Molecular Vision 2009; 15:927-936 <http://www.molvis.org/molvis/v15/a97>
Received 7 August 2008 | Accepted 1 May 2009 | Published 9 May 2009

© 2009 Molecular Vision

Comparative genomic mapping of uncharacterized canine retinal ESTs to identify novel candidate genes for hereditary retinal disorders

B. Zangerl,[1] J. L. Johnson,[2] J. Pillardy,[3] Q. Sun,[3] C. André,[4] F. Galibert,[4] G.M. Acland,[2] G.D. Aguirre[1]

[1]Department of Clinical Studies, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA; [2]J. A. Baker Institute, College of Veterinary Medicine, Cornell University, Ithaca, NY; [3]Computational Biology Service Unit, Cornell Theory Center, Cornell University, Ithaca, NY; [4]CNRS UMR 6061, Institut de Génétique et Développement de Rennes, Faculté de Médecine, Rennes, France

Purpose: To identify the genomic location of previously uncharacterized canine retina-expressed expressed sequence tags (ESTs), and thus identify potential candidate genes for heritable retinal disorders.
Methods: A set of over 500 retinal canine ESTs were mapped onto the canine genome using the RHDF$_{5000-2}$ radiation hybrid (RH) panel, and the resulting map positions were compared to their respective localization in the CanFam2 assembly of the canine genome sequence.
Results: Unique map positions could be assigned for 99% of the mapped clones, of which only 29% showed significant homology to known RefSeq sequences. A comparison between RH map and sequence assembly indicated some areas of discrepancy. Retinal expressed genes were not concentrated in particular areas of the canine genome, and also were located on the canine Y chromosome (CFAY). Several of the EST clones were located within areas of conserved synteny to human retinal disease loci.
Conclusions: RH mapping of canine retinal ESTs provides insight into the location of potential candidate genes for hereditary retinal disorders, and, by comparison with the assembled canine genome sequence, highlights inconsistencies with the current assembly. Regions of conserved synteny between the canine and the human genomes allow this information to be extrapolated to identify potential positional candidate genes for mapped human retinal disorders. Furthermore, these ESTs can help identify novel or uncharacterized genes of significance for better understanding of retinal morphology, physiology, and pathology.

Heritable disorders of the retina often inflict devastating harm on the lives of affected individuals and their families. Progress in identifying the genetic causes of these diseases has accelerated in recent years, but still presents serious challenges; in many cases nothing is known of the genetic loci involved. Currently, for diseases where the genetic locus has been mapped, close to 30% have not yet had the responsible gene or mutation identified (RetNet). A limiting factor to this progress is an incomplete catalog of the genes expressed in the retina [1].

Animals also suffer from heritable retinal disorders. This is a source of concern when it afflicts companion or working animals and yet affected animals can be a valuable resource when used as models of human disorders. The broadly similar features of retinitis pigmentosa (RP) in human families and progressive retinal atrophy (PRA) in dogs have prompted extensive, productive, and mutually beneficial comparative studies of these homologous disorders [2].

Correspondence to: Barbara Zangerl, Ryan Veterinary Hospital, #2050, University of Pennsylvania, 3900 Delancey Street, Philadelphia, PA, 19104; Phone: (215) 898-6068; FAX: (215) 573-9853; email: bzangerl@vet.upenn.edu

Expressed sequence tags (ESTs) have played an extremely powerful role in the identification and cataloguing of new and tissue-specific genes [3]. The combination of EST discovery with radiation hybrid (RH) mapping has been invaluable to the development, assembly, and annotation of the human gene map [4], a critical first step to an assessment of individual disease loci. However, identification of disease-causing genes within mapped loci that typically have 0.5–10 cM intervals is problematic as the human genome can harbor 5–300 potential candidate genes for a single disorder in these intervals [5]. Comparative genomics is often used to further refine these results. This has recognized limitations if the identification of homolog genes is based only on the alignment of cDNA clones, which may not always cover areas of sufficient conservation between the compared species [6]. Previous efforts in the definition of the canine genome included combinations of gene identification and mapping [7], which was further improved by the release of a 7.6x draft sequence (Genome) [8]. Hence, the dog constitutes a unique resource for disease genetics.

A recently developed, normalized, canine retina-expressed sequence tag library [9], supported by an online database (DOG EST or DOG EST Project), currently catalogs

7,047 individual clones comprising more than 4,000 unique transcripts. About two-thirds of the clones in this library were identified as orthologous to annotated RefSeq sequences, and, of those, 41% corresponded, by sequence identity, to previously annotated canine cDNA entries. However, 1,418 of the transcripts remain to be annotated with certainty. In the present study, a subset of 553 clones was mapped using the well established canine/hamster hybrid panel $RHDF_{5000-2}$ [10, 11] to locate the corresponding retinal genes within the dog genome and, by comparative genomics, to infer their location in human.

Comparison of EST map positions obtained by RH and sequence mapping overcomes the disadvantages and limitations of each method alone, and offers the most reliable chromosomal locations for each EST. Such an integrated map provides substantial information on potential new candidate genes for heritable retinal disorders.

METHODS

Clone identification and primer design: A set of 1,418 ESTs could not be identified with a high level of confidence by alignment with RefSeq genes or previously mapped canine cDNAs from the previously reported canine retinal EST library [9]. This particular set was then selected for mapping on the RH panel. A total of 35 EST sequences had to be excluded from primer design based on their respective sequences. Primer pairs for the remaining EST sequences were obtained in one of two ways. 227 were manually designed using the primer select option of the Lasergene™ software (Dnastar, Inc., Madison, WI). 1,156 were selected automatically using the batch primer design feature of GeneLooper™ (GeneHarbor, Inc., Rockville, MA) with an oligo Tm of 68 °C, and allowing primers within 100–150 bp of either end of the available sequence. From the latter 1,156 primer pairs, those resulting in PCR products smaller than 100 bp were excluded, and primers were sorted for increasing penalty scores assigned by the software; this yielded 920 primer pairs that were accepted for analysis. Together with the initial 227 pairs, a total of 1,147 EST sequences were thus assigned to the mapping project (Appendix 1).

