SUPPLEMENTAL MATERIALS

METHODS

Target Locus Selection

In order to increase capture efficiency, probes were designed in regions of the vertebrate genome that had two properties: conservation and uniqueness. To ensure probe conservation, the multiz46way vertebrate alignment was profiled using a sliding window of 120bp (the probe length, see below) after reducing the alignment to include only the five model species. Two conservation scores were computed for each nucleotide site, an indel score (number of gaps within ±60bp of the nucleotide site) and an identity score (average percent sequence similarity across species). In order to ensure probe uniqueness, we profiled sequences from the same reduced alignment using a sliding window of 15bp. For each nucleotide site, we computed a 15mer copy-number score as the number of times the 15mer existed in the relevant genome, maximized across the five species (higher score indicates less conserved). Similarly, a 60mer copy-number score was also computed as the occurrence of each observed 60mer when a 60bp sliding window was used. We disqualified sites for several reasons. (1) We disqualified sites in which the indel score was greater than five. This requirement ensured that probes would hybridize efficiently. (2) We disqualified sites in which the 15mer copy-number score was greater than 2000 at any site within 700 base pairs of the site. This action reduced the potential of captured repetitive elements to act as probes, resulting in chaining of off-target fragments, The threshold of 2000 was chosen because it retained the best ~50% of the loci. (3) We disqualified sites in which the 60mer copy-number score was greater than 10 within 120bp of the site. This

action was performed to reduce the potential for probes to hybridize off-target. Candidate probe regions ~120bp in length were then chosen from the remaining sites by eye using plots of the identity score and conservation score as a function of nucleotide site. Chosen regions contained high identity scores but were immediately flanked by regions of low identity scores on at least one side. Chosen regions also contained low copy-number scores, although increasing copy-number scores were tolerated with increasing distance from the candidate probe regions. In a final step to ensure uniqueness, candidate probe region sequences from each of the five species were searched for in the relevant genome (Blast 2.2.23+, Camacho et al. 2009), and candidates with greater than one significant match (e-value < 0.00001) for any species were removed from the list. A total of 512 final probe regions remained.

DNA extraction and library preparation

Where necessary, genomic DNA was extracted from tissue samples using the E.Z.N.A.® Tissue DNA Kit (Omega Bio-Tek , Norcross, GA; Supplemental Table 1). Quantification of DNA concentration was performed with the Qubit™ fluorometer (Invitrogen, Carlsbad, CA). For all samples, 1 µg of genomic DNA in 50 µl of elution buffer (QIAGEN) was placed into a 1.7 mL microcentrifuge tube and subjected to a 2 min cycle of sonication consisting of alternating 30 sec on-off pulses in a model UCD-200TM-EX Bioruptor *(*Diagenode, Sparta, NJ). A 5µl aliquot of each sample was electrophoresed in a gel composed of 1% Bio-Rad ultra agarose in 1X TAE for 90 min at 105V to visualize the effectiveness of sonication. The observed size range of the fragments was 150–1100 bp. The remainder of each sonicated sample was used to prepare barcoded libraries according to the protocol of Meyer and Kircher (2010)

with the following modifications: no addition of Tween 20 was made to the SPRI (Agilent) bead suspension, the 20 min SPRI bead dry time was eliminated, and the indexing PCR cycling program was replaced with a 2-step program consisting of 1 cycle of 30 sec at 98°C, 12 cycles of 10 sec at 98°C followed by 20 sec at 72°C, 1 cycle of 5 min at 72°C, then hold at 4°C. Three separate barcoded libraries were prepared from each sonicated genomic DNA sample using different barcodes for each one in order to allow subsequent pooling of multiple species libraries with different insert sizes (see below).

Size selection, enrichment, and sequencing

Libraries were size-selected to produce, for each species, one indexed library containing inserts of 375bp, a second indexed library containing inserts of 575bp, and a third indexed library containing inserts of 775bp. Size selected libraries were prepared by electrophoresing the indexed libraries as described above but with a run time of 2.5 hrs, followed by band excision with gel cutting pipette tips, and gel purification using commercial kits. Samples were loaded into gels in alternate lanes to avoid possible cross contamination. In order to eliminate UV and ethidium bromide exposure to the libraries, size markers were loaded into the two wells at the edges of each gel; the outermost ladder lane from each side was subsequently removed and stained with EtBr, marked at the desired band location, and then replaced in its original position alongside the remainder of the gel to serve as a reference for band excision. Once the desired bands were removed for purification (using 1.1 x 6.5 mm gel tips), the gel was EtBr stained and photographed for verification of successful selection at either 500, 700 or 900 bp. Purification kit (QIAGEN or E.Z.N.A.®) columns were eluted twice with 30 µl 70°C molecular grade H_2O.

Prior to sequence capture, each library was amplified with primers IS5 and IS6 using the indexing PCR cycling program of Meyer and Kircher (2010) but with an annealing temperature of 68°C. PCR clean-up was performed with a QIAquick (QIAGEN) kit and columns were eluted with H_2O as previously described.

Libraries of each insert size were pooled across species, corresponding to the three insert sizes. Each multi-species pool was captured separately using the previously described probe mix (which was obtained via an Agilent Custom SureSelect kit containing a single pool of all probes for all species) and the SureSelect protocol for multiplexed sequencing (Agilent, 2010) with one exception: during hybridization, the SureSelect Indexing Block #3 was replaced with a 50 µM indexed blocking reagent prepared by mixing equal volumes of 300 µM solutions of adapter and index oligos according to recommendations provided by SEQanswers.com (unpub. protocol: High-Throughput Indexed Library Preparation and Pooled Agilent Exome Enrichment For Illumina sequencing platform, Schellenberg Lab, University of Pennsylvania School of Medicine).

To compare the efficiency and quality of data produced by current Illumina instruments in order to assess their utility for future anchored enrichment experiments, next-generation sequencing was performed on both the Illumina MiSeq Personal Sequencing System (375bp insert size multi-species pool only) and the Illumina HiSeq 2000 sequencing system (all three multi-species pools were combined in a single lane). In order to ensure equal coverage across the 375, 575, and 775 insert size libraries, captured libraries were pooled at 12%, 30%, and 58%, respectively. MiSeq sequencing was performed at the Chesterton Research Park Illumina sequencing facility in the United Kingdom, whereas the HiSeq 2000 sequencing was performed at the Illumina sequencing facility in Hayward, CA. Paired-end 150bp reads were produced on

the MiSeq, whereas paired-end 100bp reads were produced on the HiSeq 2000. Both sequencing runs utilized the v3 TrueSeq PE Cluster kit and included a separate 8bp indexing read (index sequences are given in Supplemental Table 2). The HiSeq reads were processed by the CASAVA v1.8 pipeline, whereas the MiSeq reads were processed by the MiSeq Reporter software.

