Recombination patterns reveal information about centromere location on linkage maps

Limborg, Morten T.; McKinney, Garrett J.; Seeb, Lisa W.; Seeb, James E.

Published in:
Molecular Ecology Resources

Link to article, DOI:
10.1111/1755-0998.12484

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2016) **16**, 655–661 doi: 10.1111/1755-0998.12484

Recombination patterns reveal information about centromere location on linkage maps

MORTEN T. LIMBORG,*†‡ GARRETT J. MCKINNEY,* LISA W. SEEB* and JAMES E. SEEB*
*School of Aquatic and Fishery Sciences, University of Washington, 1122 NE Boat Street, Box 355020, Seattle, WA 98195, USA,
†National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark

Abstract

Linkage mapping is often used to identify genes associated with phenotypic traits and for aiding genome assemblies. Still, many emerging maps do not locate centromeres – an essential component of the genomic landscape. Here, we demonstrate that for genomes with strong chiasma interference, approximate centromere placement is possible by phasing the same data used to generate linkage maps. Assuming one obligate crossover per chromosome arm, information about centromere location can be revealed by tracking the accumulated recombination frequency along linkage groups, similar to half-tetrad analyses. We validate the method on a linkage map for sockeye salmon (*Oncorhynchus nerka*) with known centromeric regions. Further tests suggest that the method will work well in other salmonids and other eukaryotes. However, the method performed weakly when applied to a male linkage map (rainbow trout; *O. mykiss*) characterized by low and unevenly distributed recombination – a general feature of male meiosis in many species. Further, a high frequency of double crossovers along chromosome arms in barley reduced resolution for locating centromeric regions on most linkage groups. Despite these limitations, our method should work well for high-density maps in species with strong recombination interference and will enrich many existing and future mapping resources.

Keywords: centromeres, genomic architecture, genotyping by sequencing, linkage mapping, recombination

Received 1 July 2015; revision received 29 October 2015; accepted 4 November 2015

Introduction

High-throughput sequencing technology has greatly accelerated the construction of dense linkage maps in nonmodel species (e.g. Amores *et al.* 2011; Miller *et al.* 2012; Everett & Seeb 2014). The emerging maps allow unprecedented insights into the genomic architecture of adaptive divergence in the wild (e.g. Chutimanitsakun *et al.* 2011; Gagnaire *et al.* 2013; Richards *et al.* 2013). These studies catalyse our understanding of the number, genomic location and colocation of loci affected by natural selection. Unfortunately, centromeres are rarely included on linkage maps because the additional mapping efforts such as mapping half-tetrads add significant logistical hurdles (cf., Thorgaard *et al.* 1983; Brieuc *et al.* 2014).

Centromeres represent a fundamental component of chromosomal structure and function (Henikoff *et al.* 2001), and information about centromere location is vital for properly understanding genomes. Studies describing genetic divergence have shown striking patterns in either centromeric or telomeric regions (Carneiro *et al.* 2009; Ellegren *et al.* 2012). Hence, knowledge about chromosome type (i.e. acrocentric vs metacentric) is of paramount importance for understanding the underlying architecture of adaptive traits. However, interpretations often suffer from the lack of known centromeres on reference maps and genomes; this has impoverished interpretations of results plotted along high-density maps (e.g. Wang *et al.* 2012; Gagnaire *et al.* 2013; Carlson *et al.* 2015) or genome assemblies (e.g. Ellegren *et al.* 2012; Tine *et al.* 2014; Xu *et al.* 2014).

Here, we demonstrate a straightforward method to identify centromeric regions on linkage maps by phasing the same recombination data used to construct the map. We validate the method by comparing phased centromere placement with more direct centromere placement using half-tetrad analysis in sockeye salmon, *Oncorhynchus nerka*. Finally, we provide test examples that highlight advantages and limitations of the method for mapping centromeres in different sexes and taxa.

