Title

Identifying homomorphic sex chromosomes from wild-caught adults with limited genomic resources

Journal

Molecular Ecology Resources, 17(4)

ISSN

1755-098X

Authors

Brelsford, Alan
Lavanchy, Guillaume
Sermier, Roberto
et al.

Publication Date

2017-07-01

DOI

10.1111/1755-0998.12624

Peer reviewed

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2017) **17**, 752–759 doi: 10.1111/1755-0998.12624

Identifying homomorphic sex chromosomes from wild-caught adults with limited genomic resources

ALAN BRELSFORD,*[1a] GUILLAUME LAVANCHY,*[a] ROBERTO SERMIER,* ANNA RAUSCH† and NICOLAS PERRIN*

*Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland†Department of Integrative Zoology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

Abstract

We demonstrate a genotyping-by-sequencing approach to identify homomorphic sex chromosomes and their homolog in a distantly related reference genome, based on noninvasive sampling of wild-caught individuals, in the moor frog *Rana arvalis*. Double-digest RADseq libraries were generated using buccal swabs from 30 males and 21 females from the same population. Search for sex-limited markers from the unfiltered data set (411 446 RAD tags) was more successful than searches from a filtered data set (33 073 RAD tags) for markers showing sex differences in heterozygosity or in allele frequencies. Altogether, we obtained 292 putatively sex-linked RAD loci, 98% of which point to male heterogamety. We could map 15 of them to the *Xenopus tropicalis* genome, all but one on chromosome pair 1, which seems regularly co-opted for sex determination among amphibians. The most efficient mapping strategy was a three-step hierarchical approach, where *R. arvalis* reads were first mapped to a low-coverage genome of *Rana temporaria* (17 My divergence), then the *R. temporaria* scaffolds to the *Nanorana parkeri* genome (90 My divergence), and finally the *N. parkeri* scaffolds to the *X. tropicalis* genome (210 My). We validated our conclusions with PCR primers amplifying part of *Dmrt1*, a candidate sex determination gene mapping to chromosome 1: a sex-diagnostic allele was present in all 30 males but in none of the 21 females. Our approach is likely to be productive in many situations where biological samples and/or genomic resources are limited.

Keywords: genotyping by sequencing, RAD tags, *Rana arvalis*, Ranidae, sex chromosome turnover, sex determination

Received 9 May 2016; revision received 11 August 2016; accepted 16 August 2016

Introduction

In sharp contrast with mammals and birds, many cold-blooded vertebrates present homomorphic (i.e. morphologically undifferentiated) sex chromosomes. Along with occasional XY recombination, one main cause of homomorphy resides in high rates of sex chromosome turnover, during which new sex-determining mutations replace established sex chromosomes before they had time to degenerate (Schartl 2004; Volff *et al.* 2007). Documenting and quantifying turnover rates constitute an important step towards understanding the evolutionary forces acting on sex determination mechanisms (Beukeboom & Perrin 2014; Bachtrog *et al.* 2014; Dufresnes *et al.* 2015). However, the absence of cytogenetic differentiation hinders the identification of sex chromosomes and detection of homologies, especially when chromosomal fusions and translocations hamper comparisons between species. In this context, molecular genetic approaches have the potential to provide a powerful alternative to cytogenetics.

Sex-linked markers directly distinguish male- from female-heterogametic systems (e.g. Berset-Brändli *et al.* 2006; Gamble & Zarkower 2014), which might then provide first insights onto the patterns of turnover. By analysing patterns of heterogamety in 63 species of frogs and newts, Hillis & Green (1990) uncovered a high frequency of transitions throughout the phylogenetic history of amphibians, with at least seven switches to male heterogamety from an ancestral female-heterogametic system. A similar approach in 37 species of geckos unveiled a high rate of independent transitions from an ancestral TSD system to either ZW or XY systems (Gamble *et al.* 2015). Changes in heterogamety, however, might substantially underestimate sex chromosome turnovers, insofar as sex linkage may also display considerable variation within groups homogeneous for

Correspondence: Nicolas Perrin, Fax: 0041 21 692 41 65;
E-mail: nicolas.perrin@unil.ch
[1]Present address Biology Department, University of California, Riverside, CA 92521, USA
[a]These authors contributed equally to this work.