Initial read processing

Reads received from the genome centers were first filtered for quality following the Illumina recommendations contained within the raw read files (MiSeq Reporter; HiSeq: CASAVA v1.8 pipeline; chastity threshold = 0.6). After quality filtering, reads were demultiplexed (sorted into separate files) using the 8bp indexes. Reads for which the index sequence did not match exactly to one of the 30 expected indexes were removed. Recall that each index sequence differed by a minimum of two base pairs from all other index sequences used. Two-fold degeneracy of indexes is necessary to avoid an excessive number of misidentified reads (Meyer and Kircher 2010).

Since the sequencing libraries contained some proportion of short fragments (~20% for HiSeq reads and ~35% for MiSeq reads), we developed an approach to identify and merge paired reads for which the two sequencing reads overlapped (i.e., reads derived from fragments with length less than 200 bp and 300 bp for HiSeq and MiSeq reads, respectively). Using a merging script written by ARL, we performed the following steps for each pair: (1) slide read 1 and the reverse-complement of read 2 past each other and for each position, compute the probability that the number of observed matches would occur by chance. An equal base frequency assumption was utilized to substantially reduce computation time. (2) If the minimum probability score

(across all possible overlap positions) is less than 10E-10 for any position, the reads were merged with a degree of overlap corresponding to the lowest probability score. When conflicts in base calls were evident the base call from the read with the higher quality score for that base was used. New quality scores were computed as the posterior probability that the base called is incorrect following Rodrigue et al. (2010). The Merger program is described in more detail in Rokyta et al. (in review).

We also screened the reads for PCR duplicates, which are replicate fragments sequenced in separate clusters but generated from the same initial library fragment.
PCR duplicates can be generated when a very small number of DNA fragments (e.g. the pool of post-enrichment library fragments) is amplified by PCR to yield a final library with low diversity relative to the number of sequencing reads. Since our probes were designed to capture a small percentage of the genome (0.02% to 0.09% depending on genome size) and a small amount of each library is used during the capture process (50ng), there exists the potential for PCR duplicates to exist in the set of sequencing reads. The issue arises when the number of fragments from the target regions drop so low (e.g., during capture) that sampling error occurs when the captured DNA is re-amplified prior to sequencing. Before proceeding with the assembly, we identified PCR duplicates by concatenating the two reads from each pair and identifying the unique set of sequences for all concatenated pairs (accounting for the reverse ordering of the reads). The sequences and quality scores of exactly one pair of each type was written to a second file. A similar procedure was performed for merged reads. Note that this method, while it removes the majority of PCR duplicates, is not perfect since sequencing error causes the sequence of two PCR duplicates to differ.

Read assembly

We developed a quasi-*de novo* approach to the read assembly since we desired comparable results across species and mapping to whole genomes was not possible for all 10 species. The quasi-*de novo* approach involved three steps: (1) mapping reads to probe region sequences, (2) gathering unmapped mates into the assembly, and (3) refining read positions within the assembly. For step (1), reads corresponding to each read type were mapped separately in SeqMan NGen 3 (DNASTAR, Inc.) to each of the five reference sequence sets (40 read types * 5 references = 200 assemblies). The 40 read types refer to the 30 libraries sequenced on the HiSeq 2000 plus the 10 libraries sequenced on the MiSeq. Default assembly parameters for the initial reference assembly were used except: MatchSize was increased to 25 and MinMatchPercent was lowered to 50 to improve mapping of nonmodels to divergence model references. Step (2) consisted of adding unmapped mates to the assembly (an unmapped mate is a read that was not mapped but that has a mate that was mapped). This step is necessary since a read will not be retained in the NGen assembly unless it overlaps with the probe region. All reads mapping to the probe region were joined into a single assembly for each locus by species combination. The optimal assembly for each locus (among the 200) was identified as the one in which the maximum number of reads mapped to the target regions in the NGen assembly.

Step (3) involved iteratively extending the optimal alignment. We began by fixing the relative position of reads mapping in the optimal assembly. All other reads began with unfixed positions. Positions of the unfixed reads were then optimized by (a) taking a consensus of fixed reads, (b) identifying for each read the position maximizing the proportion of bases matching the consensus sequence, and (c) fixing the position of reads with ≥25bp of overlap with the

consensus and ≥90% of overlapping bases matching the consensus. Steps (a)-(c) were repeated until the number of fixed reads no longer increased between iterations (typically requiring 2-8 iterations). Reads with unfixed positions at the end of the process were removed from the assembly. Note that this method will only incorporate pairs for which one of the two reads extended sufficiently into the probe region (expectation: 100% for 375bp, 79% for 575bp, and 46% for 775bp insert sizes). This fact makes our estimates of enrichment, coverage, and locus length conservative. Long fragments binding to probes in the center of the fragment will not be represented in these assemblies. Although this procedure will affect coverage estimates across each locus, it is sufficient for producing long, high quality assemblies from which high quality consensus sequences can be obtained. We found the quasi-*de novo* approach to be better than a strict *de novo* approach (as implemented in NGen) since it requires at least one read in each pair to map to the single-copy probe region, thus reducing the number of assembly errors caused by repetitive elements. The extension approach works well because of the very low rate of indel sequencing errors for Illumina reads (<0.01%; Glenn 2011). True indels were accommodated during the manual alignment adjustment (see below).

Assembly inspection and refinement

Recognizing that all downstream analyses rely on accurate assemblies, we subjected all assemblies to several quality control measures. We first noticed that the number of assembled reads for each locus was bimodally distributed with a break at approximately 60 reads. Inspection of assemblies below the threshold indicated that these were poor assemblies often

containing a large proportion of possibly mis-indexed reads (as evidenced by a high degree of mismatches among mapped reads). Assemblies above the threshold appeared to be proper. Loci corresponding with assemblies containing fewer than 60 reads were removed from further analysis (deemed unusable for phylogenetic analysis).

Assemblies for each locus passing the threshold were treated with an automated assembly refinement step designed to remove obvious sequencing errors and aid in efficient manual inspection. The refinement began with the taking of a temporary 90% consensus sequence of mapped reads. Sites with <10 called bases were marked as N in the consensus. For each site with an unambiguous consensus base, we computed the probability that the number of mismatches to the consensus sequence base would occur under a sequencing-error-only model (one locus one allele) with an error rate of 1%. Sites with a p-value of >1% demonstrated sufficient evidence that mismatched were the result of sequencing error. Those mismatched base calls were converted to the consensus base. Manual inspection suggested that this procedure did not affect sites with evidence of a true polymorphism, excessive sequencing error, or assembly error.

