1   **A glance at recombination hotspots in the domestic cat**

2

3   Hasan Alhaddad[1], Chi Zhang[2], Bruce Rannala[3], and Leslie A. Lyons[4]

4   [1]College of Science, Department of Biological Sciences, Kuwait University,

5   Safat, Kuwait, 13060

6   [2]Department of Bioinformatics and Genetics, Swedish Museum of Natural

7   History, Box 50007, SE-104 05 Stockholm, Sweden

8   [3]Genome Center and Department of Evolution and Ecology, University of

9   California - Davis, Davis, CA 95616, USA

10   [4]Department of Veterinary Medicine and Surgery, College of Veterinary

11   Medicine, University of Missouri-Columbia, Columbia, MO, 65211

12

13   Corresponding author:

14   Hasan Alhaddad

15   Kuwait University

16   Safat, Kuwait, 13060

17   Email: hassan.alhaddad@ku.edu.kw

18   hhalhaddad@gmail.com

19   Tel: +965 9770 1635

20   Running title: Recombination hotspots in cats

1   Abstract

2   Recombination has essential roles in increasing genetic variability within a

3   population and in ensuring successful meiotic events. The objective of this study

4   is to (i) infer the population scaled recombination rate (ρ), and (ii) identify and

5   characterize localities of increased recombination rate for the domestic cat, *Felis*

6   *silvestris catus*. SNPs ($n = 701$) were genotyped in twenty-two cats of Eastern

7   random bred origin. The SNPs covered ten different chromosomal regions (A1,

8   A2, B3, C2, D1, D2, D4, E2, F2, X) with an average region size of 850 Kb and an

9   average SNP density of 70 SNPs/region. The Bayesian method in the program

10 *inferRho* was used to infer regional population recombination rates and hotspots

11 localities. The regions exhibited variable population recombination rates and four

12 decisive recombination hotspots were identified on cat chromosome A2, D1, and

13 E2 regions. No correlation was detected between the GC content and the locality

14 of recombination spots. The hotspots enclosed L2 LINE elements and MIR and

15 tRNA-Lys SINE elements in agreement with hotspots found in other mammals.

16

17 **Keywords:** recombination, rho, recombination hotspot, domestic cat, *Felis*

18 *silvestris catus*

1 **Introduction**

2   Recombination is a major source of genetic variation within sexually

3 reproducing organisms, and is necessary for the proper alignment and segregation

4 of homologous chromosomes during meiosis. Variation is achieved via

5 recombination when new combinations of parental alleles across loci are

6 generated and transmitted to succeeding generations. Lack of recombination may

7 cause failure of meiotic division, or formation of gametes with chromosome

8 number abnormalities (aneuploidy), which are often detrimental.

9   Localized chromosomal regions with elevated recombination rates are

10 referred to as recombination "hotspots" (Steinmetz *et al*, 1982). Known hotspots

11 in mice and human are generally 1 - 2 kb regions of high recombination rates

12 surrounded by regions of low recombination (Paigen and Petkov, 2010). In

13 humans, recombination hotspots are distributed about every 200 Kb (McVean *et*

14 *al*, 2004) and over 25 000 hotspots have been identified (Myers *et al*, 2005). The

15 first human recombination hotspot was identified using sperm typing via PCR in a

16 region that harbors GC-rich mini-satellite (MS32) as a molecular signature

17 (Jeffreys *et al*, 1998).

18   Several genomic features exhibit correlations with recombination hotspots.

19 The GC content has been found to be positively correlated with recombination

20 hotspots in humans (Myers *et al*, 2005), dog (Axelsson *et al*, 2012), pig

21 (Tortereau *et al*, 2012), and chicken (Groenen *et al*, 2009) but not in mice (Wu *et*

1   *al*, 2010). Long terminal repeats (LTR), long interspersed elements (LINE), and

2   short interspersed elements (SINE) were also observed to be positively correlated

3   with the locality of recombination hot spots in human (Lee *et al*, 2011; Myers *et*

4   *al*, 2005), mice (Wu *et al*, 2010) and pigs (Tortereau *et al*, 2012). DNA motifs

5   (*cis*) have been identified to be associated with recombination hot spots both in

6   humans (Myers *et al*, 2005; Myers *et al*, 2008; Zheng *et al*, 2010) and in other

7   organisms (Comeron *et al*, 2012). In addition to the *cis* elements presented by the

8   motifs, a *trans* element, PRDM9, was found to be a major determinant of

9   recombination hotspots in human and mice (Baudat *et al*, 2010; Brunschwig *et al*,

10   2012) but not in dog and related wild relatives (Auton *et al*, 2013; Munoz-Fuentes

11   *et al*, 2011). PRDM9 is thought to bind to a DNA motif via a zinc finger domain,

12   altering the chromatin structure through methylation, and recruiting

13   recombination molecular machinery (Myers *et al*, 2010).

14   Coarse-scale recombination in cats has previously been investigated

15   through linkage map analyses using sparsely distributed microsatellite markers

16   (Menotti-Raymond *et al*, 1999; Menotti-Raymond *et al*, 2009). The objective of

17   this study is to investigate fine-scale recombination rates, and recombination

18   hotspots, in the domestic cat using population-level data for dense SNPs in ten

19   selected genomic regions on ten different chromosomes.