RH amplification and data collection: Each PCR product was amplified under standard conditions using 25 ng DNA template, 1 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 µM of each respective primer, and 0.5 U Taq polymerase.

Following an initial denaturation step at 94 °C for 2 min, 35 cycles were performed at 94 °C for 15 s, 52 °C, 55 °C, 58 °C, or 62 °C, respectively (Appendix 1) for 15 s, and 72 °C for 40 s. This was then followed by a final extension at 72 °C for 5 min. All PCR products were visualized on 2% agarose gels.

In the first step, primer pairs were used to amplify canine genomic DNA. Those that failed to amplify a single product of the expected size were not continued in the analysis (result F, Appendix 1). Primers that reliably amplified canine genomic DNA (result D, Appendix 1) were then tested for differential amplification of dog and hamster DNA. Those that yielded either no detectable hamster amplicon or a product significantly different in size from the canine amplicon (result RH and M, Appendix 1) were subsequently used to amplify all 118 cell lines of the canine and hamster hybrid panel $RHDF_{5000-2}$ together with previously used hamster and dog DNA as controls. PCR results were manually transferred into a computer file registering presence or absence of amplification for each cell line.

Mapping algorithm: The data were analyzed against the previously published $RHDF_{5000-2}$ map [12]. ESTs were assigned to chromosomes, based on the chromosomal location of the closest linked previously mapped RH marker [12], using the MultiMap [13] best-two-points function. Chromosomes for which marker density was too low to yield a consistent whole-chromosome linkage group were partitioned into blocks, and individual maps were created of each block. A framework map of each block was first created using MultiMap at a logarithm of the odds (LOD) of 3 or more and then filled in at continually lowering LOD scores. Chromosomes were oriented using markers previously placed on meiotic and RH maps [12].

BLAST results and graphic display: Genomic localization for each EST was assigned for the current canine genome sequence (Genome Gateway; CanFam2, May 2005) using BLASTN (cutoff value $E \leq 1e^{-100}$). The genomic position thus identified for each EST and selected anchor markers were displayed in a graph relative to the size of the respective chromosome. This was done using software developed at the Cornell University Computational Biology Service Unit (available from the authors upon request). RetNet (RetNet; status January 2009) was searched to identify the map location of hereditary retinal disorders on the human genome. For all such human disease loci mapped to intervals less than 25 Mb, the homologous region of the canine genome was established using BioMart, and refined according to the UCSC genome browser multiple alignments, disregarding alignments of less than 20 kb in length (Appendix 2).

RESULTS

Performance of retinal ESTs on the $RHDF_{5000-2}$ panel: From the initial set of 1,418 ESTs with no detected homology to previously known sequences, amplification was attempted for a subset of 1,147 markers. Of these, 998 (87%) amplified a unique PCR product from canine genomic DNA without optimization, and 711 (62%) amplified a consistent and distinctive product from the $RHDF_{5000-2}$ panel cell lines (Table 1). Roughly half of the ESTs tested could be scored satisfactorily for each of the 118 cell lines (Figure 1A). The overall presence of each EST marker on the panel (Figure 1B) was similar to previously published results (e.g., average retention frequency 22% in [12]). Furthermore, linkage to at

Table 1. EST loci retained at each experimental step.

EST loci	Number	Percent
Tested	1,147	100.0
Amplified on dog DNA	998	87.0
Amplified satisfactorily on the RH5000–2 panel	711	62.0
Readable scores for all cell lines	555	48.4
Linked to at least one other marker	553	48.2
Mapped in unique position	501	43.7

To assess quality and efficiency of primer design, all primers were first amplified on dog DNA. Subsequently, working primer pairs were amplified on the RH panel, which proved to be the least efficient step. Of those satisfactorily amplifying on the panel, almost 80% could be scored in all cell lines, and only two of the scored ESTs did not provide sufficient linkage to at least on other marker. Finally, 90% of all linked markers were assigned to unique positions in the chromosome, thus yielding results for 44% of all tested primers without further optimization.

least one other marker present in the RHDF5000–2 panel was found; this was supported by most two-point LOD scores higher than 10 (Figure 1C, Appendix 3).

Distribution of newly mapped retinal ESTs in the canine genome: Of 553 ESTs linked to known markers, 501 could be mapped to unique positions, ranging from 3 to 34 ESTs per chromosome (Table 2, Appendix 4). An additional 48 ESTs were linked, but not uniquely placed on the RH map. The remaining four EST loci collapsed with one of the other loci (Appendix 3) and were not counted as separate transcripts, reducing the total number of individually displayed loci to 549. The average number of ESTs per Mb mapped on the RH panel was calculated for each chromosome. These data were then compared to results obtained for all EST clones contained in the library (Table 2). With respect to the size of the chromosomes, relatively fewer clones were represented on the RH map for *Canis familiaris* chromosome (CFA) 19, CFA32, and CFAX; clones represented in the complete library appeared to be evenly distributed throughout the genome with the exception of CFA19, which was synteny to portions of human chromosomes 2 and 4 [10]. It remains to be elucidated whether this is due to underrepresentation of retinal clones on CFA19 or, more likely, poorer annotation of this chromosome.

Comparison of RH and sequence maps: For each chromosome, the linkage groups comprising all RH mapped retinal ESTs were aligned to the current canine genome sequence (Appendix 5). Apparent micro-rearrangements of markers among the maps were the most common discrepancy observed, particularly where markers could not be positioned on the RH map with high confidence. This is a familiar problem for markers located toward the end of a linkage group, e.g., CFA6 and CFA10 (Appendix 5).