Correspondence: Morten T. Limborg, Fax: 206-685-7471; E-mail: morten.limborg@gmail.com
‡ Present address: Centre for GeoGenetics, University of Copenhagen, 1350 Copenhagen K, Denmark

Materials and methods

Mapping data

We used genotype data from the mapping families presented in Limborg *et al.* (2015) to illustrate how consideration of recombination patterns in mapping data can be used to infer both chromosome type and location of centromeric regions. Two family data sets were generated from a single female sockeye salmon. *Family 1* consisted of 77 gynogenetic diploid offspring genotyped for 2562 loci that were used to identify centromeric regions through half-tetrad analysis (Thorgaard *et al.* 1983). *Family 2* consisted of a gynogenetic haploid family (n = 94) genotyped for 3245 loci that were used to construct a linkage map. Loci were scored using RAD sequencing and 5′-nuclease genotyping (Limborg *et al.* 2015).

Identifying chromosome type and centromeres

Here, we phased the data available from the haploid *Family 2* to demonstrate an alternative approach for placing centromeres. Markers that are polymorphic within a mapping parent contain information about the allelic phase within that parent. After the construction of a linkage map, one can plot, for each offspring, the underlying phased genotypes along each linkage group (Fig. 1). This framework allows the observation of individual recombination events along linkage groups. We can then estimate a proxy for the recombination frequency between the terminal markers and each subsequent marker along a linkage group. We obtain recombination frequencies (RF_m) for intervals between each marker (m) and the two terminal loci on a linkage group. Thus, along the entire length of each linkage group, we plot two patterns of RF_m, one proceeding from each end, with each terminal marker assigned as a reference starting point (RF_m = 0).

Interpretation of RF_m

Certain assumptions are needed for the phasing method to work properly. First, strong recombination interference along chromosome arms should occur. Also, crossovers are needed on each side of the centromere in metacentric chromosomes. These assumptions are generally met in the sockeye salmon linkage map (Limborg *et al.* 2015) and lead us to expect exactly one crossover per arm in most meiotic products. Hence, the method should apply to any eukaryote genome characterized by strong recombination interference along chromosome arms.

It is important to consider the recombinant type of meiotic products captured in the form of haploid offspring. In haploid offspring, all four meiotic products

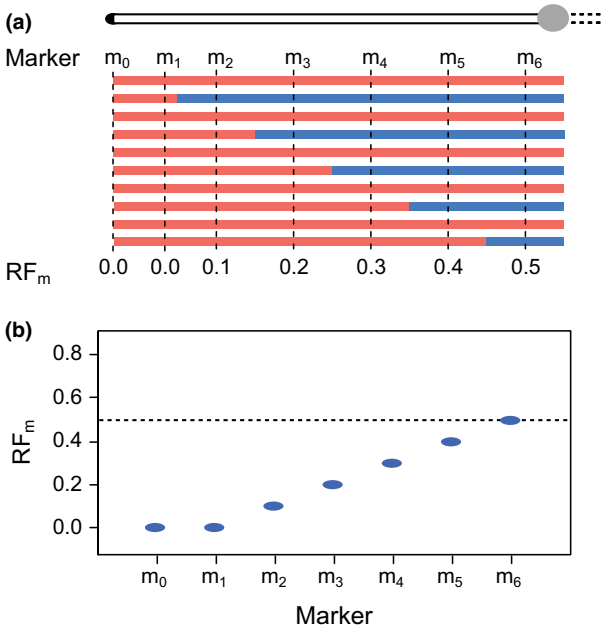

Fig. 1 Conceptual illustration of using phased genotypes to estimate recombination frequency (RF_m). (a) Illustration of how to calculate RF_m along a single chromosome arm from phased genotype data. The ten horizontal bars represent phased genotypes along a single chromosome arm for ten haploid individuals. Alternate phases are shown in red and blue. Starting at one terminal marker (m_0) and moving along the linkage group, RF_m record the fraction of offspring with an observed crossover (i.e. phase shift) between m_0 and all subsequent markers ($m_{1–6}$). (b) Resulting plot of the RF_m values estimated in (a) above. When capturing all four meiotic products individually, as done in haploid mapping, RF_m values are expected to vary from 0.0 at m_0 to ~0.5 when a full chromosome arm has been covered because only half the offspring will be recombinants.