All refined assemblies were then inspected by eye and classified into four categories: (0) Excellent, no manual adjustment required, (1) Very Good, minor manual adjustment required; (2) Good, moderate manual adjustment required, and (3) Poor, probably not possible to resolve errors. Typically, assemblies in category 1 contained a small number of low quality reads, those in category 2 contained an NGen assembly error involving large assembly sections that were misaligned, and those in category 3 had substantial evidence of gene duplication and/or alignment errors. In order to make manual inspection of 3796 assemblies possible, the work was divided across seven members of the Lemmon Labs. In order to minimize bias, assemblies were

divided by locus rather than by species (assemblies for all ten species for a given locus were evaluated by the same person). Assemblies were inspected using Geneious Pro (version 5.5.1, Drummond et al. 2011).

Assemblies were manually adjusted in order to ensure the quality of the final consensus sequences. These assemblies were adjusted in one or more of the following ways: regions misaligned by NGen in the initial mapping assembly reads were repaired, regions containing indels were improved, regions containing microsatellites were adjusted, and individual sequencing reads were trimmed or removed to eliminate obvious cases of sequencing error. Once these adjustments were made, the assemblies were considered final.

RESULTS

Sequence quantity and quality

Both the HiSeq and MiSeq sequencers produced a large quantity of high-quality sequence data. The MiSeq produced over 13 million 150bp-reads (6.6 million pairs, ~2 billion bp total), whereas the HiSeq 2000 produced over 31 million reads (>15 million pairs, 31 billion bp total) after quality filtering based on Illumina recommendations (4.6% of HiSeq and 6.7% of MiSeq reads were removed). Average reported quality scores (corresponding to percent accuracy) remained high across the length of the reads (Supplemental Figure 5) and were consistent across the two platforms. Read 1 was consistently higher quality than Read 2 on both platforms. Quality increased for the first 20 bases then declined for the remainder of the bases remaining above 29 (99.87% reported accuracy) for the 100bp HiSeq reads and above 24 (99.6% reported accuracy) for the MiSeq.

Indexing reads allowed the vast majority of reads to be unambiguously assigned to one of the 30 libraries (recall that the 8bp indexes differed by at least 2 bp). Reads not matching exactly to one of the 30 indexes were not utilized in downstream analyses (1.7% and 0.7% of reads passing quality filter for the HiSeq and MiSeq, respectively, were excluded).

A substantial portion of the sequenced fragments were short enough to allow overlapping reads. Libraries varied with respect to abundance of overlapping reads (4% to 38% for HiSeq and 7% to 73% for MiSeq). As indicated by Supplemental Figure 6a, cases in which reads overlap can be readily identified as long as a sufficient degree of overlap existed (a minimum of 20 bp of overlap is required to obtain significance using a high-stringency threshold of 10E-10). Merging overlapped reads is beneficial in that it produces longer, higher quality reads (Supplemental Figure 6b) and allows many sequencing errors to be corrected.

The frequency of PCR duplicates was modest and most pronounced for libraries pooled at higher concentrations in the final library pool. Between 13% and 26% of reads from the 735bp insert size libraries, for example, were removed due to evidence of PCR duplication (recall that all but one copy of each unique sequence pair were identified as a PCR duplicates and removed). Libraries with small insert sizes, which were present in lower concentrations in the final library pool, showed a lower incidence of PCR duplication (3% to 18%). The number of unique reads plotted as a function of the number analyzed is given in Supplemental Figure 7.

Assembly results

Initial reference assemblies in NGen produced some coverage for all loci for all ten samples. Each assembly typically required less than one hour and 16GB of RAM. In general, the greatest number of reads mapped to the reference sequence from the closest model species.

Exceptions did occur, however, especially for cases when coverage was low for a particular species-by-gene combination (e.g., for non-model species). Suspecting that the mapped reads in this case may be reads with misidentified indexes, we plotted the number of mapped reads as a function of locus for each species. This plot indicated a bimodal distribution in the number of mapped reads with a natural break at ~60 reads. Assemblies of loci with less than 60 reads contained a majority of reads that had a large degree of among-read mismatches compared to assemblies of loci with greater than 60 reads. Thus, loci with fewer than 60 reads mapped were considered not captured and were not analyzed further.

Extension of the initial assembly was also straightforward. Between two and eight extension iterations were typically required, with the number generally increasing with the number of reads initially mapped. In general, the distance between reads within a pair corresponded to the insert size of the relevant library. As expected, reads from short-insert libraries were clustered in the center of the assembled locus whereas reads from long-insert libraries were distributed throughout the assembled locus. Merged reads overlapped substantially with the probe region.

Manual inspection and adjustment of the assemblies was a time consuming but necessary step to ensure that occasional assembly artifacts did not result in error in downstream analyses. Manual inspection of each of the 3796 assembled loci required between 0.5 and 3 minutes totaling ~45 researcher-hours. The majority of loci with sufficient coverage required minor or no manual adjustments (46%, 22%, 6%, 0.5% and 0.26% scored 0, 1, 2, 3, and 4, respectively). A summary of the scores are given in Supplemental Table 1. The distribution of scores was surprisingly consistent across species, except for the fact that the non-model species had a substantially higher proportion of uncaptured loci. Manual adjustments of loci with a quality

score of 1 were simple, typically involving trimming or removing a small number of reads that were clearly outliers. Adjustments of loci ranking 2 were somewhat more involved, requiring correction of the alignment in regions containing an indel, homopolymer, or microsatellite. Only 9 of the 320 manually-adjusted assemblies could not be adequately adjusted and were thus not analyzed further. These problematic assemblies showed evidence of gene duplication as evidenced by incongruent patterns of SNPs.

REFERENCES

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T.L. 2009. BLAST+: architecture and applications. BMC Bioinform. 10:421.

Drummond A.J., Ashton B., Buxton S., Cheung M., Cooper A., Heled J., Kerse M., Moir R., Stones-Havas S., Sturrock S., Thierer T., Wilson A. 2010. Geneious v5.1. Available from http://www.geneious.com

Glenn T.C. 2011. Field guide to next-generation DNA sequencers. Mol.Ecol. Resources 11:759–769.

Meyer M., Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 2010; doi:10.1101/pdb.prot5448

Rodrigue S., Materna A.C., Timberlake S.C., Blackburn M.C., Malmstrom R.R., Alm E.J., Chisholm S.W. 2010. Unlocking Short Read Sequencing for Metagenomics. PLoS One 5:e11840

Roykta D.R., Lemmon A.R., Margres M.J.., Arnow K. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). BMC Genomics. *In Review*.