Fourteen markers diverged in placement between the RH and sequence maps. Eight of these yielded best two-point linkage to a single chromosome, but could not be mapped in a unique position on the RH map (Appendix 5; e.g., DR010005B10H04 on CFA22). Therefore, correct genomic position of these ESTs should be assumed based on the sequence alignment, since it is the better supported method for these clones. The other six ESTs (DR01007A10C09 and DR010025B10A10 on CFA4, DR010009A20D03 on CFA6, DR010013A10E04 on CFA7, DR010027B10D03 on CFA17, and DR010015B20A07 on CFA19) were in complete disagreement with the RH map and the sequence assemblies (Appendix 5; no placements on D). Without additional information, correct genomic location of these ESTs cannot be determined.

Indication of genome sequence alignment problems: A subset of ESTs proved difficult to align to a unique area of the sequence assembly. Despite the high cut-off value ($E \leq 1e^{-100}$), 22 markers that were RH mapped to 16 different chromosomes showed significant sequence alignment to more than one chromosome. In each case, the multiple alignments included the chromosome assigned by the RH map. A consensus chromosomal position for these markers was thus based on the RH map (Appendix 5, e.g., DR010017A21D11 on CFA14, DR01005A20E07 on CFA15, placement on C, D, and E).

In addition to the aforementioned issues in placing individual markers, some genomic regions revealed consistent complications. Several markers RH mapped to CFA17 did not align to the same chromosome in the canine genome sequence draft; DR010027B10D03 mapped to CFA3, while DR010006A10A08, DR010025A10C12, DR010010B20B05, and DR010024A10E04 yielded multiple hits on CFA17, in addition to several alignments against alignments against a part of the sequence assembly not assigned to chromosomes (chrUN). Similarly, marker DR010026A20B10 on CFA11 aligned to multiple genomic locations. It is also worth mentioning that the three ESTs RH mapped to CFAY have significant sequence homology to locations on CFAX in the genome sequence assemblies.

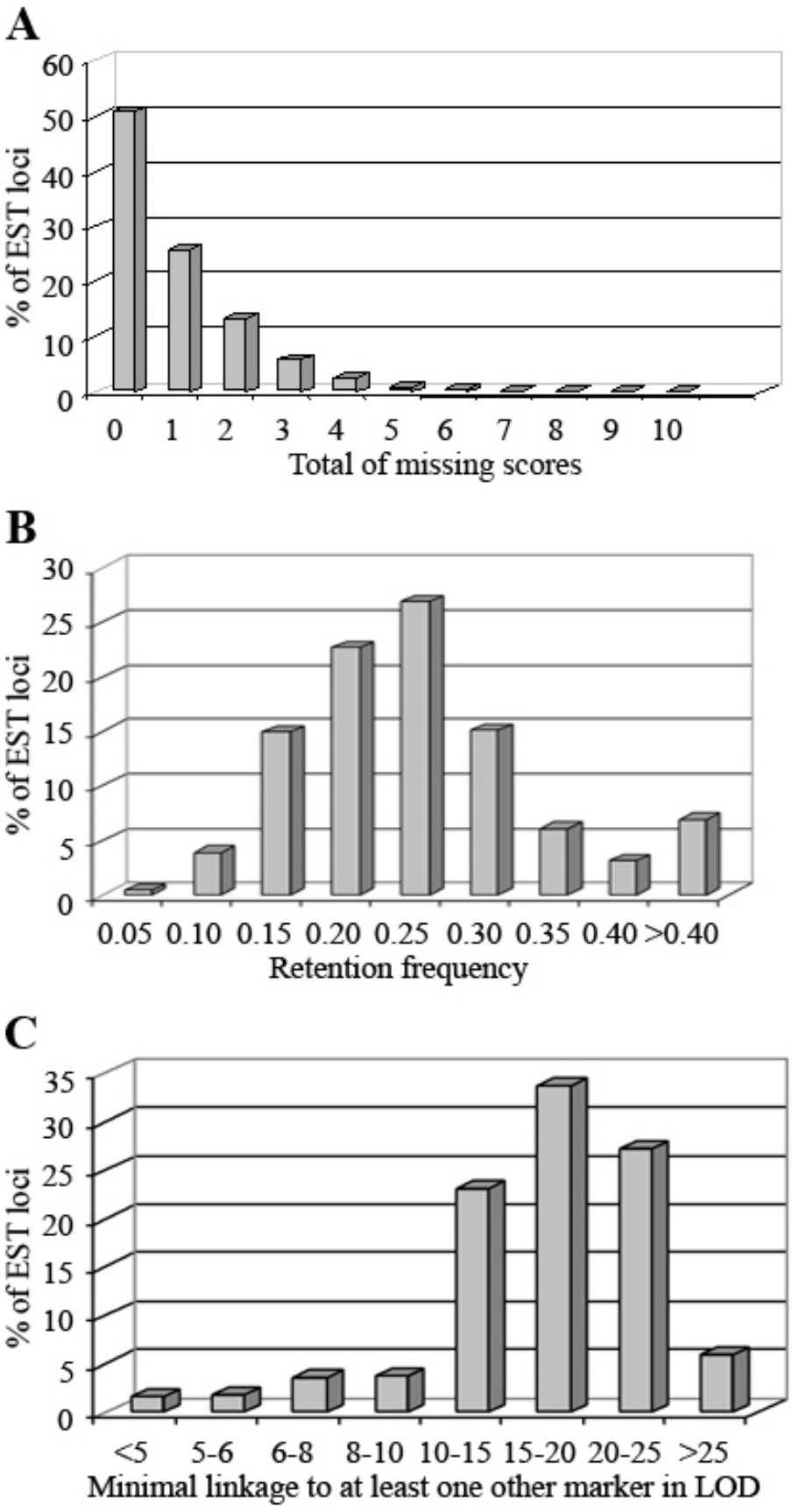


Figure 1. Quality control for retinal clones on the $RHDF_{5000-2}$ panel. A total of 555 retinal clones were amplified from 118 cell lines representing the $RHDF_{5000-2}$ panel. For each locus, we assessed both, the overall number of cell lines that could unambiguously be scored, and the number of cell lines amplifying the respective EST, for quality. Half of the loci were scored in each individual line with the balance of loci missing only few scores (**A**). The respective retention frequency resulting from amplification scores, on average, was 0.22 and showed a distribution that is similar to previously published data [12] using this panel (**B**). The good quality performance of EST amplification resulted in highly supported linkage to known markers (**C**) with most of the LOD scores above 10.