are represented; however, almost always only two of the four chromatids on each side of the centromere engage in recombination during meiosis (Roeder 1997). We therefore expect ~50% of haploid gametes to represent meiotic recombinants for each chromosome arm. When scoring values of RF_m along a linkage group, recombination events are recorded whenever a phase shift compared to the terminal reference marker (m_0) is observed (Fig. 1a,b). With strong interference, this should translate into values of RF_m ranging from 0.0 at the reference marker (m_0) to ~0.5 at the opposite end of the same arm. We are interested in identifying linkage regions corresponding to individual arms; therefore, we round RF_m estimates >0.5 down to 0.5 to simplify the identification of centromeres by smoothing plotted RF_m patterns.

We further expect that acrocentric chromosomes should be distinguishable from metacentric chromosomes even when no a priori information about centromere location exists. For metacentric chromosomes,

the phase of the terminal reference marker is expected to be independent from phases of most markers on the arm opposite the centromere due to a general lack of crossover interference along entire metacentric chromosomes (Colombo & Jones 1997; Demarest *et al.* 2011; Limborg *et al.* 2015). Therefore, for acrocentric chromosomes, RF_m will be 0.0 at the terminal reference marker and increases linearly towards a value of 0.5 at the opposite end (see pattern in Fig. 1b) – regardless of which reference marker represents the centromere. For metacentric chromosomes, plots of RF_m should translate into a flat pattern where RF_m plateau along the chromosome arm opposite the terminal reference marker reflecting random phasing on the opposite arm in relation to the reference marker (i.e. $RF_m \sim 0.5$).

By scoring RF_m from both ends of a linkage group, centromeres on metacentric chromosomes are then expected to reside in regions on the linkage group where values of RF_m counted from either end intersect around a value of 0.5. Plotting RF_m alone is thus expected to efficiently distinguish metacentric from single-armed (acro- or telocentric) chromosomes although RF_m alone will not identify the centromeric end on single-armed chromosomes. The resolution of most linkage maps in nonmodel species may prevent distinction between telocentric and acrocentric chromosomes (Levan *et al.* 1964). In the following, we assume that chromosomes always have a short p arm and refer to single-armed chromosomes as acrocentric (Levan *et al.* 1964). In conclusion, we expect plots of RF_m to broadly distinguish acrocentric from metacentric chromosomes as well as to identify centromeric regions on metacentric chromosomes.

The paired families of Limborg *et al.* (2015) offered a perfect opportunity to compare the phasing approach of centromere mapping to known centromeric regions defined from the more direct inference from half-tetrads.

Application of RF_m to other example data sets

We applied the phasing method to genotype data from published linkage maps that lacked information about centromere locations. First, we demonstrated the method in a map for pink salmon, *O. gorbuscha* (Limborg *et al.* 2014), based on gynogenetic haploid mapping data similar to that in haploid *Family* 2 for sockeye salmon presented above. We then applied the method to a male map for rainbow trout, *O. mykiss* (Miller *et al.* 2012), built from androgenetic doubled haploids. Lastly, we plotted RF_m patterns for a plant genome using the doubled haploid barley, *Hordeum vulgare*, family presented in Chutimanitsakun *et al.* (2011).

Results

We plotted paired RF_m estimates for each linkage group. Comparing plots of RF_m to centromeric regions previously defined in Limborg *et al.* (2015) revealed patterns that reflected acrocentric chromosomes (paired straight lines; Fig. 2a) or metacentric chromosomes (mirrored hockey stick shapes; Fig. 2b). This pattern was consistent across all linkage groups (Fig. S1, Supporting information). These results clearly illustrate the usefulness of considering RF_m to distinguish acrocentric from metacentric chromosomes in a female salmonid linkage map. For metacentric chromosomes, regions where RF_m estimated from either telomeric end converge around a value of 0.5 define centromeric regions with comparable power to the more direct inference from half-tetrad analyses (Figs 2 and S1, Supporting information).