SUPPLEMENTAL FIGURE CAPTIONS

Supplemental Figure 1. Comparison of low-sensitivity (LS, i.e. the quasi-*de novo* approach) and high-sensitivity (HS) approaches of determining whether a locus was successfully captured (see Methods for a description of these approaches). Nine example loci are presented. The top row presents loci in which neither approach detected successful capture. The center row presents loci in which only the high-sensitivity approach detected successful capture. The bottom row presents loci in which both approaches detected successful capture. In each graph the left-most peak (top of peak not shown) containing the large majority of the reads represents the null distribution, with ~40 matches per read occurring by chance. For some of the loci, a second peak is present with the percent of matches per read exceeding 55 (the high-sensitivity threshold chosen). This second peak contains reads mapping to the locus represented by the panel.

Supplemental Figure 2. Bowtie mapping of reads to *Gallus* genome. Distribution of mapped *Gallus* reads across *Gallus* chromosomes and mitochondrion is shown. Reads were mapped to galGal3 using Bowtie with m=1 (only uniquely mapped reads were retained). Average coverage is taken across 10000bp segments. Circles indicate positions of target regions.

Supplemental Figure 3. Bowtie mapping of reads to *Danio* genome. Distribution of mapped *Danio* reads across *Danio* chromosomes and mitochondrion is shown. Reads were mapped to danRer6 using Bowtie with m=1 (only uniquely mapped reads were retained). Average coverage is taken across 10000bp segments. Circles indicate positions of target regions.

Supplemental Figure 4. Bowtie mapping of reads to *Mus* genome. Distribution of mapped *Mus* reads across *Mus* chromosomes and mitochondrion is shown. Reads were mapped to mm9 using Bowtie with m=1 (only uniquely mapped reads were retained). Average coverage is taken across 10000bp segments. Circles indicate positions of target regions.

Supplemental Figure 5. Read quality. Average Illumina reported quality scores are shown as a function of position in read. Paired-end sequencing was performed on the Illumina MiSeq and HiSeq 2000 platforms. Read 1 indicates the forward read and Read 2 the reverse read. The HiSeq produced 100bp paired-end reads whereas the MiSeq produced 150bp paired-end reads. Quality scores shown are the average across reads passing the Illumina quality filter. Quality scores of 40, 30, and 20 correspond to 99.99%, 99.9%, and 99% accuracy, respectively. HiSeq reads were processed with the CASAVA v1.8 pipeline, whereas MiSeq reads were processed with the MiSeq Reporter. Reported quality scores were in Illumina 1.8+ format.

Supplemental Figure 6. Increased read length and quality through merging. Each set of paired-end reads was evaluated for evidence that it originated from a fragment less than 200bp for HiSeq reads (100bp read in each direction) and 300bp for MiSeq reads (150bp read in each direction), as would be expected if read 1 and read 2 are overlapping. (a) For each potential fragment length, the probability of the number observed matches occurring by chance was computed and compared to a significance threshold (p = 10E-10). Pairs were merged into a single read when significance was observed (yellow point) and left paired if no significance was observed (gray points). New quality scores were computed for merged reads as the posterior probability that the new base call is correct (see Supplemental Materials). Panel (b) indicates how the average quality scores are enhanced for 150bp fragments. The blue and red lines indicate average quality scores for Reads 1 and 2, respectively. The light gray line indicates the

greatly increased quality scores of paired reads after merging. Note that the low-quality region becomes high quality when merged. This figure is based on Figure 1 of Rokyta et al. (in review), which describes more details of the Merger program used to merge the overlapping reads.

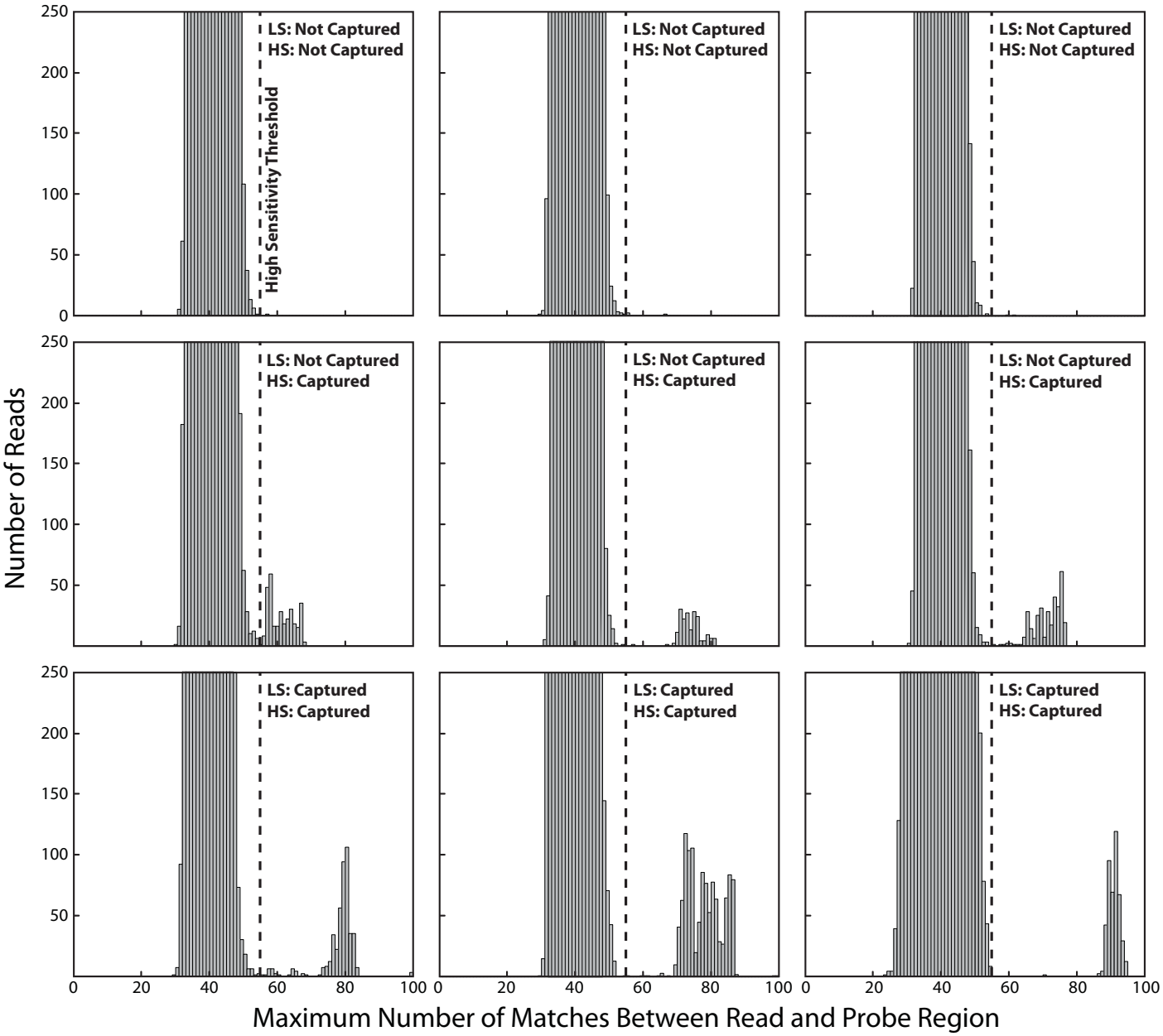
Supplemental Figure 7. PCR duplicates. The number of unique pairs is plotted as a function of the number of pairs analyzed. Each curve represents a different library. A paired sequencing read was considered to be a duplicate if the sequences in the pair were found more than once from the same library (see Supplemental Methods).