20   **Materials and methods**
21   **Samples and genotypes**

1   SNP genotype data of twenty-two cats of Eastern random bred origin were

2   obtained as previously described in (Alhaddad *et al*, 2013). The data were

3   generated using a custom illumina GoldenGate array that represent ten different

4   cat chromosomal regions (A1, A2, B3, C2, D1, D2, D4, E2, F2, X) and were

5   composed of 1536 markers nearly equally distributed across the ten regions.

6   Markers' positions were updated to their location in the 6.2 cat genome assembly

7   (http://genome.ucsc.edu/). To ensure successful inference, three criteria were

8   used to filter the SNPs: (i) SNPs mapped to a single chromosomal location on the

9   most recent genome assembly of cat with 100% sequence match, (ii) SNPs

10   exhibiting genotype rate of $\geq 80\%$, and (iii) SNPs possessing a minor allele

11   frequency of ≥ 0.1. The final dataset, included in this study, is composed of 701

12   markers distributed over the ten regions. The genomic locations of the regions

13   and related summaries are shown in Table 1 (see also Figure 1 and Supplementary

14   Information 2).

15   Three SNPs residing in a recombination hotspot on chromosome E2

16   region (see results) were chosen for genotype validation via sequencing. Primers

17   were designed to flank the SNPs (Supplementary Information 2) and PCR was

18   performed using DNA Engine Gradient Cycler (MJ Research, GMI, Ramsey,

19   MN) using the following conditions. For each reaction, 2µl of DNA was used in

20   a 1.5mM magnesium concentration with 1µM primers in a total reaction volume

21   of 20µl. The annealing temperature of all primers was 62°C. The PCR protocol

1   was as follows: initial denaturation at 94°C for 5 min followed by 40 cycles of

2   94°Cx45 sec, 62°Cx20 sec, 72°Cx30 sec, and a final extension at 72°C for 20

3   min. The PCR products were purified with ExoSap (USB, Cleveland, OH) per

4   the manufacture's recommendations and directly sequenced using the BigDye

5   terminator Sequencing Kit v3.1 (Applied Biosystems, Foster City, CA) as

6   previously implemented (Bighignoli *et al*, 2007). Sequences were verified and

7   aligned using the software sequencer version 4.10 (Gene Codes Corp., Ann

8   Arbor, MI).

9   **Recombination inference**

10   The program *inferRho* uses the coalescent with recombination model

11   (Hudson, 1991; Kingman, 1982a; Kingman, 1982b) in a full Bayesian framework

12   to infer recombination rates and hot spots along the chromosome regions (Wang

13   and Rannala, 2008; Wang and Rannala, 2009). The evolutionary relationship of

14   the SNPs is represented by the ancestral recombination graph (ARG), which is an

15   unobserved random variable that is integrated over using Markov Chain Monte

16   Carlo (MCMC). In the variable recombination rate model, the crossing-over rate

17   is $\rho_i = 4N_e c_i$, where N_e is the effective population size, c_i is the recombination rate

18   per generation in cM/Mb between marker i and $i+1$ ($i = 1,\ldots, k-1$, and k is the

19   total number of SNPs). The total recombination rate (c_i) consists of the

20   background crossing-over rate (following a prior gamma distribution) and

1   recombination hotspots (arising according to a Markov process) (Wang and

2   Rannala, 2008; Wang and Rannala, 2009).

3   To make it computationally tractable, each chromosomal region was

4   divided into blocks, with 20 SNPs per block (e.g., chromosome A1 region with 53

5   SNPs was divided into 3 blocks, 20 SNPs in the first and second block and 13

6   SNPs in the third block). Blocks from each chromosomal region are assumed to

7   share the same population size parameter (θ) but have an independent ARG for

8   each block. Preliminary runs were executed to determine the appropriate length

9   and thinning interval of the MCMC chain for each chromosomal dataset. The last

10   half of the MCMC samples (1000 samples) was used to estimate the

11   recombination rates (ρ_i) and to plot them with respect to the SNP marker

12   positions. The distance-weighted mean recombination rate for each chromosomal

13   region was calculated using the following equation:

14
$$\bar{\rho} = \sum_{i=1}^{k-1} \rho_i d_i \bigg/ \sum_{i=1}^{k-1} d_i ,$$

15   in which d_i is the distance between marker i and $i+1$, and k is the total number of

16   markers in the region.

17   **Designation of recombination hot spots**

18   We calculated the Bayes factors (Kass and Raftery, 1995) to locate the

19   positions of hotspots and represent relative odds of a hotspot being present. Each

20   chromosomal region was divided into bins (200 bp per bin) for analysis of

1   posterior samples to estimate the probability of having hotspot (p_j) for each bin.

2   The corresponding probability of hotspot for each bin (q_j) from the prior was

3   obtained by running the program without data (e.g., constant likelihood) but

4   maintaining the same sample size and marker positions. The Bayes factor (BF) is

5   defined as the ratio of the posterior and prior odds:

6
$$BF = \frac{p_j/(1-p_j)}{q_j/(1-q_j)}.$$

7   The total number of bins depends on the length of the region, thus the

8   range of j is variable among the ten chromosome regions. Recombination

9   hotspots are defined to have at least two consecutive bins with $BF \geq 100$ (cf.

10   section 3.2 in Kass&Raftery (1995)).

11   **GC content and genomic elements analyses**

12   The GC content was calculated for the sequences of each bin using the

13   function *CG.content* of package *APE* in R (Paradis *et al*, 2004).

14   Variation and repeat elements within each of the ten chromosomal regions

15   were downloaded from UCSC genome browser using *RepeatMasker* for the v6.2

16   cat genome assembly. Elements were analyzed separately for bins with $BF < 100$

17   and $BF \geq 100$ (hotspot bins). The elements within neutral bins ($BF < 100$) were

18   to provide the general overview of the elements within each chromosomal region.