Although our method worked well with data from a single female sockeye salmon, this does not mean that it will work well in all circumstances. Tests on linkage maps from other species with a more diverse set of

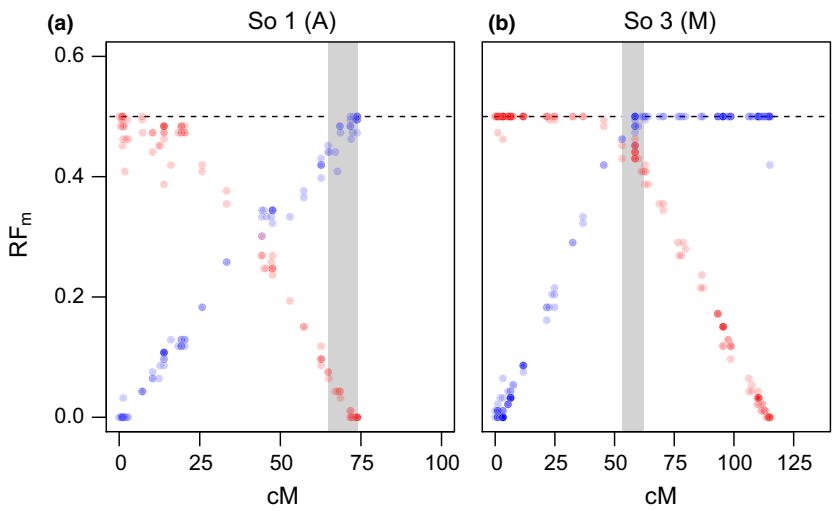

Fig. 2 Plots of RF_m counted from the left (cM = 0; blue circles) and from the right (cM = 75; red circles). The two plots illustrate: (a) the acrocentric (A) So1, and (b) the metacentric (M) So3 linkage groups. Grey bars depict known centromeric regions as identified in Limborg *et al.* (2015). For these haploid family data, RF_m values are expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm.

chromosomal characteristics revealed that our method works for some mapping data, whereas other data sets presented some limitations.

First, when we applied the method to the recently published map for pink salmon (Limborg *et al.* 2014), RF_m patterns clearly depict a karyotype dominated by metacentric chromosomes (Fig. S2, Supporting information), consistent with the known pink salmon karyotype (Phillips & Rab 2001). This high-resolution map was based on a gynogenetic haploid family in a species with strong interference.

Second, the rainbow trout male map was characterized by generally short linkage groups and with many loci binned to the same map location. Consequently, we also observed a reduced resolution for the RF_m method (Figs 3a and S3, Supporting information). At best, weak inference about chromosome type can be made for some of the larger, presumably metacentric, linkage groups (e.g. LGs WS01-WS08), but the RF_m patterns remain inconclusive for most linkage groups in this male map.

Finally, in the barley map, most linkage groups were longer than 150 cM in accordance with reduced levels of interference; this resulted in patterns of wide regions characterized by RF_m values of 0.5 because estimates from both sides plateaued after approximately 50 cM from either end (Figs 3b and S4, Supporting information). One linkage group (4H) was shorter (~130 cM), with putatively fewer double crossovers on each arm, and revealed a more satisfactory resolution for defining the putative centromeric region (Fig. S4, Supporting information). Estimates of RF_m for all data sets are given in Table S1 (Supporting information).

Discussion

Here, we demonstrate a method to identify centromeres on metacentric chromosomes with similar accuracy to that obtained from half-tetrad analyses. Our RF_m method provides a straightforward approach to improve many linkage maps by providing information on chromosome type and centromere location using existing data. Comparing the RF_m method with half-tetrad data for sockeye salmon revealed high precision for locating centromeres on metacentric chromosomes. The method also provided convincing results for a previously published linkage map for pink salmon that lacked information about the location of centromeric regions. Lastly, although we validated the method for a haploid family, the method will also work in diploid families where both parents contribute alleles to offspring. In diploid families, estimates of RF_m can be obtained in a similar fashion if restricting analyses to loci that only segregate in a single parent (e.g. ♀ = AB, ♂ = AA), where the allelic phase of the segregating parent can be unequivocally phased for estimating RF_m.