Supplemental Figure 8. Estimated tradeoff between the number of captured loci, the median length of captured loci, and the number of individuals pooled in one Illumina HiSeq sequencing lane (indicated as the number contained in each point). This figure is identical to Figure 7 in the main manuscript but shows the tradeoff expected when DNA is fragmented using the 450bp Covaris protocol. Genome size (GS) and divergence time with nearest model (DT) are shown for each species.

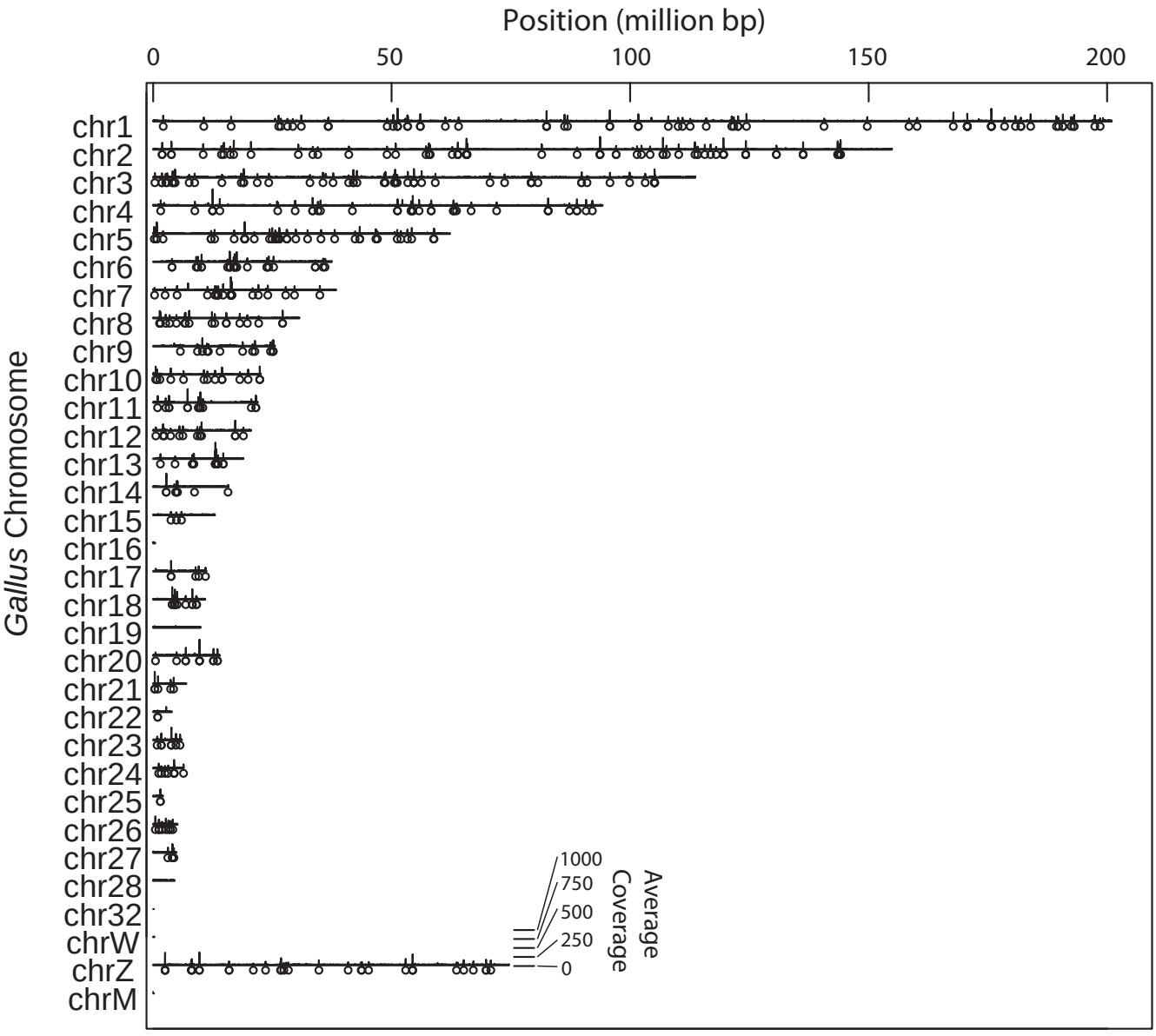
Supplemental Figure 9. Estimated tradeoff between the number of captured loci, the median length of captured loci, and the number of individuals pooled in one Illumina HiSeq sequencing lane (indicated as the number contained in each point). This figure is identical to Figure 7 in the main manuscript but shows the tradeoff expected when DNA fragments of approximately 775bp are size-selected using a gel. Genome size (GS) and divergence time with nearest model (DT) are shown for each species.

Supplemental Figure 10. Estimated tradeoff between the number of captured loci, the median length of captured loci, and the number of individuals pooled in one sequencing lane (indicated as the number contained in each point). This figure is identical to Figure 7 in the main manuscript but shows the tradeoff expected when the library is prepared using the Nextera sample preparation kit (which produces fragments approximately 300bp in length) and sequenced on the Illumina MiSeq personal genome sequencer with paired 150bp reads. Genome size (GS) and divergence time with nearest model (DT) are shown for each species.

