we need to keep in mind in using the approach described here. Some candidate genes may map to a predicted position based on reference to the anchored linkage map but may be misleading because of the placement of a duplicated (paralogous) gene in one species that is not present in the reference genome. A second way the anchored linkage map can yield misleading information is when there are multiple regions of synteny with other teleost fish in the same region of the genome, that are not recognized as such because of the still crude nature of our genetic map. Nonetheless, the value of this approach is seen in the examples described above, and initially misleading candidate proposals can generally be sorted out as additional markers around the gene of interest are isolated and analyzed.

Materials and Methods

Microsatellite Isolation and Linkage Map Construction. Our first published linkage map (12) was generated by using recombination frequencies of CA-dinucleotide repeat microsatellites, given that they are widely common and distributed frequently throughout the genomes of teleost fish. A microsatellite library was created by using genomic DNA isolated from a surface morph of *A. mexicanus* and digested to completion with the restriction enzyme *Sau3AI*. Fragments between 500 and 700 bp in size were subcloned into BamHI-linearized pBluescriptSK(+). Ligated clones were electroporated into SURE electrocompetent cells (Stratagene) and probed by using either P^{32} or digoxigenin end-labeled $(CA)_{12}$ oligonucleotides. Positive clones were sequenced with T3 and T7, and primers were designed around the putative polymorphic regions of the clones, as described (12). Primers were designed only to microsatellites containing 10 or more CA repeats using the web-based primer design tool, Primer3 (http://fokker.wi.mit.edu/primer3/input.htm).

We obtained >600 microsatellites from this genomic library. Primers were initially tested in a genotypic analysis of our original two crosses: an F_2 Pachón cave × surface cross and a Molino cave × surface backcross. The Pachón F_2 cross contained 539 individuals genotyped with 259 markers; the Molino backcross contained 111 individuals genotyped with 322 markers. In addition, we used a subset of the same primers to type two additional Tinaja Cave × surface F_2 crosses: Tina110 (190 individuals with 103 markers), and Tina111 (285 individuals with 207 markers).

Because the three cave populations are isolated from one another, the loci segregating in the various crosses comprised overlapping subsets of the total. Because all of the founding populations are outbred, each cross exhibited a mixture of segregation types. JoinMap 3.0 (Kyazma) is capable of mapping populations with mixed segregation and its map integration function was used to combine linkage groups from between two and four crosses, where appropriate, to produce the integrated map.

To integrate the map, we created a JoinMap project with separate data nodes for the four different mapping crosses. Within each of the crosses, groups of loci based on cosegregation (lod 5.0) were created. These groups were then inspected manually and united by content among crosses as follows: we started with a linkage group in the Molino cross because it had the largest number of markers. We then searched for the presence of each marker from the Molino group in the groups of the other three crosses to identify corresponding linkage groups. We then inspected each marker in the corresponding groups to determine whether they would conscript new groups from the other crosses. If so, new markers in the new groups were checked, in turn. Typically, between two and four other groups were united with the original Molino group by this process. Related groups from different crosses identified in this way were then integrated by JoinMap into groups which were then mapped by using default settings. The resulting integrated map had 28 linkage groups with 400 markers and a length of 1,783 cM.

The most current draft of the *Astyanax* linkage map is coarse owing to the moderate number of genomic markers currently used in linkage analyses. Future efforts are focused on increasing the resolution of the map through the addition of hundreds of additional markers. A higher resolution map will collapse the number of linkage groups to the proper of chromosomes (i.e., 25) in *Astyanax* and clarify potentially problematic areas of the current anchored linkage map (e.g., interrupted synteny, microinversions, etc.).

BLAST Search Protocol. Each of 326 genomic sequences flanking the microsatellites housed in our genomic library of informative markers was BLAST searched against a variety of physical genomes. All searches were carried out by using the Ensembl Genome Browser (www.ensembl.org/index.html) using the BlastN search engine with Search Sensitivity set to ''No Optimisation.'' To return hits quickly, the number of returned hits was set to a limit of 10; the E value cutoff was set to the default of 10.

A variety of physical genomes were used in this analysis, including the following teleost fish species: *D. rerio* (Assembly 47), Fugu (Assembly FUGU 4.0), Medaka (Assembly HdrR), Stickleback (Assembly BROAD S1), *Tetraodon* (Assembly TETRAODON 8); the amphibian *Xenopus* (Assembly JGI 4.1); amniote species: Chicken (Assembly WASHUC2), Mouse (Assembly National Center for Biotechnology Information m37), Human (Assembly National Center for Biotechnology Information 36); the urochordate *Ciona* (Assembly JGI 2); invertebrate species: *Drosophila* (Assembly BDGP 4.3), *C. elegans* (Assembly WS 180); and the yeast, *Saccharomyces* (Assembly SGD1.01).

ACKNOWLEDGMENTS. This work was supported by a collaborative grant from the National Science Foundation (IOS-0821982) (to C.J.T. and R.B.) and a grant from the National Institutes of Health (1RO3EYE016783, to R.B.). Postdoctoral fellowship support was provided to J.B.G. through a Kirschstein National Institutes of Health–National Research Service Award (1F32 GM081439-01).