Some analyses may require a concrete definition of centromeric regions (cf., Limborg *et al.* 2015), which will necessitate designation of a general threshold value of RF_m to define centromeric locations. In our sockeye salmon example, defining a region flanked by the first markers reaching a value of $RF_m \sim 0.45$ counted from each end seems to compare well with the regions defined based on half-tetrad analyses (Fig. S1, Supporting information). However, the optimal threshold value is likely to vary depending on the resolution and density of a given map.

Comparison to alternative methods

The phasing method offers a cost-effective and straightforward approach for including information about centromere locations on both existing and future linkage maps. Performing half-tetrad analyses to directly estimate gene–centromere distances is nontrivial. Historically, generation of mapping and half-tetrad families have been split into independent studies (e.g. Lindner

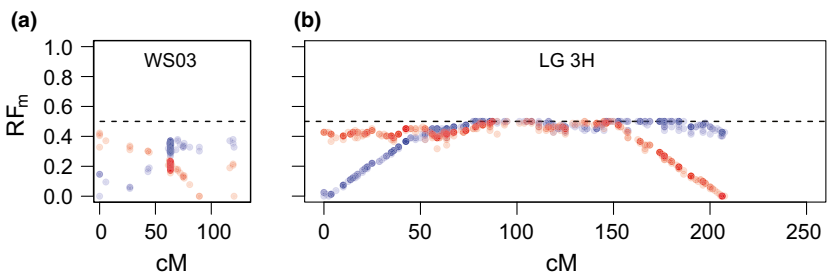

Fig. 3 Examples with data where the RF_m method has limitations for locating centromeres. Graphs show RF_m plotted along a linkage group for a male rainbow trout (a) and along a barley linkage group (b). RF_m values are counted from the left (cM = 0; blue circles) and from the right (red circles). Both the rainbow trout and the barley linkage maps were constructed from doubled haploid families; RF_m values are therefore expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm or when at least one crossover has occurred in each offspring.

et al. 2000; Morishima *et al.* 2001; Reid *et al.* 2007; Li *et al.* 2008) or, in a few cases, relied on large and expensive data sets including both types of crosses (Brieuc *et al.* 2014).

Other strategies to identify centromeres include cytogenetic methods targeting heterochromatin proteins known to encapsulate centromeres (Fransz *et al.* 2000; Kelly *et al.* 2006). However, this method also provides additional hurdles in terms of the expertise and analyses needed. The added complexity and cost of most alternative approaches means that many maps for nonmodel species may continue to lack information about centromere locations.

The RF_m method can reduce the effort for placing centromeres on acrocentric chromosomes. Large families of half-tetrads were used in the studies of Brieuc *et al.* (2014) and Limborg *et al.* (2015) to obtain adequate resolution for locating centromeric regions. Considering that the resolution provided by the haploid mapping family to locate centromeres was similar to that obtained with the half-tetrads, we argue that a much smaller sample size of half-tetrads would suffice to simply distinguish centromeric ends from telomeric ends on acrocentric chromosomes. Alternatively, if centromere placement exists for a related species, or population, then centromeres on acrocentric chromosomes may be inferred by combining comparative mapping (cf., Kodama *et al.* 2014; McKinney *et al.* 2015) with information from RF_m.

Limitations of the method

It should be noted that the RF_m method may fail to detect centromeres on small metacentric chromosomes characterized by short arms (i.e. $\ll$50 cM) where many meiotic products may completely lack observed crossovers along some arms. This is exemplified by the small metacentric linkage group So28 in the sockeye salmon map (Fig. S1, Supporting information).

In many species, one sex often shows reduced recombination rates (Lagercrantz & Lydiate 1995; Broman *et al.* 1998), which can lead to low resolution of linkage maps. Maps with low resolution are often characterized by having many loci binned to the same location, and this obstructs the resolution of RF_m. In salmonids, for example, males exhibit lower recombination rates compared to females (e.g. Lien *et al.* 2011; Everett *et al.* 2012), and when we plotted RF_m along a male linkage map for rainbow trout, signals of centromere locations were vague. Indeed, to further test the generality of such sex difference, we further estimated RF_m for a diploid family of Chinook salmon (*O. tshawytscha*) using data from McKinney *et al.* (2015). Female recombination data provided an adequate resolution compared to the haploid female map for pink salmon, whereas the male map proved inadequate for inferring centromeric locations (data not shown). We conclude that for diploid crosses, RF_m only appears reliable when the female parent, or the heterogametic sex, is used.