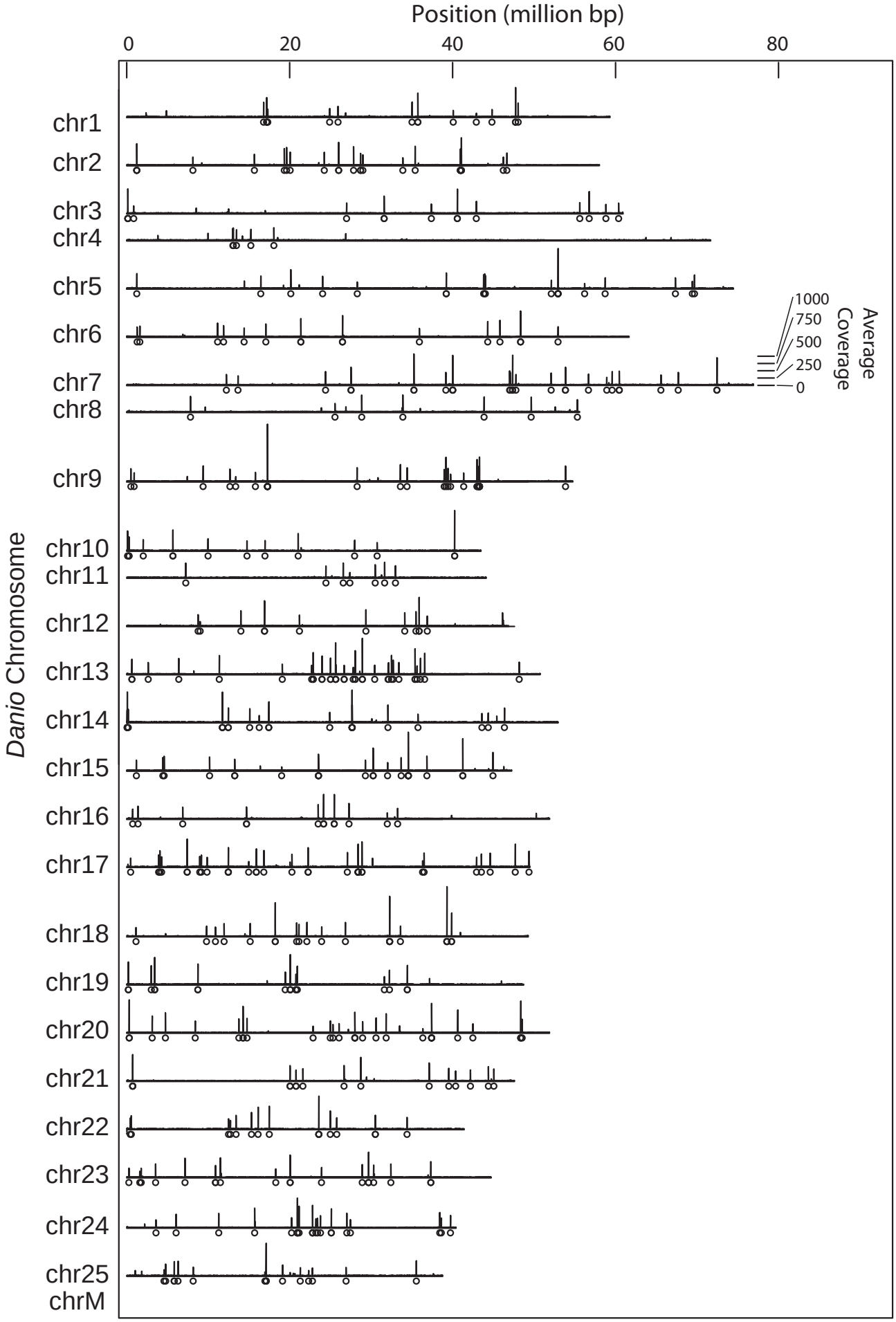
Supplemental Figure 1



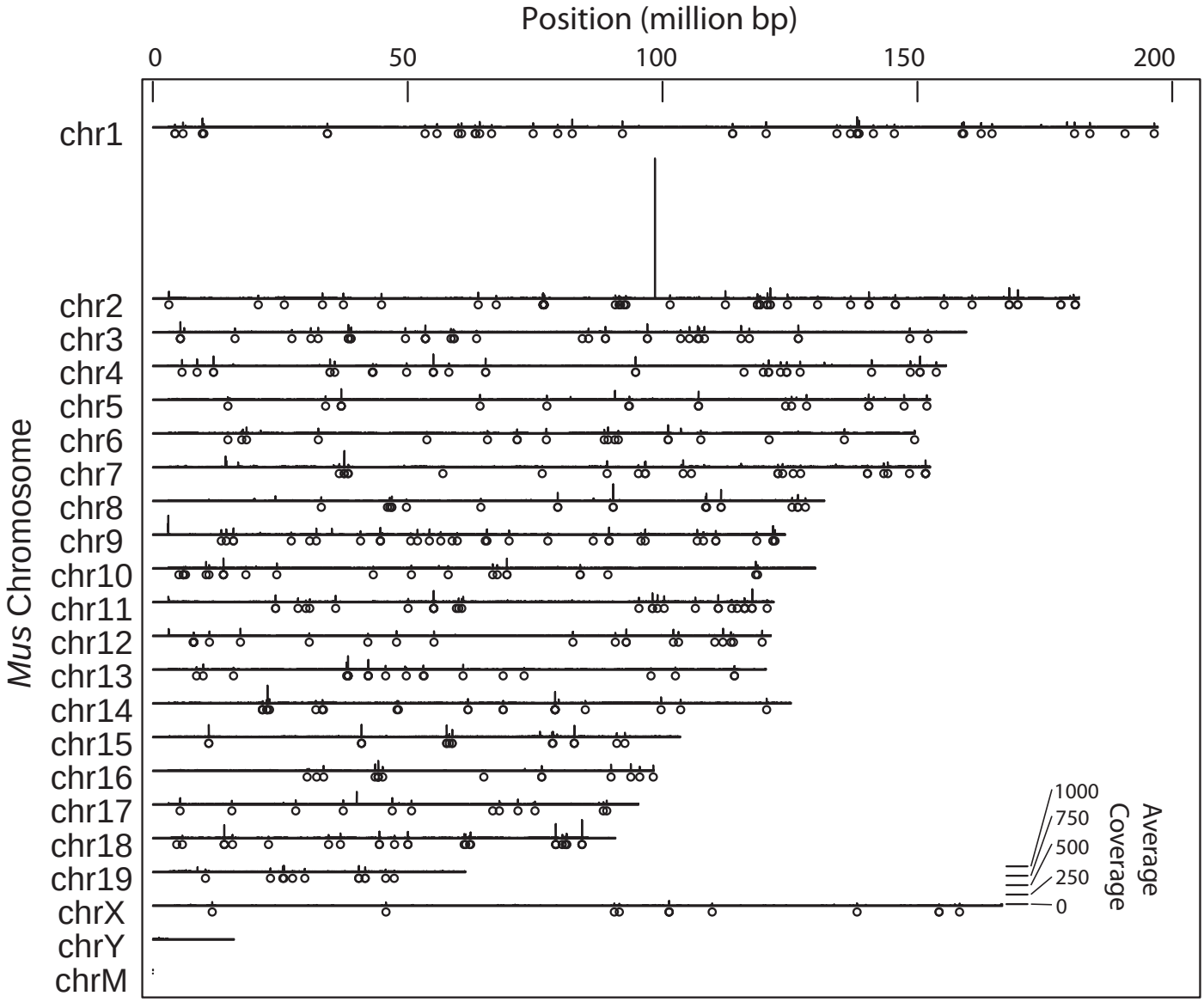
Supplemental Figure 2



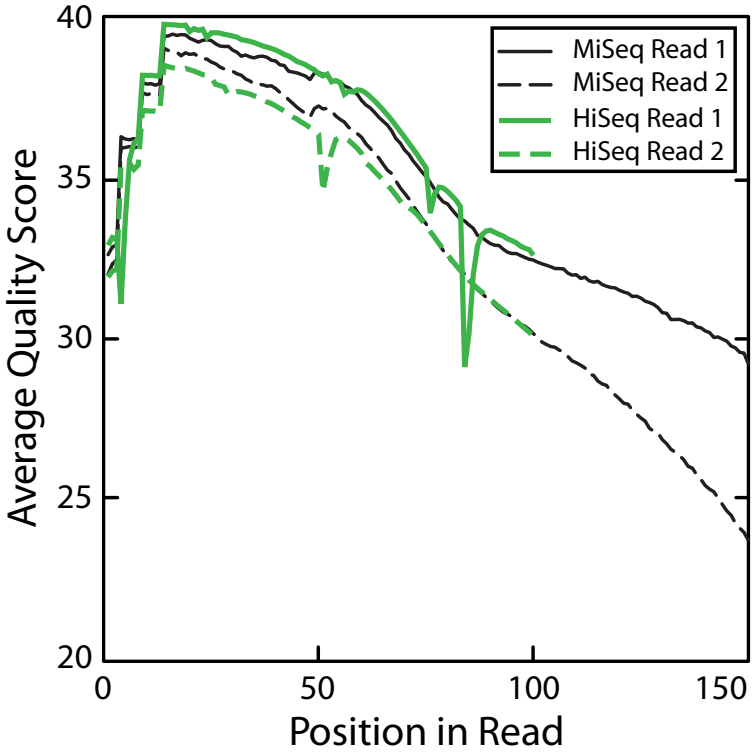