19   **Results**
20   **Samples and genotypes**

1   The dataset is composed of 701 SNPs on ten different cat chromosomal

2   regions with an average of 70 SNPs/region. The chromosome E2 region harbors

3   the highest number of SNPs ($n = 90$) while chromosome X region has the lowest

4   ($n = 37$). SNPs are distributed across the regions with an average distance

5   between SNPs of 12 Kb and a range of 145 bp – 651 Kb (Table 1). Four SNPs

6   were selected for genotype verification using direct sequencing. The sequencing

7   results showed were concordant to the genotypes obtained previously using the

8   genotyping platform.

9   **Recombination rates and hotspots**

10   The ten regions on chromosomes: A1, A2, B3, C2, D1, D2, D4, E2, F2,

11   and X, were analyzed using *inferRho*. The lengths of the MCMC chains were

12   determined as: 600 000 iterations for A1, C2; 800 000 iterations for A2, D1, D4,

13   F2; 1 000 000 iterations for B3, D2, E2; and 500 000 iterations for X. Two

14   thousand posterior samples for each chromosomal dataset were obtained after

15   thinning, while the first half were discarded as burn-in. Using a parallel

16   computing approach, each run could be accomplished within one to two weeks on

17   a cluster with 2 Opteron 270 (2.0 GHz) processors per node.

18   Population recombination rate (ρ) is plotted for each region (Figure 1 a-c,

19   Supplementary Information 2a). The mean recombination rate (ρ) across all

20   regions is 200 per Mb. The E2 region exhibits the highest mean rate (309 /Mb)

21   whereas X region has the lowest (77 /Mb) (Table 1). The difference between the

1   background recombination rate and whole recombination rate is indicative of

2   recombination spots. This difference is most noticeable in regions of

3   chromosomes A2, D1, and E2 (Figure 1 a-c).

4   The posterior probability of hotspot was calculated for bins of size 200 bp

5   in each region (Figure 1 d-f, Supplementary Information 2b). Chromosomes A2,

6   D1, and E2 show distinctly high posterior probabilities (> 0.6) in four localized

7   areas (Figure 1 d-f). The posterior probabilities for the other seven chromosome

8   regions (A1, B3, C2, D2, D4, F2, and X) are less than 0.2, indicating little support

9   for elevation of recombination rate in these areas.

10   The Bayes factors (Figure 1 g-i, Supplementary Information 2c) are

11   consistent in pattern with the posterior probabilities (Figure 1 d-f, Supplementary

12   Information 2b) and recombination rates (Figure 1 a-c, Supplementary

13   Information 2a). Approximately 99% ($n = 42{,}531$) of bins were classified as

14   "neutral" ($BF < 100$) across all regions examined. The hot spots were found in

15   only three chromosomal regions (A2, D1, and E2) and represented $\sim 0.13\%$ ($n =$

16   57) of all bins studied. Summaries of the numbers and distribution of bins within

17   each chromosomal region are provided in Table 1. The hotspots had size of 3 Kb

18   in A2, 1.8 Kb in D1, and 1.8 Kb for the first and 4.6 Kb for the second in E2. The

19   distance between the two hotspots in E2 was ~37.4 Kb (Table 2).

20   **GC content analysis**

1   The $\log_{10}$ of the Bayes factor was plotted as a function of the GC content

2   in Supplementary Information 3. Pearson's correlation test revealed a positive

3   correlation between GC and $\log_{10}$(Bayes factor) (cor $= 0.077$, $p < 0.0001$) but was

4   not suggestive of strong correlation. Moreover, no significant differences in the

5   mean GC content of each class of bins were observed (t-test, $p = 0.05$)

6   (Supplementary Information 3). The mean GC contents of the hotspots are shown

7   in Table 2.

8   **Repeat elements analysis**

9   The four hotspot regions contained 22 repeat and variation elements

10   (Table 2, Figure 2a). SINE elements constitute the highest proportion (40%) of

11   the elements present in the hotspots followed by LINE elements (27%). Within

12   LINE elements, L2 elements were present in three of the four hotspot regions.

13   MIR family elements were present in all hotspot regions and tRNA-Lys family

14   elements are present in three of the four elements. Low complexity, long terminal

15   repeats, simple repeats, and DNA elements were inconsistently present across the

16   hotspot regions.

17   Variation and repeat elements within neutral regions where investigated to

18   get a general view of the general distribution of repeat elements (Figure 2b). In

19   neutral regions, the SINE elements represent the highest proportion of elements,

20   34 %. The tRNA-Lys SINE elements constitute 22% and MIR SINE elements

21   represent 12%. The second highest repeat elements was the LINE elements, 29%,

1   where L1 elements represent ~17% and L2 elements represent 9.5% of all

2   elements in the neutral regions.

3   **Discussion**

4   Advances in population genetic theory and technology have made it

5   possible to estimate recombination rates directly from genotype data on

6   population samples, overcoming the limitations of sperm typing or using large

7   extended families (Hellenthal and Stephens, 2006). The strategy of the

8   population genetics based approach is to use information on the number of

9   recombination events that have occurred in the history of the population, which

10   can be detected by modeling the patterns of genetic variation expected to be

11   present in randomly selected individuals.