Furthermore, the RF_m method will also have reduced resolution for chromosomes characterized by long arms ($\gg$50 cM) with incomplete interference such as in the barley example. In those cases, the value of RF_m will level off after ~50 cM because the rate of double crossovers steadily increases when moving along an arm, obscuring signalling of the actual centromere location.

Hence, the method appears most promising for chromosomes with an average of one crossover per arm as seen in species with strong interference. Crossover interference is a common phenomenon, but varies between sexes (Drouaud *et al.* 2007), among species (Segura *et al.* 2013) and even among chromosomes within a genome (Lian *et al.* 2008). The method described here may work well on an array of taxa characterized by strong interference and short chromosomes such as birds (Tiersch & Wachtel 1991; Zhang *et al.* 2014), many mammals (Segura *et al.* 2013) and at least within the salmonid lineage among teleosts (Thorgaard *et al.* 1983; Lindner *et al.* 2000; Limborg *et al.* 2015); however, results may vary among species and testing is warranted.

Future perspectives

Centromeres are an essential component of the genomic landscape, and genetic maps lacking centromeres are incomplete. In humans, for example, subtelomeric regions are known to harbour evolutionary dynamic regions (Mefford & Trask 2002). Inference of similar patterns in other taxa requires information about the location of centromeres. Indeed, recent studies on speciation divergence report nonrandom distributions of divergence peaks in relation to centromeres and telomeres, patterns that suggest different roles of these genomic regions during adaptive divergence (Ellegren *et al.* 2012). Unfortunately, comparisons of centromeric with telomeric regions have often been impeded by missing, or equivocal, identification of centromeres on linkage maps (e.g. Gagnaire *et al.* 2013; Santure *et al.* 2013; Limborg *et al.* 2014) or on genome assemblies (e.g. Ellegren *et al.* 2012; Jones *et al.* 2012). Further, linkage maps are also crucial in guiding de novo genome assemblies (Lewin *et al.* 2009; Ellegren 2014; Tine *et al.* 2014), and linkage maps with centromeres will have an increased value in such efforts. The RF_m method presented here can help mitigate existing limitations and improve linkage maps by providing a cheap and fast approach to distinguish acrocentric from metacentric chromosomes and locate centromeres on metacentric chromosomes.

Acknowledgements

We thank Carita E. Pascal for assistance with molecular analyses. Ryan K. Waples and Wesley A. Larson are thanked for insightful discussions during the work. We further thank Darin Combs and colleagues for providing assistance and hatchery facilities to conduct the family crosses at Issaquah Hatchery (WA, USA). MTL was supported by the Danish Council for Independent Research's career programme Sapere Aude (Grant # 12-126687).

References

Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH (2011) Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics*, **188**, 799–808.

Brieuc MS, Waters CD, Seeb JE, Naish KA (2014) A dense linkage map for Chinook salmon (*Oncorhynchus tshawytscha*) reveals variable chromosomal divergence after an ancestral whole genome duplication event. *G3: Genes, Genomes, Genetics*, **4**, 447–460.

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. *The American Journal of Human Genetics*, **63**, 861–869.

Carlson BM, Onusko SW, Gross JB (2015) A high-density linkage map for *Astyanax mexicanus* using genotyping-by-sequencing technology. *G3: Genes, Genomes, Genetics*, **5**, 241–251.

Carneiro M, Ferrand N, Nachman MW (2009) Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics*, **181**, 593–606.

Chutimanitsakun Y, Nipper RW, Cuesta-Marcos A *et al.* (2011) Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics*, **12**, 4.

Colombo PC, Jones GH (1997) Chiasma interference is blind to centromeres. *Heredity*, **79**, 214–227.

Demarest BL, Horsley WH, Locke EE *et al.* (2011) Trans-centromere effects on meiotic recombination in the zebrafish. *Genetics*, **187**, 333–336.