Table 2. ESTs and markers mapped per chromosome.

CFA	CFA size, MB	blocks number	RH markers uniquely mapped	ESTs uniquely mapped	ESTs linked	EST/MB RH panel	ESTs database	EST/MB database
1	137	4	135	34	1	0.26	895	6.53
2	99	2	90	20	0	0.20	733	7.40
3	105	4	85	18	4	0.21	580	5.52
4	100	2	93	16	0	0.16	578	5.78
5	99	3	108	26	0	0.26	599	6.05
6	87	2	79	16	1	0.20	462	5.31
7	94	4	114	15	0	0.16	673	7.16
8	86	1	76	13	3	0.19	689	8.01
9	77	4	95	16	3	0.25	522	6.78
10	80	3	60	15	3	0.23	702	8.78
11	86	1	96	14	2	0.19	592	6.88
12	85	2	113	20	1	0.25	564	6.64
13	75	1	54	8	1	0.12	331	4.41
14	72	1	75	9	1	0.14	559	7.76
15	75	1	76	16	0	0.21	497	6.63
16	73	3	64	18	2	0.27	368	5.04
17	80	2	80	17	1	0.23	435	5.44
18	66	1	79	16	1	0.26	543	8.23
19	66	3	57	4	0	0.06	212	3.21
20	66	1	93	13	5	0.27	507	7.68
21	61	4	86	8	1	0.15	408	6.69
22	61	1	53	7	6	0.21	393	6.44
23	61	1	51	7	2	0.15	318	5.21
24	73	1	51	16	0	0.22	369	5.05
25	60	2	68	14	1	0.25	451	7.52
26	48	1	50	9	1	0.21	389	8.10
27	57	1	67	15	1	0.28	532	9.33
28	55	1	53	19	1	0.36	310	5.64
29	51	1	53	8	0	0.16	313	6.14
30	47	2	42	15	2	0.36	415	8.83
31	50	2	34	8	2	0.20	265	5.30
32	51	1	29	3	0	0.06	447	8.76
33	41	1	39	11	0	0.27	215	5.24
34	50	1	41	6	1	0.14	221	4.42
35	38	1	24	5	0	0.13	187	4.92
36	41	1	44	4	1	0.12	220	5.37
37	40	1	47	6	0	0.15	214	5.35
38	38	2	24	5	0	0.13	149	3.92
X	139	3	52	8	0	0.06	927	6.67
Y	27	2	9	3	0	0.11	N/A	N/A

Each chromosome was mapped in individual linkage groups (blocks, column 3) containing previously mapped markers (column 4, reference [12]), ESTs mapped in unique positions (column 5), and ESTs linked, but not ordered on the chromosome (column 6). The number of RH-mapped ESTs per MB for each chromosome (column 7) was compared to the number of all ESTs currently in the database for each chromosome (column 8) per MB (column 9) to assess distribution of retinal expressed genes throughout the genome.

Identification of potential candidate genes for human retinal disorders: The homologous canine genome locations for 28 mapped human retinal disease loci (RetNet), for which disease causing genes remain unknown, were identified on the canine sequence assembly (Figure 2, Appendix 2). On average, these identified homologous canine segments were 83% of the size

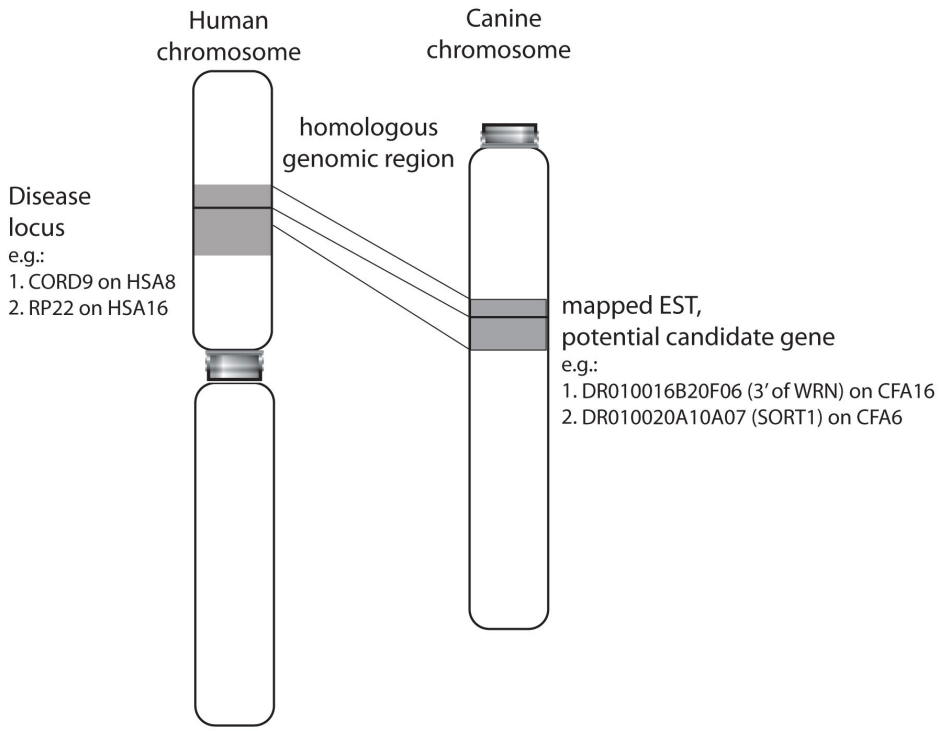


Figure 2. Identification of potential candidate genes for human retinal disease Human genomic intervals for known diseases (e.g., CORD9 on HSA8, RP22 on HSA16) were mapped against the canine genome to identify homologous regions, and EST within these regions of interest (e.g., DR010016B20F06 on CFA16, DR010020A10A07 on CFA6). A comprehensive list of these disorders and the number of corresponding ESTs contained within our library is given in Appendix 2. ESTs mapped in the presented research are also illustrated on the respective chromosomes in Appendix 5. Details on all clones can be obtained through a web database (DOG EST or DOG EST Project) to obtain insights into corresponding transcripts (e.g.: 1. WRN, 2. SORT1). We suggest that this tool provides new positional candidate genes for mapped human retinal disorders. This would allow for the identification of mutations in genes that are thus far unknown or have not yet been linked to retinal disorders, after the exclusion of conventional candidate genes.