Supplemental Figure 3



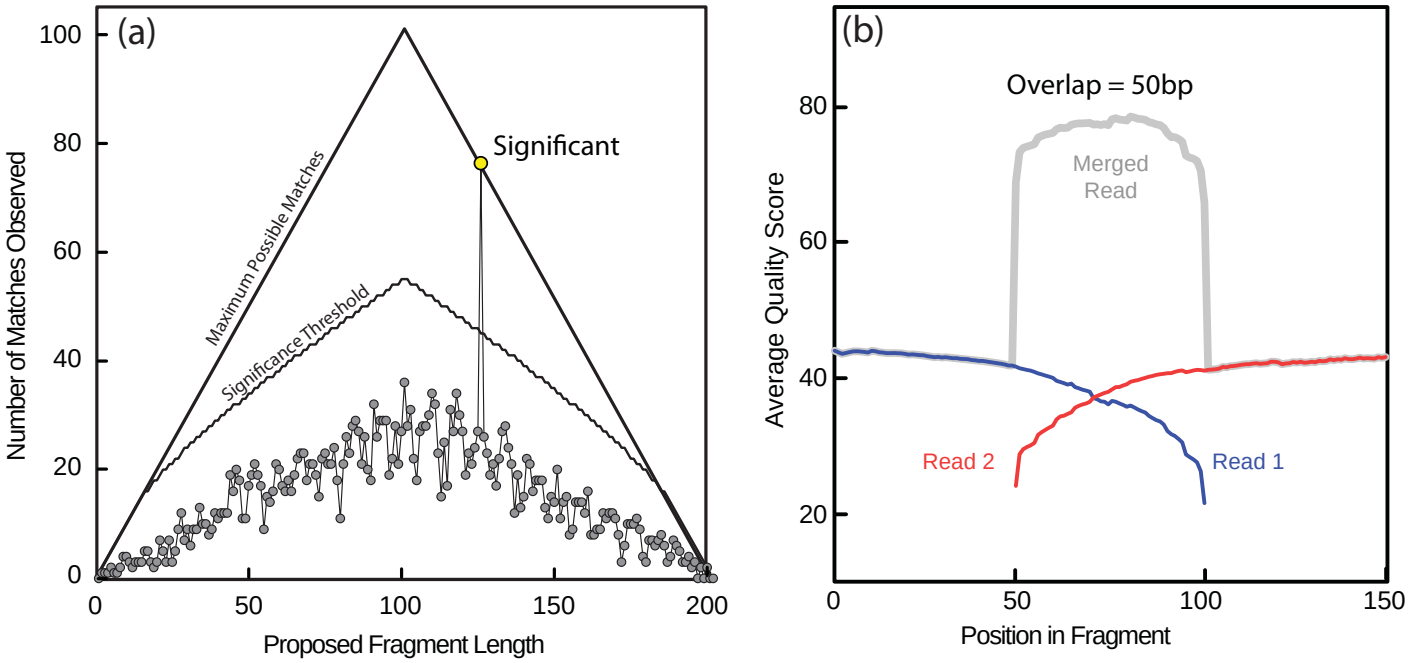
Supplemental Figure 4



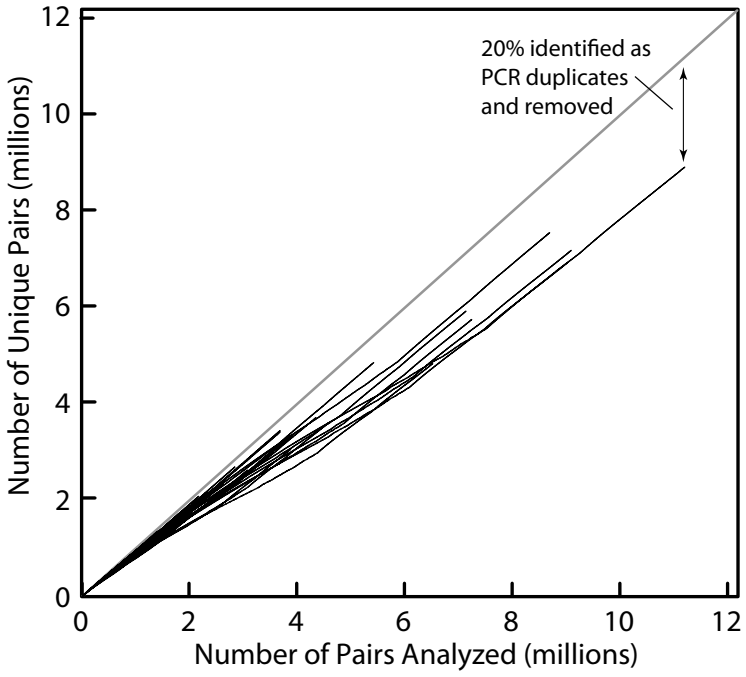
Supplemental Figure 5



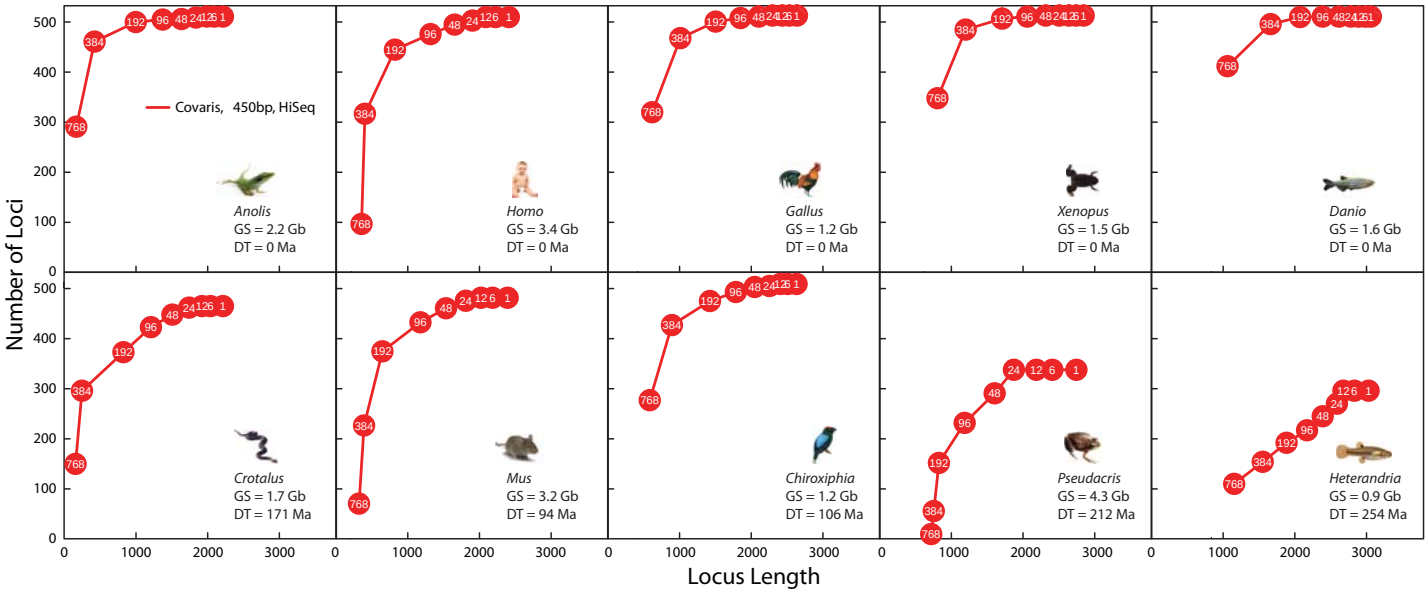