Drouaud J, Mercier R, Chelysheva L *et al.* (2007) Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *PLoS Genetics*, **3**, 1096–1107.

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, **29**, 51–63.

Ellegren H, Smeds L, Burri R *et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, **491**, 756–760.

Everett MV, Seeb JE (2014) Detection and mapping of QTL for temperature tolerance and body size in Chinook salmon (*Oncorhynchus tshawytscha*) using genotyping by sequencing. *Evolutionary Applications*, **7**, 480–492.

Everett MV, Miller MR, Seeb JE (2012) Meiotic maps of sockeye salmon derived from massively parallel DNA sequencing. *BMC Genomics*, **13**, 521.

Fransz PF, Armstrong S, de Jong JH *et al.* (2000) Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell*, **100**, 367–376.

Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483–2497.

Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science*, **293**, 1098–1102.

Jones FC, Grabherr MG, Chan YF *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Kelly JM, McRobert L, Baker DA (2006) Evidence on the chromosomal location of centromeric DNA in Plasmodium falciparum from etoposide-mediated topoisomerase-II cleavage. *Proceedings of the National Academy of Sciences, USA*, **103**, 6706–6711.

Kodama M, Brieuc MS, Devlin RH, Hard JJ, Naish KA (2014) Comparative mapping between Coho Salmon (*Oncorhynchus kisutch*) and three other salmonids suggests a role for chromosomal rearrangements in the retention of duplicated regions following a whole genome duplication event. *G3: Genes, Genomes, Genetics*, **4**, 1717–1730.

Lagercrantz U, Lydiate DJ (1995) RFLP mapping in *Brassica nigra* indicates differing recombination rates in male and female meioses. *Genome*, **38**, 255–264.

Levan A, Fredga K, Sandberg AA (1964) Nomenclature for centromeric position on chromosomes. *Hereditas*, **52**, 201–220.

Lewin HA, Larkin DM, Pontius J, O'Brien SJ (2009) Every genome sequence needs a good map. *Genome Research*, **19**, 1925–1928.

Li Y, Cai M, Wang Z *et al.* (2008) Microsatellite–centromere mapping in large yellow croaker (*Pseudosciaena crocea*) using gynogenetic diploid families. *Marine Biotechnology*, **10**, 83–90.

Lian J, Yin Y, Oliver-Bonet M *et al.* (2008) Variation in crossover interference levels on individual chromosomes from human males. *Human Molecular Genetics*, **17**, 2583–2594.

Lien S, Gidskehaug L, Moen T *et al.* (2011) A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics*, **12**, 615.

Limborg MT, Waples RK, Seeb JE, Seeb LW (2014) Temporally isolated lineages of pink salmon reveal unique signatures of selection on distinct pools of standing genetic variation. *Journal of Heredity*, **105**, 741–751.

Limborg MT, Waples RK, Allendorf FW, Seeb JE (2015) Linkage mapping reveals strong chiasma interference in sockeye salmon: implications for interpreting genomic data. *G3: Genes, Genomes, Genetics*, **5**, 2463–2473.

Lindner KR, Seeb JE, Habicht C *et al.* (2000) Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. *Genome*, **43**, 538–549.

McKinney GJ, Seeb LW, Larson WA *et al.* (2015) An integrated linkage map reveals candidate genes underlying adaptive variation in Chinook salmon (*Oncorhynchus tshawytscha*). *Molecular Ecology Resources*. doi:10.1111/1755-0998.12479 [Epub ahead of print].

Mefford HC, Trask BJ (2002) The complex structure and dynamic evolution of human subtelomeres. *Nature Reviews Genetics*, **3**, 91–102.

Miller MR, Brunelli JP, Wheeler PA *et al.* (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, **21**, 237–249.

Morishima K, Nakayama I, Arai K (2001) Microsatellite-centromere mapping in the loach, *Misgurnus anguillicaudatus*. *Genetica*, **111**, 59–69.

Phillips R, Rab P (2001) Chromosome evolution in the Salmonidae (Pisces): an update. *Biological Reviews*, **76**, 1–25.