12   The model accounts for coalescent and recombination events in an

13   Ancestral Recombination Graph (ARG). Markers will have coalescent trees that

14   are likely to vary across the genome. In theory, all markers in a chromosome are

15   correlated by an ARG. However, the size of the ARG might grow much faster

16   than linear with the increasing number of SNPs, making it computational

17   intractable. In practice, the data are usually partitioned into blocks. Blocks from

18   the same chromosomal region are assumed to share the same population size

19   parameter (θ) but have an independent ARG for each block. For fast

20   computation, some methods use an approximate likelihood instead of the full

21   likelihood calculation, for example, the composite-likelihood method

1   implemented in *LDhat* (McVean *et al*, 2002) and PAC-likelihood method

2   implemented in *PHASE* (Stephens and Donnelly, 2003). Approximate-likelihood

3   methods may be feasible to apply to large genomic regions, but may lack power

4   to detect a moderate or low rate of recombination. Full-likelihood method in

5   *inferRho* uses all of the information in the data and should therefore provide more

6   accurate estimates (Wang and Rannala, 2008; Wang and Rannala, 2009).

7   The analysis of recombination hotspots in cats, presented here, constitutes

8   the first application of the program *inferRho* to non-human data and the first

9   analysis of fine-scale recombination rates in cats. The population recombination

10   rate was found to be variable between the regions analyzed and, as expected, the

11   mean recombination rate of X chromosome regions was lower than that of any

12   autosomal regions. This variation in recombination rates and the notable reduced

13   rate outside of the pseudoautosomal region of the X chromosome are in

14   agreement with observations of recombination in human (Myers *et al*, 2005) and

15   dog (Axelsson *et al*, 2012). The latter result is expected due to that fact that

16   recombination outside of the pseudoautosomal region occurs only in females.

17   Four decisive hot spots were identified on three chromosomes: A2, D1 and

18   E2. The localities of the hotspots are in agreement with the localities of increased

19   posterior probabilities and the general topography of recombination rates as

20   expected. Acknowledging the limitation posed by the total size of the regions

21   analyzed (total ~ 8.5 Mb) compared to the size of the genome and the lack of

1   power to perform correlation and element enrichment analyses, the following

2   observations have been made: (i) Unlike hotspots in humans (Myers *et al*, 2005),

3   dog (Axelsson *et al*, 2012), pig (Tortereau *et al*, 2012), and chicken (Groenen *et*

4   *al*, 2009), which exhibit positive correlation with GC content, the hotspots in cats,

5   identified in this study, show no distinct positive correlation with GC content (this

6   is similar to mice) (Wu *et al*, 2010). (ii) Genomic features in the four hot spots

7   are consistent with those found in other mammals. Three of the decisive hotspots

8   contained at least one L2 LINE element, which is consistent with features of

9   hotspots in human and chimpanzee (Lee *et al*, 2011; Myers *et al*, 2010; Myers *et*

10   *al*, 2008). Similarly, MIR and tRNA-Lys SINE elements were found in hotspots

11   of cats, which is consistent with the positive correlation found for hotspots in

12   human (Lee *et al*, 2011), and pig (Tortereau *et al*, 2012). (iii) The general

13   similarity of the hotspot repeat elements to those of other mammals suggests

14   similar recombination mechanism and probable involvement of PRDM9. (iv)

15   This similarity suggests there may be evolutionary differences between cats and

16   dogs, especially since recent studies found that PRDM is not functional in dog

17   (Axelsson *et al*, 2012; Munoz-Fuentes *et al*, 2011) and that recombination

18   hotspots in dogs are associated with CpG and promoter regions (Auton *et al*,

19   2013).

20   This study represents a glimpse of recombination hotspots in cats and only

21   an initial step toward understanding recombination in cats. As cat resources

1   developing, genome-wide analyses may be performed and more definitive

2   conclusions will reach. Nonetheless, understanding the recombination landscape

3   will shed light on the mechanism of recombination in cats compared to other

4   species, the pattern of variation generated by recombination in cats, and should

5   lead to better implementation of efficient disease mapping strategies in cats.

1 **Reference**

2

3 Alhaddad H, Khan R, Grahn RA, Gandolfi B, Mullikin JC, Cole SA *et al* (2013).

4 Extent of linkage disequilibrium in the domestic cat, Felis silvestris catus, and its

5 breeds. *Plos One* **8**(1)**:** e53537.

6

7 Auton A, Rui Li Y, Kidd J, Oliveira K, Nadel J, Holloway JK *et al* (2013).

8 Genetic recombination is targeted towards gene promoter regions in dogs. *ArXiv*

9 *e-prints* **arXiv:1305.6485**.

10

11 Axelsson E, Webster MT, Ratnakumar A, Ponting CP, Lindblad-Toh K (2012).

12 Death of PRDM9 coincides with stabilization of the recombination landscape in

13 the dog genome. *Genome Res* **22**(1)**:** 51-63.

14

15 Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M *et al* (2010).

16 PRDM9 is a major determinant of meiotic recombination hotspots in humans and

17 mice. *Science* **327**(5967)**:** 836-840.

18

19 Bighignoli B, Niini T, Grahn RA, Pedersen NC, Millon LV, Polli M *et al* (2007).

20 Cytidine monophospho-N-acetylneuraminic acid hydroxylase (CMAH) mutations

21 associated with the domestic cat AB blood group. *BMC Genet* **8:** 27.