Supplemental Figure 6



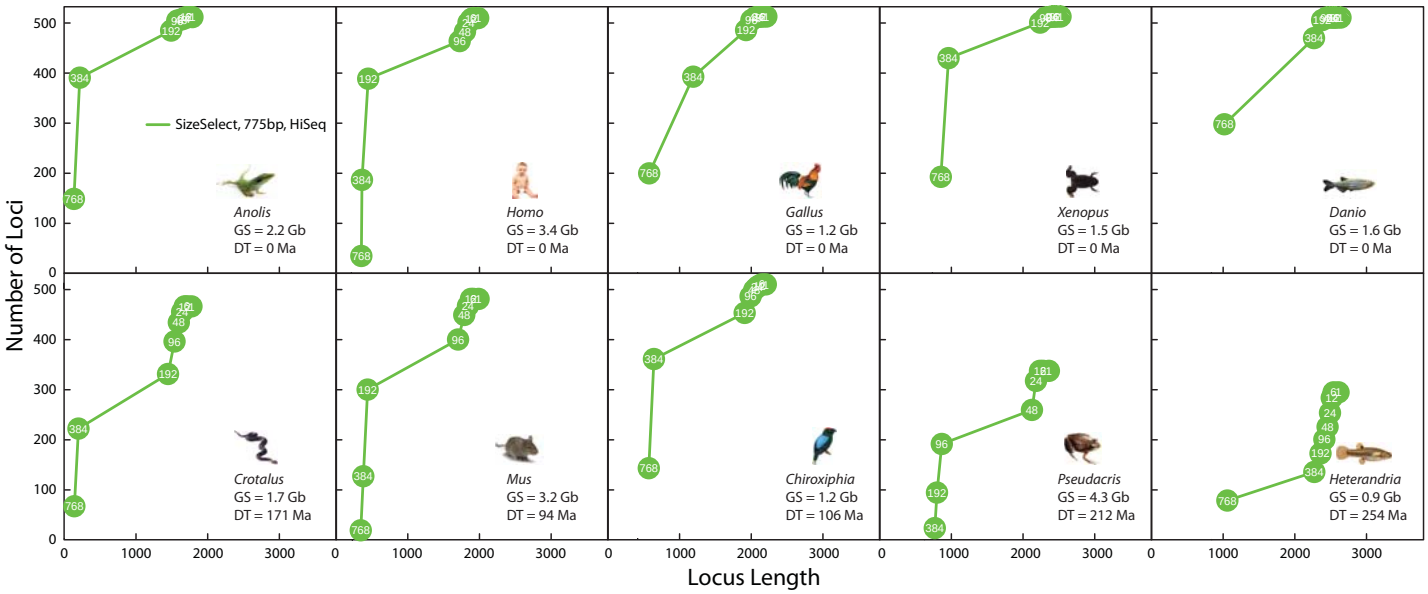
Supplemental Figure 7



Supplemental Figure 8



Supplemental Figure 9



Supplemental Figure 10