Reid DP, Smith C-A, Rommens M *et al.* (2007) A genetic linkage map of Atlantic halibut (*Hippoglossus hippoglossus* L.). *Genetics*, **177**, 1193–1205.

Richards PM, Liu MM, Lowe N *et al.* (2013) RAD-Seq derived markers flank the shell colour and banding loci of the *Cepaea nemoralis* supergene. *Molecular Ecology*, **22**, 3077–3089.

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. *Genes & Development*, **11**, 2600–2621.

Santure AW, De Cauwer I, Robinson MR *et al.* (2013) Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. *Molecular Ecology*, **22**, 3949–3962.

Segura J, Ferretti L, Ramos-Onsins S *et al.* (2013) Evolution of recombination in eutherian mammals: insights into mechanisms that affect recombination rates and crossover interference. *Proceedings of the Royal Society B-Biological Sciences*, **280**, 20131945.

Thorgaard GH, Allendorf FW, Knudsen KL (1983) Gene-centromere mapping in rainbow trout: high interference over long map distances. *Genetics*, **103**, 771–783.

Tiersch TR, Wachtel SS (1991) On the evolution of genome size of birds. *Journal of Heredity*, **82**, 363–368.

Tine M, Kuhl H, Gagnaire PA *et al.* (2014) European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications*, **5**, 5770.

Wang N, Fang LC, Xin HP, Wang LJ, Li SH (2012) Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. *BMC Plant Biology*, **12**, 148.

Xu P, Zhang X, Wang X *et al.* (2014) Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nature Genetics*, **46**, 1212–1219.

Zhang G, Li C, Li Q *et al.* (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, **346**, 1311–1320.

M.T.L., L.W.S. and J.E.S. conceived the study. M.T.L. developed the method with input from G.M.K. M.T.L. drafted the manuscript, and all authors commented on earlier versions.

Data accessibility

The genotype data for the gynogenetic haploid and diploid sockeye salmon families are available from Limborg *et al*. (2015) at Dryad with doi: http://dx.doi.org/10.5061/dryad.q675s. All other data sets are available through the original publications: pink salmon (http://dx.doi.org/10.5061/dryad.pp43m; Limborg *et al.* 2014), rainbow trout (appendix S2 in Miller *et al.* 2012), barley (http://wheat.pw.usda.gov/ggpages/maps/OWB/; Chutimanitsakun *et al.* 2011), Chinook salmon (http://dx.doi.org/10.5061/dryad.j7245; McKinney *et al.* 2015).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Plots of RF_m counted from the left (cM=0; blue circles) and from the right (red circles) with known centromeric regions (grey bars) for all LGs in the haploid sockeye salmon map presented in Limborg *et al*. (2015). The known centromeric regions were used to characterize whether a LG represents a acrocentric (A) or a metacentric (M) chromosome. For haploid families, RF_m values are expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm with strong interference. Note that LG So18 was split into two linkage groups in Limborg *et al*. (2015); this led to an amputated construction of one arm (So18B) 2014).

Fig. S2 Results for the haploid female linkage map for *O. gorbuscha* (Limborg *et al*. 2014) characterized by high-density marker spacing and small chromosome arms. Plots illustrate the distribution of RF_m values counted from the left (cM=0; blue circles) and from the right (red circles). For haploid families, RF_m values are expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm with strong interference.

Fig. S3 Results for a doubled haploid male linkage map for *O. mykiss* (Miller *et al*. 2012) characterized by low-density marker spacing and small chromosome arms. Plots illustrate the distribution of RF_m values counted from the left (cM=0; blue circles) and from the right (red circles). For doubled haploid families, RF_m values are expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm with strong interference.

Fig. S4 Results for the doubled haploid barley, *Hordeum vulgare*, map (Chutimanitsakun *et al*. 2011) characterized by high-density marker spacing and long chromosome arms. Plots illustrate the distribution of RF_m values counted from the left (cM=0; blue circles) and from the right (red circles). For doubled haploid families, RF_m values are expected to plateau around 0.5 (dotted horizontal line) when covering a full chromosome arm with strong interference.

Table S1 Estimated RFm values for all loci on each linkage map for the different species analyzed in this study.