of the corresponding human genome sequence interval, while the interval homologous to CORS2 (HSA11p12-q13.3) was covered by only 49%. The latter is likely because it spans the HSA11 centromere. Candidate ESTs were identified for these locations based on their positional annotation to identify potential candidate genes for the respective disorders. For only one locus, X-chromosomal RP24, no potential candidate clones mapped to the proposed homologous disease interval. For twenty of these comparative genome regions in the dog, which include clones placed on the presented RH maps (Appendix 2, shaded), the respective diseases intervals are also illustrated on the canine chromosome maps (Appendix 5).

DISCUSSION

Comparison of mapping approaches: EST clones from the previously published canine retinal library [9] were re-aligned against RefSeq sequences, with a criterion cut-off value of $E \leq 1e^{-3}$ accepted as establishing homology. This yielded annotation for 80% of clones evaluated. In the present study, hypothesizing that the remaining, unidentified sequences might represent previously unrecognized retinal genes or control elements relevant to retinal function, 553 such retinal clones were mapped to the $RHDF_{5000-2}$ panel with high confidence, and of these, 501 were mapped to unique positions (Table 1, Figure 1, and Appendix 3). The resulting RH map thus integrated these ESTs into the overall canine genome. Among the mapped ESTs, 159 sequences (29%) aligned to known RefSeq sequences, but the balance (71%; 394) remained unidentified.

Direct comparison of this integrated RH map to the assembled canine genome sequence became possible during the progress of the present project. This permitted independent confirmation of the location of mapped clones, comparative evaluation of the different mapping methods, and, as a consequence, overcame some of the limitations of each method. RH mapping is relatively unreliable toward the end of linkage groups, and does not provide highly reliable order for markers located close together. However, some of these problems have recently been addressed in a refined map for this panel [14]. Nonetheless, mapping of markers by sequence alignment is subject to errors in assembly of the underlying sequence. For example, gene duplications, pseudo genes, and other sequence similarities within a single genome can create both ambiguities in assembly, and multiple significant alignments for a single marker; this was recognized in 4% of the presented data. Furthermore, alignment of expressed sequences, which may be spliced, against a genome assembly, can be problematic since the spliced out sequence (e.g., introns) can cause difficulties in evaluating alignment scores.

The greater part of inconsistencies between placement on the RH and sequence maps were minor. These were more

common at the ends of chromosomes rather than the middle. In particular, the multiple sequence alignments for ESTs mapping near the centromere of CFA11 and parts of CFA17 suggest sequence or assembly problems, or even potential genome duplications. In general, major disagreements between the compared maps were rare and resulted in only 1% of the chromosomal placement of ESTs to remain unresolved. While the data indicate room for improvement, these issues are likely to be resolved with the improvement of both RH maps and genome sequence assemblies. The synergistic combination of different approaches provides higher fidelity genome maps than does any single method alone [15].

Genomic coverage of retinal ESTs: Overall, retinal ESTs analyzed in the current study do not appear to be concentrated in particular areas of the genome, but were distributed rather evenly throughout. There were, however, some departures from this general observation. CFA19, CFA32, and CFAX were relatively underrepresented by ESTs mapped in the present study, but only CFA19 showed the same underrepresentation when adjusted for all ESTs from the same library (Table 2). This suggests that the initial annotations were more efficient for CFA32 and CFAX and, thus, fewer clones were selected for additional mapping and annotation. For CFA19, however, in addition to a lower representation of retinal EST clones, one of the four ESTs RH-mapped to this chromosome aligned to a different chromosome (CFA27) in the canine genome sequence assembly, indicating a potential error in one or the other map.

We attempted to identify canine ESTs that potentially might represent novel candidate sequences for 28 mapped loci for hereditary human retinal disorders. Surprisingly, the only such locus that did not yield canine retinal ESTs contained neither in the complete EST library nor was presented in the RH mapped subset, was X-linked. This is especially remarkable considering the well recognized strong conservation of the mammalian X chromosome [16,17]. More importantly, all three ESTs mapped to CFAY on the $RHDF_{5000-2}$ panel aligned close to the centromere of CFAX based on sequence homology. Based on the absence of CFAY sequences from the CanFam2 assembly, the correct assignment of the RH map linkage group to this chromosome cannot be verified by sequence alignments. One therefore is left to hypothesize whether this finding represent incorrect linkage in the RH map or functional transcripts are indeed encoded on CFAY. However, since the genomic CFAX sequence aligning to one of the clones, DR010006B10E05, contained a microsatellite that is not present in the EST, it is possible these transcripts have become inactivated on CFAX but are still functionally present on CFAY. Further studies will be necessary to confirm and interpret this potentially interesting finding.