1. Jeffery WR (2008) Emerging model systems in evo-devo: Cavefish and microevolution of development. *Evol Dev* 10:265–272.
2. Wilkens H, Strecker U (2003) Convergent evolution of the cavefish. *Astyanax (Characidae, Teleostei)*: Genetic evidence from reduced eye-size and pigmentation. *Biol J Linn Soc Lond* 80:545–554.
3. Hubbs CL, Innes WT (1936) The first known blind fish of the family Characidae: A new genus from Mexico. *Occas Pap Mus Zool Univ Mich* 342:1–7.
4. Wilkens H (1988) in *Evolutionary Biology*, eds Hecht MK, Wallace B (Plenum, New York), pp 271–367.
5. Jeffery WR (2001) Cavefish as a model system in evolutionary developmental biology. *Dev Biol* 231:1–12.
6. Stock DW, Jackman WR, Trapani J (2006) Developmental genetic mechanisms of evolutionary tooth loss in cypriniform fishes. *Development* 133:3127–3137.
7. Boudriot F, Reutter K (2001) Ultrastructure of the taste buds in the blind cave fish. *Astyanax jordani* (''*Anoptichthys*'') and the sighted river fish *Astyanax mexicanus* (*Teleostei, Characidae*). *J Comp Neurol* 434:428–444.
8. Teyke T (1990) Morphological differences in neuromasts of the blind cavefish. *Astyanax hubbsi* and the sighted river fish *Astyanax mexicanus*. *Brain Behav Evol* 35:23–30.
9. Parzefall J, Fricke D (1991) Alarm reaction and schooling behaviour in population hybrids of. *Astyanax fasciatus* (*Pisces, Characidae*). *Mem Biospelol* 18:29–32.
10. Burchards H, Dölle A, Parzefall AJ (1985) Aggressive behaviour of an epigean population of *Astyanax mexicanus* (*Characidae*, *Pisces*) and some observations of three subterranean populations. *Behav Proc* 11:225–235.
11. Kavanau JL (2005) Evolutionary approaches to understanding sleep. *Sleep Med Rev* 9:141–152.
12. Protas ME, *et al.* (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat Genet* 38:107–111.
13. Protas M, Conrad M, Gross JB, Tabin C, Borowsky R (2007) Regressive evolution in the Mexican cave tetra, *Astyanax mexicanus. Curr Biol* 17:452–454.
14. Protas M, *et al.* (2008) Multi-trait evolution in a cave fish, *Astyanax mexicanus. Evol Dev* 10:196–209.
15. Fink SV, Fink WL (1996) in *Interrelationships of Fishes*, eds Stiassny MLJ, Parenti LR, Johnson GD (Academic, New York), pp 209–249.
16. Oetting WS (2000) The tyrosinase gene and oculocutaneous albinism type 1 (OCA1): A model for understanding the molecular biology of melanin formation. *Pigment Cell Res* 13:320–325.
17. Brilliant MH (2001) The mouse p (pink-eyed dilution) and human *P* genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. *Pigment Cell Res* 14:86–93.
18. Boissy RE, *et al.* (1996) Mutation in and lack of expression of tyrosinase-related protein-1 (*TRP-1*) in melanocytes from an individual with brown oculocutaneous albinism: A new subtype of albinism classified as ''OCA3.'' *Am J Hum Genet* 58:1145–1156.
19. Newton JM, *et al.* (2001) Mutations in the human orthologue of the mouse. *underwhite* gene (*uw*) underlie a new form of oculocutaneous albinism, OCA4. *Am J Hum Genet* 69:981–988.
20. Yager J (1981) Remipedia, a new class of crustacea from a marine cave in the Bahamas. *J Crustac Biol* 1:328–333.
21. Hervant F, Mathieu J, Durand JP (2000) Metabolism and circadian rhythms of the European blind cave salamander *Proteus anguinus* and a facultative cave dweller, the Pyrenean newt (*Euproctus asper*). *Can J Zool* 78:1427–1432.
22. Leys R, Cooper SJ, Strecker U, Wilkens H (2005) Regressive evolution of an eye pigment gene in independently evolved eyeless subterranean diving beetles. *Biol Lett* 1:496–499.
23. van den Akker E, *et al.* (2001) Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development* 128:1911–1921.
24. Prince VE, Joly L, Ekker M, Ho RK (1998) Zebrafish. *hox* genes: Genomic organization and modified collinear expression patterns in the trunk. *Development* 125:407–420.
25. Strickler AG, Byerly MS, Jeffery WR (2007) Lens gene expression analysis reveals down-regulation of the anti-apoptotic chaperone αA-crystallin during cavefish eye degeneration. *Dev Genes Evol* 217:771–782.
26. Keen TJ, Inglehearn CF (1996) Mutations and polymorphisms in the human peripherin-*RDS* gene and their involvement in inherited retinal degeneration. *Hum Mutat* 8:297–303.
27. Fairbank PD, *et al.* (2006) Shroom2 (APXL) regulates melanosome biogenesis and localization in the retinal pigment epithelium. *Development* 133:4109–4118.