22

1 Brunschwig H, Levi L, Ben-David E, Williams RW, Yakir B, Shifman S (2012).

2 Fine-scale Map of Recombination Rates and Hotspots in the Mouse Genome.

3 *Genetics* **112.141036**

4

5 Comeron JM, Ratnappan R, Bailin S (2012). The many landscapes of

6 recombination in Drosophila melanogaster. *PLoS Genet* **8**(10)**:** e1002905.

7

8 Groenen MA, Wahlberg P, Foglio M, Cheng HH, Megens HJ, Crooijmans RP *et*

9 *al* (2009). A high-density SNP-based linkage map of the chicken genome reveals

10 sequence features correlated with recombination rate. *Genome Res* **19**(3)**:** 510-

11 519.

12

13 Hellenthal G, Stephens M (2006). Insights into recombination from population

14 genetic variation. *Curr Opin Genet Dev* **16**(6)**:** 565-572.

15

16 Hudson RR (1991). Gene genealogies and the coalescent process. *Oxford Survey*

17 *in Evolutionary Biology* **7:** 1-44.

18

19 Jeffreys AJ, Murray J, Neumann R (1998). High-resolution mapping of

20 crossovers in human sperm defines a minisatellite-associated recombination

21 hotspot. *Mol Cell* **2**(2)**:** 267-273.

1

2   Kass RE, Raftery AE (1995). Bayes Factors. *Journal of the American Statistical*

3   *Association* **90**(430)**:** 773-795.

4

5   Kingman JFC (1982a). The coalescent. *Stochastic Processes and their*

6   *Applications* **13**(3)**:** 235-248.

7

8   Kingman JFC (1982b). On the Genealogy of Large Populations. *Journal of*

9   *Applied Probability* **19:** 27-43.

10

11   Lee YS, Chao A, Chen CH, Chou T, Wang SY, Wang TH (2011). Analysis of

12   human meiotic recombination events with a parent-sibling tracing approach. *BMC*

13   *Genomics* **12:** 434.

14

15   McVean G, Awadalla P, Fearnhead P (2002). A coalescent-based method for

16   detecting and estimating recombination from gene sequences. *Genetics* **160**(3)**:**

17   1231-1241.

18

19   McVean GA, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P (2004).

20   The fine-scale structure of recombination rate variation in the human genome.

21   *Science* **304**(5670)**:** 581-584.

1

2   Menotti-Raymond M, David VA, Lyons LA, Schaffer AA, Tomlin JF, Hutton

3   MK *et al* (1999). A genetic linkage map of microsatellites in the domestic cat

4   (Felis catus). *Genomics* **57**(1)**:** 9-23.

5

6   Menotti-Raymond M, David VA, Schaffer AA, Tomlin JF, Eizirik E, Phillip C *et*

7   *al* (2009). An autosomal genetic linkage map of the domestic cat, Felis silvestris

8   catus. *Genomics* **93**(4)**:** 305-313.

9

10 Munoz-Fuentes V, Di Rienzo A, Vila C (2011). Prdm9, a major determinant of

11 meiotic recombination hotspots, is not functional in dogs and their wild relatives,

12 wolves and coyotes. *Plos One* **6**(11)**:** e25498.

13

14 Myers S, Bottolo L, Freeman C, McVean G, Donnelly P (2005). A fine-scale map

15 of recombination rates and hotspots across the human genome. *Science*

16 **310**(5746)**:** 321-324.

17

18 Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS *et al*

19 (2010). Drive against hotspot motifs in primates implicates the PRDM9 gene in

20 meiotic recombination. *Science* **327**(5967)**:** 876-879.

21

1   Myers S, Freeman C, Auton A, Donnelly P, McVean G (2008). A common

2   sequence motif associated with recombination hot spots and genome instability in

3   humans. *Nat Genet* **40**(9)**:** 1124-1129.

4

5   Paigen K, Petkov P (2010). Mammalian recombination hot spots: properties,

6   control and evolution. *Nat Rev Genet* **11**(3)**:** 221-233.

7

8   Paradis E, Claude J, Strimmer K (2004). APE: Analyses of Phylogenetics and

9   Evolution in R language. *Bioinformatics* **20**(2)**:** 289-290.

10

11   Steinmetz M, Minard K, Horvath S, McNicholas J, Srelinger J, Wake C *et al*

12   (1982). A molecular map of the immune response region from the major

13   histocompatibility complex of the mouse. *Nature* **300**(5887)**:** 35-42.

14

15   Stephens M, Donnelly P (2003). A comparison of bayesian methods for haplotype

16   reconstruction from population genotype data. *Am J Hum Genet* **73**(5)**:** 1162-

17   1169.

18

19   Tortereau F, Servin B, Frantz L, Megens HJ, Milan D, Rohrer G *et al* (2012). A

20   high density recombination map of the pig reveals a correlation between sex-

21   specific recombination and GC content. *BMC Genomics* **13:** 586.

1

2   Wang Y, Rannala B (2008). Bayesian inference of fine-scale recombination rates

3   using population genomic data. *Philos Trans R Soc Lond B Biol Sci* **363**(1512)**:**

4   3921-3930.

5

6   Wang Y, Rannala B (2009). Population genomic inference of recombination rates

7   and hotspots. *Proc Natl Acad Sci U S A* **106**(15)**:** 6215-6219.

8

9   Wu ZK, Getun IV, Bois PR (2010). Anatomy of mouse recombination hot spots.

10 *Nucleic Acids Res* **38**(7)**:** 2346-2354.