Candidate genes for human retinal disorders: Despite considerable progress in recent years, there has been a steady gap between identification of the mapped loci for heritable human retinal disorders, and characterization of the underlying gene and causative mutations (D graph). In part, this reflects an incomplete understanding of the genes critical for retinal development, function, and maintenance. Currently, 48 such loci are recognized for which no causative sequence change has been defined. For 20 loci, the candidate region is either not precisely defined or extends over more than 25 Mb. For the remaining 28 loci, we attempted to identify retina-expressed ESTs that might represent potential candidate genes, and were able to do so for all except the aforementioned X-linked RP24 locus. It should be acknowledged that some of these clones may be redundant—i.e., more than one EST might represent the same transcript. Thus, the number of potential retinal expressed candidate genes could be overestimated in our data set. Furthermore, genomic regions for some human disease loci also do not align unambiguously between human and dog, (e.g., MRST aligned to several areas on CFA3 and CFA30), and, consequently, we may not have covered the complete homologous disease intervals in the dog.

Confidence in the RH map location of an EST, and, concomitantly in its potential as a positional candidate, is strengthened when confirmed by sequence alignment with the corresponding genomic interval. This was achieved for the majority of ESTs (93.6%) in the current study. However, three clones potentially located within disease relevant areas were mapped to different genomic locations with the RH map when compared to the genomic sequence. DR010007A10C09, which aligns within the MRST homologous interval on CFA3 in the sequence assembly, was mapped with high confidence to CFA4 on the RH map (Appendix 5, CFA4, C). In contrast, DR010030A10D12 was located within the CORD9 interval on CFA16 in the RH map, but showed sequence homology just outside this interval in the sequence assembly; the same applies to DR010030A10G12 in the RP23 interval on CFAX. Despite these minor inconsistencies, however, homologous canine intervals have successfully been established for most of the human loci extracted from RetNet (Figure 2, Appendix 2).

The importance of correctly assigned location and tissue specificity of potential disease candidate genes has recently been demonstrated for progressive rod-cone degeneration (*prcd*) in the canine model. This disorder had previously been mapped to CFA9, and known retinal expressed genes within the disease critical interval had been excluded from causative association with disease [18-20]. Once the EST library was screened for positional candidates, transcripts expressing a novel gene were identified harboring the mutation responsible for *prcd* [21]. Subsequently, the gene was found to also cause RP in humans. It should be remarked that this effort was initially complicated by the fact that the disease critical interval was incorrectly assembled to CFA18 in the first pass of the canine genome assembly CanFam1. However, the

canine draft sequence has been revised and is in high agreement with the presented RH mapping data.

A major advantage of comparative genomics lies in the validation of results and preliminary data in the absence of large patient cohorts to repeat or further narrow linkage data of known disease loci. One of the most recent genes identified to contribute to vision loss in humans is *PROM1* [22]. A missense mutation in the gene has been linked to formerly mapped loci MCDR2 and STGD4 in a screen of 12 positional candidate genes within the minimal overlapping disease interval of 12cM between markers D4S1582 and GATA1582G03. This region corresponds to 9.7 Mb (bp 65,599,776 −71,744,479; 90,673,048–94,195,452) located on CFA3 in the dog. The presented library contains 10 transcripts for this genomic location; the most abundant one of those, eight clones comprising contig 1265, represents the canine version of *PROM1*. Thus, a cross species comparison may have been able to assist in the ranking of candidate genes for this disease. With the high fidelity of genomic location established, this resource is well suited to assist with the investigation of known and new disease loci to speed up first steps in the identification and understanding of retinal disorders.

In conclusion, a comparative map of retinal genes allows identification of new candidate genes for retinal disorders in both dog and human, and provides a further step toward the complete categorization of genes relevant to retinal development and degeneration. Results of this study are integrated into the web database at DOG EST or DOG EST Project, and are publically accessible.

ACKNOWLEDGMENTS

This research was supported by NEI/NIH grants EY-06855, EY-13132, and EY-017549, the Foundation Fighting Blindness, The Morris Animal Foundation, The Van Sloun Fund for Canine Genetic Research, Pfizer, Inc., The University of Pennsylvania Research Foundation, and the Laura J. Niles Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Eye Institute, the National Institutes of Health, or the other sources of support identified.