11

12 Zheng J, Khil PP, Camerini-Otero RD, Przytycka TM (2010). Detecting sequence

13 polymorphisms associated with meiotic recombination hotspots in the human
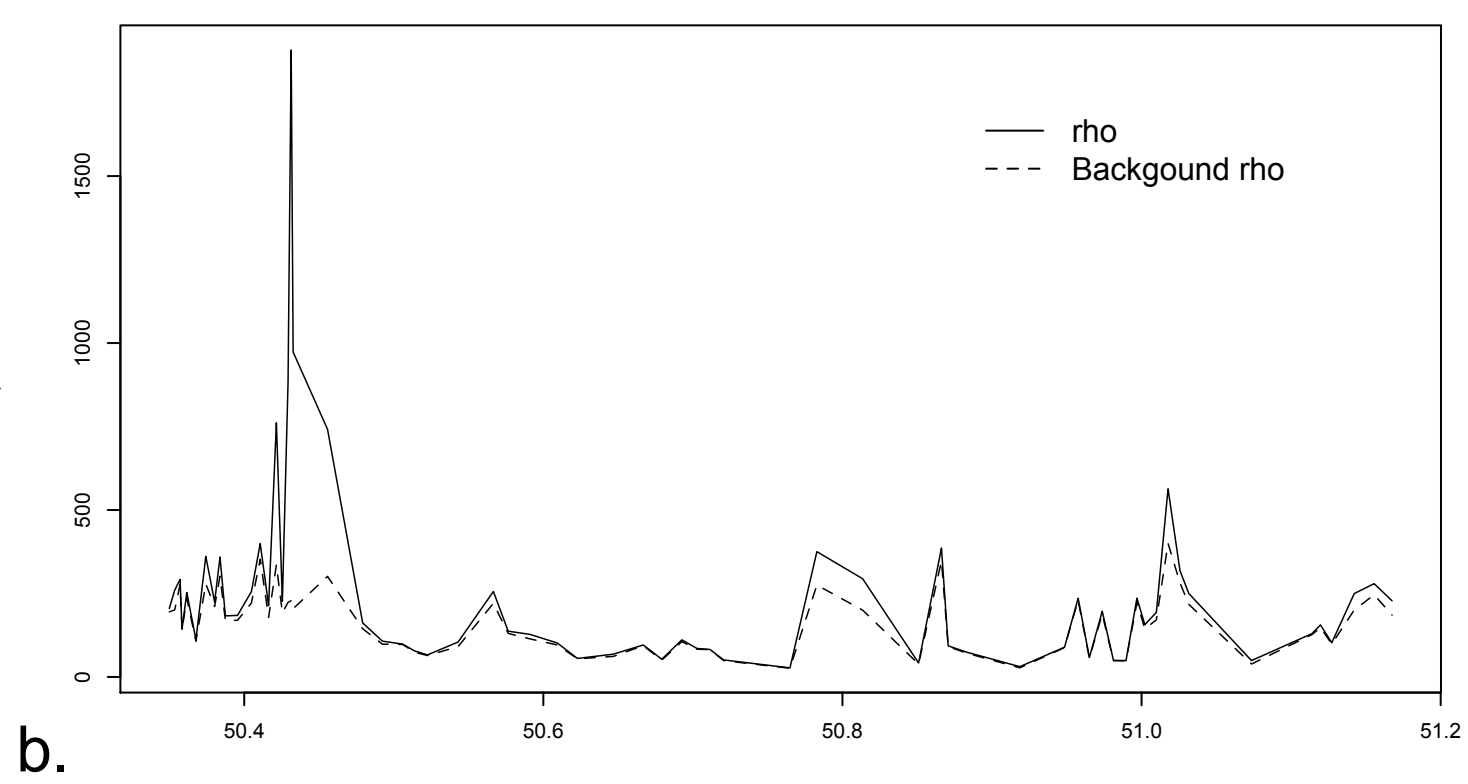
14 genome. *Genome Biol* **11**(10)**:** R103.
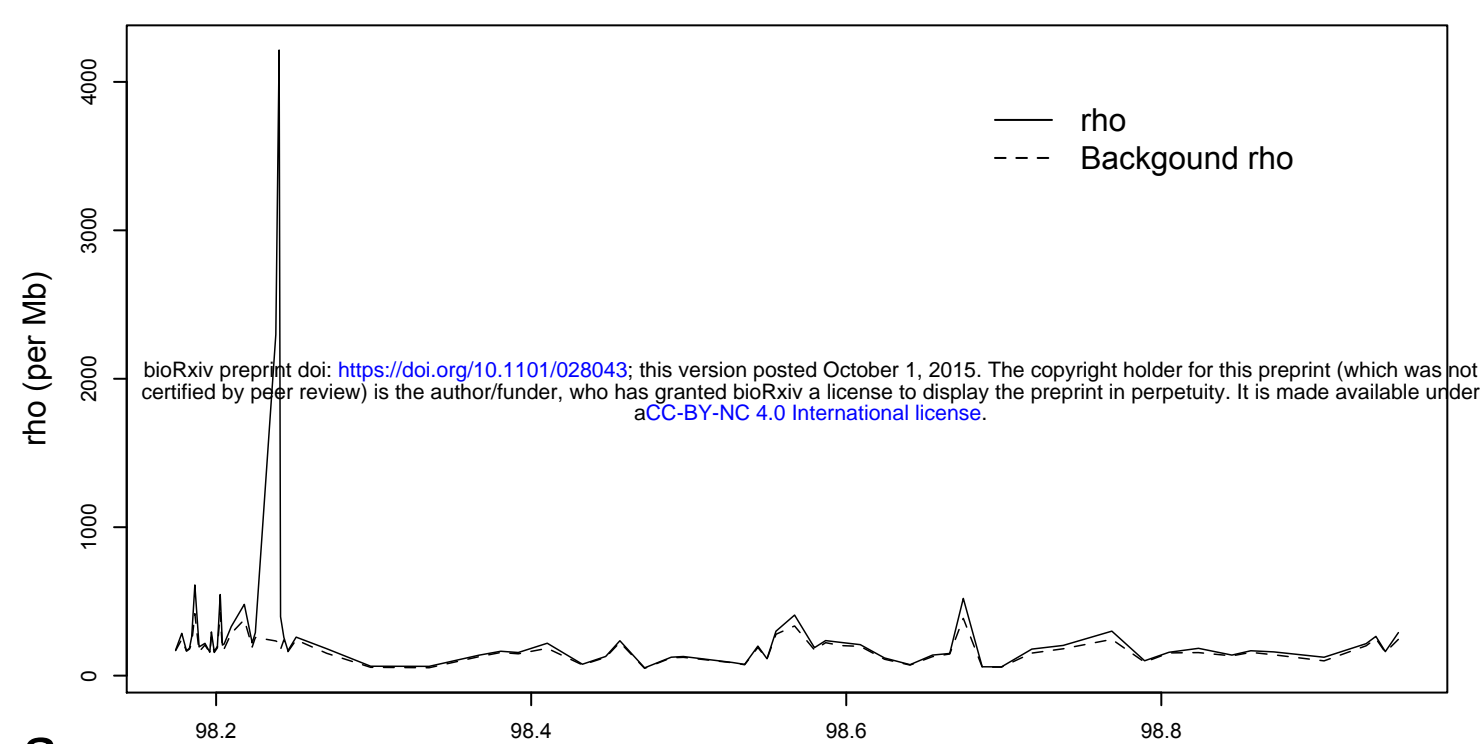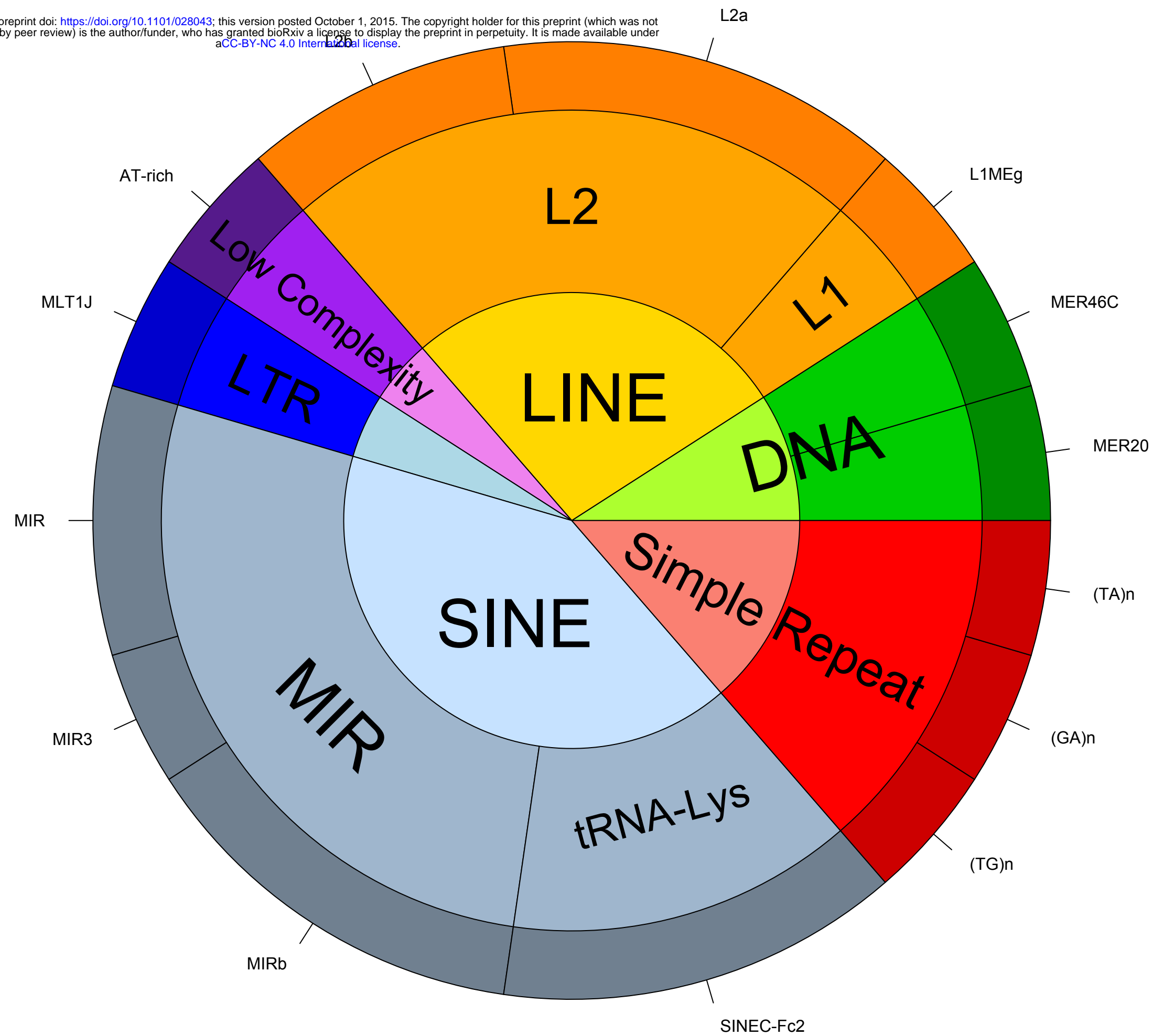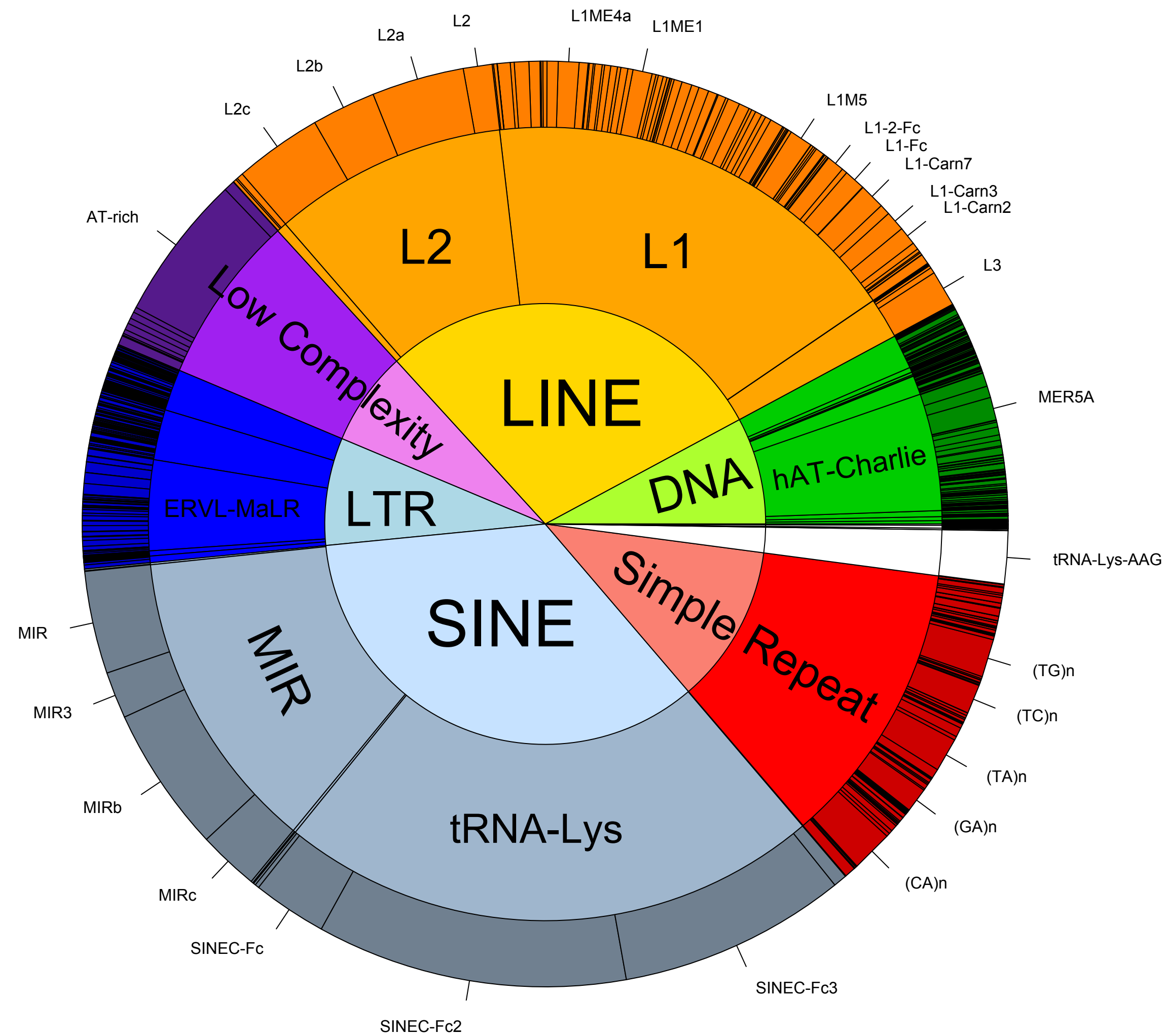
15

16

17

1 **Figure legends**

2 **Figure 1** Recombination overview of chromosomes A2, D1, and E2 regions. (a-

3 c) Posterior recombination rates (population size scaled) across chromosomes A2,

4 D1, and E2 regions, respectively. Solid line shows the whole recombination rates,

5 while the dashed line shows the background recombination rates. (d-f) Posterior

6 probability of hotspots along chromosomes A2, D1, and E2 regions, respectively.

7 (g-i) Bayes factor of hotspots for chromosomes A2, D1, and E2 regions,

8 respectively. Horizontal dotted line corresponds to Bayes factor of 100 in a $\log_{10}$

9 scale. Position and distribution of SNPs included in the recombination analysis

10 are at the bottom.

11

12 **Figure 2** Overview of variation and repeat elements in the cat chromosomal

13 regions studied. a) Variation and repeat element ($n = 22$) within four hotspots.

14 All elements within the hotspots are shown at the tips of the outer circle. b)

15 Variation and repeat element ($n = 16,798$) within "neutral" bins. Elements

16 present >100 within the warm-spots are shown at the tips of the outer circle. The

17 inner circle represents the elements' classes, middle circle represents elements'

18 families, and outer circle represent the individual elements.

a.

b.