REFERENCES

1. Schulz HL, Goetz T, Kaschkoetoe J, Weber BH. The retinome - defining a reference transcriptome of the adult mammalian retina/retinal pigment epithelium. BMC Genomics 2004; 5:50. [PMID: 15283859]
2. Aguirre GD, Acland GM. Models, mutants and man: Searching for unique phenotypes and genes in the dog model of inherited retinal degeneration. In: Ostrander EA, Giger U, Lindblad-Toh K, editors. The Dog And Its Genome. Cold Spring Harbor (NY): Cold Spring Harbor Press; 2006. p. 291–325.
3. Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC. Complementary DNA sequencing: Expressed sequence tags and human genome project. Science 1991; 252:1651-6. [PMID: 2047873]
4. Deloukas P, Schuler GD, Gyapay G, Beasley EM, Soderlund C, Rodriguez-Tomé P, Hui L, Matise TC, McKusick KB, Beckmann JS, Bentolila S, Bihoreau M, Birren BB, Browne J, Butler A, Castle AB, Chiannilkulchai N, Clee C, Day PJ, Dehejia A, Dibling T, Drouot N, Duprat S, Fizames C, Fox S, Gelling S, Green L, Harrison P, Hocking R, Holloway E, Hunt S, Keil S, Lijnzaad P, Louis-Dit-Sully C, Ma J, Mendis A, Miller J, Morissette J, Muselet D, Nusbaum HC, Peck A, Rozen S, Simon D, Slonim DK, Staples R, Stein LD, Stewart EA, Suchard MA, Thangarajah T, Vega-Czarny N, Webber C, Wu X, Hudson J, Auffray C, Nomura N, Sikela JM, Polymeropoulos MH, James MR, Lander ES, Hudson TJ, Myers RM, Cox DR, Weissenbach J, Boguski MS, Bentley DR. A physical map of 30,000 human genes. Science 1998; 282:744-6. [PMID: 9784132]
5. van Driel MA, Cuelenaere K, Kemmeren PP, Leunissen JA, Brunner HG. A new web-based data mining tool for the identification of candidate genes for human genetic disorders. Eur J Hum Genet 2003; 11:57-63. [PMID: 12529706]
6. Sharma S, Chang JT, Della NG, Campochiaro PA, Zack DJ. Identification of novel bovine RPE and retinal genes by subtractive hybridization. Mol Vis 2002; 8:251-8. [PMID: 12131876]
7. Roberts MC, Hitte CC, Hendrickson JA, Hoffmann DE, Flickinger GH, Rutherford MS, Guyon RS, Galibert FS, Mickelson JR. Characterization and radiation hybrid mapping of expressed sequence tags from the canine brain. Mamm Genome 2003; 14:203-13. [PMID: 12647243]
8. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ 3rd, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin CW, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli KP, Parker HG, Pollinger JP, Searle SM, Sutter NB, Thomas R, Webber C, Baldwin J, Abebe A, Abouelleil A, Aftuck L, Ait-Zahra M, Aldredge T, Allen N, An P, Anderson S, Antoine C, Arachchi H, Aslam A, Ayotte L, Bachantsang P, Barry A, Bayul T, Benamara M, Berlin A, Bessette D, Blitshteyn B, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Brown A, Cahill P, Calixte N, Camarata J, Cheshatsang Y, Chu J, Citroen M, Collymore A, Cooke P, Dawoe T, Daza R, Decktor K, DeGray S, Dhargay N, Dooley K, Dooley K, Dorje P, Dorjee K, Dorris L, Duffey N, Dupes A, Egbiremolen O, Elong R, Falk J, Farina A, Faro S, Ferguson D, Ferreira P, Fisher S, FitzGerald M, Foley K, Foley C, Franke A, Friedrich D, Gage D, Garber M, Gearin G, Giannoukos G, Goode T, Goyette A, Graham J, Grandbois E, Gyaltsen K, Hafez N, Hagopian D, Hagos B, Hall J, Healy C, Hegarty R, Honan T, Horn A, Houde N, Hughes L, Hunnicutt L, Husby M, Jester B, Jones C, Kamat A, Kanga B, Kells C, Khazanovich D, Kieu AC, Kisner P, Kumar M, Lance K, Landers T, Lara M, Lee W, Leger JP, Lennon N, Leuper L, LeVine S, Liu J, Liu X, Lokyitsang Y, Lokyitsang T, Lui A, Macdonald J, Major J, Marabella R, Maru K,

Matthews C, McDonough S, Mehta T, Meldrim J, Melnikov A, Meneus L, Mihalev A, Mihova T, Miller K, Mittelman R, Mlenga V, Mulrain L, Munson G, Navidi A, Naylor J, Nguyen T, Nguyen N, Nguyen C, Nguyen T, Nicol R, Norbu N, Norbu C, Novod N, Nyima T, Olandt P, O'Neill B, O'Neill K, Osman S, Oyono L, Patti C, Perrin D, Phunkhang P, Pierre F, Priest M, Rachupka A, Raghuraman S, Rameau R, Ray V, Raymond C, Rege F, Rise C, Rogers J, Rogov P, Sahalie J, Settipalli S, Sharpe T, Shea T, Sheehan M, Sherpa N, Shi J, Shih D, Sloan J, Smith C, Sparrow T, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Stone S, Sykes S, Tchuinga P, Tenzing P, Tesfaye S, Thoulutsang D, Thoulutsang Y, Topham K, Topping I, Tsamla T, Vassiliev H, Venkataraman V, Vo A, Wangchuk T, Wangdi T, Weiand M, Wilkinson J, Wilson A, Yadav S, Yang S, Yang X, Young G, Yu Q, Zainoun J, Zembek L, Zimmer A, Lander ES. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 2005; 438:803-19. [PMID: 16341006]

9. Zangerl B, Sun Q, Pillardy J, Johnson JL, Schweitzer PA, Hernandez AG, Liu L, Acland GM, Aguirre GD. Development and characterization of a normalized canine retinal cDNA library for genomic and expression studies. Invest Ophthalmol Vis Sci 2006; 47:2632-8. [PMID: 16723480]
10. Guyon R, Lorentzen TD, Hitte C, Kim L, Cadieu E, Parker HG, Quignon P, Lowe JK, Renier C, Gelfenbeyn B, Vignaux F, DeFrance HB, Gloux S, Mahairas GG, André C, Galibert F, Ostrander EA. A 1-Mb resolution radiation hybrid map of the canine genome. Proc Natl Acad Sci USA 2003; 100:5296-301. [PMID: 12700351]
11. Vignaux F, Hitte C, Priat C, Chat JC, Andre C, Galibert F. Construction and optimization of a dog whole-genome radiation hybrid panel. Mamm Genome 1999; 10:888-94. [PMID: 10441740]
12. Breen M, Jouquand S, Renier C, Mellersh CS, Hitte C, Holmes NG, Chéron A, Suter N, Vignaux F, Bristow AE, Priat C, McCann E, André C, Boundy S, Gitsham P, Thomas R, Bridge WL, Spriggs HF, Ryder EJ, Curson A, Sampson J, Ostrander EA, Binns MM, Galibert F. Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. Genome Res 2001; 11:1784-95. [PMID: 11591656]
13. Matise TC, Perlin M, Chakravarti A. Automated construction of genetic linkage maps using an expert system (MultiMap): A human genome linkage map. Nat Genet 1994; 6:384-90. [PMID: 8054979]
14. Breen M, Hitte C, Lorentzen TD, Thomas R, Cadieu E, Sabacan L, Scott A, Evanno G, Parker HG, Kirkness EF, Hudson R, Guyon R, Mahairas GG, Gelfenbeyn B, Fraser CM, André C, Galibert F, Ostrander EA. An integrated 4249 marker FISH/RH map of the canine genome. BMC Genomics 2004; 5:65. [PMID: 15363096]
15. Hitte C, Madeoy J, Kirkness EF, Priat C, Lorentzen TD, Senger F, Thomas D, Derrien T, Ramirez C, Scott C, Evanno G, Pullar B, Cadieu E, Oza V, Lourgant K, Jaffe DB, Tacher S, Dreano S, Berkova N, Andre C, Deloukas P, Fraser C, Lindblad-Toh K, Ostrander EA, Galibert F. Facilitating genome navigation: survey sequencing and dense radiation-hybrid gene mapping. Nat Rev Genet 2005; 6:643-8. [PMID: 16012527]
16. Glas R, Marshall Graves JA, Toder R, Ferguson-Smith M, O'Brien PC. Cross-species chromosome painting between human and marsupial directly demonstrates the ancient region of the mammalian X. Mamm Genome 1999; 10:1115-6. [PMID: 10556436]
17. Raudsepp T, Lee EJ, Kata SR, Brinkmeyer C, Mickelson JR, Skow LC, Womack JE, Chowdhary BP. Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. Proc Natl Acad Sci USA 2004; 101:2386-91. [PMID: 14983019]
18. Sidjanin DJ, Miller B, Kijas J, McElwee J, Pillardy J, Malek J, Pai G, Feldblyum T, Fraser C, Acland G, Aguirre G. Radiation hybrid map, physical map, and low-pass genomic sequence of the canine prcd region on CFA9 and comparative mapping with the syntenic region on human chromosome 17. Genomics 2003; 81:138-48. [PMID: 12620391]
19. Acland GM, Ray K, Mellersh CS, Gu W, Langston AA, Rine J, Ostrander EA, Aguirre GD. Linkage analysis and comparative mapping of canine progressive rod-cone degeneration (prcd) establishes potential locus homology with retinitis pigmentosa (RP17) in humans. Proc Natl Acad Sci USA 1998; 95:3048-53. [PMID: 9501213]
20. Goldstein O, Zangerl B, Pearce-Kelling S, Sidjanin DJ, Kijas JW, Felix J, Acland GM, Aguirre GD. Linkage disequilibrium mapping in the domestic dog breeds - a powerful genetic tool for the identification of disease genes. Genomics 2006; 88:541-50. [PMID: 16859891]
21. Zangerl B, Goldstein O, Philp AR, Lindauer SJ, Pearce-Kelling SE, Mullins RF, Graphodatsky AS, Ripoll D, Felix JS, Stone EM, Acland GM, Aguirre GD. Identical mutation in a novel retinal gene causes progressive rod-cone degeneration in dogs and retinitis pigmentosa in man. Genomics 2006; 88:551-63. [PMID: 16938425]
22. Yang Z, Chen Y, Lillo C, Chien J, Yu Z, Michaelides M, Klein M, Howes KA, Li Y, Kaminoh Y, Chen H, Zhao C, Chen Y, Al-Sheikh YT, Karan G, Corbeil D, Escher P, Kamaya S, Li C, Johnson S, Frederick JM, Zhao Y, Wang C, Cameron DJ, Huttner WB, Schorderet DF, Munier FL, Moore AT, Birch DG, Baehr W, Hunt DM, Williams DS, Zhang K. Mutant prominin 1 found in patients with macular degeneration disrupts photoreceptor disk morphogenesis in mice. J Clin Invest 2008; 118:2908-16. [PMID: 18654668]

Appendix 1.

To access the data, click or select the words "Appendix 1." This will initiate the download of a (pdf) archive that contains the file. Amplification details and results (F=PCR failed, D=amplified from dog genomic DNA, RH=amplified on the RH panel, M=placed on the final map) for all ESTs tested.

Appendix 2.

To access the data, click or select the words "Appendix 2." This will initiate the download of a (pdf) archive that contains the file. Human retinal disease loci mapped against the canine genome. Based on the corresponding intervals, potential candidate genes were identified from the canine retinal EST library. The number of those ESTs is provided in the last column (EST contained in the complete database/ subset al.so included in the currently presented RH mapping project). Genomic regions containing ESTs mapped in the present study (shaded) are also illustrated in Appendix 5.

Appendix 3.

To access the data, click or select the words "Appendix 3." This will initiate the download of a (pdf) archive that contains the file. Chromosomal location of each EST was assigned using the MultiMap [13] best-two points function using the previously published canine RH map [12]. Results are listed for each EST to demonstrate linkage to the closest marker (**A**) or indicate that the EST has collapsed with one of the markers and been placed in the same genomic location (**B**).

Appendix 4.

To access the data, click or select the words "Appendix 4." This will initiate the download of a (pdf) archive that contains the file. Markers were grouped into chromosomes according to results of the best-two points function (Appendix 3) and subsequently mapped within each chromosome using MultiMap [13]. Chromosomes for which marker density was too low to yield a consistent whole-chromosome linkage group were petitioned into blocks, and individual maps were created of each block. Final orientation and location of all markers on the chromosome are displayed according to previously established chromosome orientation [12]. Appendix lists all markers used for mapping, selected markers and EST locations are graphically displayed in Appendix 5.

Appendix 5. Comparative maps of retinal expressed clones in the canine genome.

To access the data, click or select the words "Appendix 5." This will initiate the download of a (pdf) archive that contains the file. Results of the EST mapping are graphically displayed for each chromosome indicating the relative position (**A**) of all mapped ESTs (**B**, blue) on the RH panel (**C**), the corresponding chromosome sequence draft (**D**), and potential alignments to nonhomologous chromosomes (**E**). Placement is indicated by horizontal lines, while oblique lines connect placements on different maps for the same EST. Chromosomes are oriented with centromeres located next to the respective legend on top of each page (**A**, 0); the artificial scale to the left (**A**) is proportional to the respective chromosome size based on the RH mapping. For orientation, selected markers are displayed on the RH map (**B**, red), and marker groups mapped in individual blocks are separated by black lines. ESTs linked, but not uniquely mapped to individual chromosome are displayed in *italic*. Intervals corresponding to mapped human disease loci (Appendix 2) are indicated on the right of each homologous region of the sequence drafts (**C**).

The print version of this article was created on 2 May 2009. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.