heterogamety. Mapping sex-linked markers to specific genomic regions is thus required as an additional step to identify sex chromosomes and establish homologies. Based on the patterns of sex linkage of enzymatic polymorphisms, Sumida & Nishioka (2000) revealed a high rate of sex chromosome turnover within Ranidae, a family otherwise highly homogeneous for male heterogamety. Interestingly, five chromosome pairs, out of the 13 pairs that form the normal complement in this family, seem to be recurrently co-opted for sex (Miura 2007), suggesting that some genomic regions are more likely than others to take a sex-determining role. More recently, Brelsford *et al.* (2013) developed a series of gene-based markers to show that a genomic region corresponding to the largest chromosome pair in *Xenopus tropicalis* (*Xt1*) had been independently co-opted for sex determination in three highly diverged lineages of frogs: *Rana temporaria* (Ranidae), *Hyla arborea* (Hylidae) and *Bufo siculus* (Bufonidae). This result underlines the potential of gene-based approaches to identify sex chromosomes, once a candidate sex-determining genomic region has been recognized.

The ongoing rise of next-generation sequencing is now opening a new chapter in sex determination research. Gamble & Zarkower (2014) recently applied a genotyping-by-sequencing (GBS) approach to identify a male-specific marker in *Anolis carolinensis*, based on a few sexed adults. A similar approach was used by Gamble *et al.* (2015) to identify the patterns of heterogamety in 12 species of geckos. Brelsford *et al.* (2016a) developed an alternative GBS approach to identify sex chromosomes in *H. arborea*; high-density linkage maps obtained from one family revealed a threefold excess of single nucleotide polymorphisms (SNPs) for linkage group 1 (the *Hyla* homolog of *Xt1*) in the male (the heterogametic sex). Homologies between *H. arborea* linkage groups and *X. tropicalis* chromosomes could be established by aligning GBS reads to a *H. arborea* low-coverage draft genome, then searching the *X. tropicalis* genome for each *H. arborea* scaffold that contained a marker on the linkage map. These results validated the potential of GBS approaches to identity sex chromosomes based on a single family, without information on offspring sex. Using a similar GBS approach, Brelsford *et al.* (2016b) further showed that adding information on offspring phenotypic sex increased power sufficiently to confidently rule out any major genetic component to sex determination in a *R. temporaria* family. Homologies with *X. tropicalis* were also established by aligning GBS reads to an *R. temporaria* draft genome, then searching the *X. tropicalis* genome for *R. temporaria* scaffolds containing a marker on the linkage map.

The present study aimed at developing a GBS approach to identify sex chromosomes based on noninvasive sampling of wild-caught adults and with limited genomic resources. In particular, we aimed at delineating the most powerful strategies to (i) identify sex-linked GBS reads, and (ii) map them on a distantly related reference genome. We focus on *Rana arvalis*, a member of the species-rich frog family Ranidae, first because this family seems characterized by a very high rate of sex chromosome turnover (so that the methodology can be readily extended to investigate patterns of turnover across the family), second because no prior information was available on this species' sex determination system, and neither were genetic map or genomic resources. Our approach provides operational rules that can be readily applied to a wide range of species with undifferentiated sex chromosomes, even in cases when biological samples and genetic resources are limited.

Materials and methods

Study species and sampling

The moor frog (*R. arvalis*), which belongs to the group of Palearctic brown frogs, diverged from *R. temporaria* approximately 17 My ago (Voituron *et al.* 2009; Gomez-Mestre *et al.* 2012). Unlike most Ranidae, which have 13 pairs of chromosomes (five large and eight small), moor frogs only possess 12 pairs (six large and six small), with the second largest pair resulting from the fusion of two small pairs (Green & Borkin 1983). Moor frogs breed in standing or slow-flowing water bodies such as swamps, moats and ponds (http://amphibiaweb.org/species/4983). While females are brownish all year long, males display a typical light blue coloration during the breeding season (Fig. 1), which allows unambiguous phenotypic sexing. In Spring 2011, 21 females and 30 males were collected at a single pond in southeastern Austria (47°10 N, 16°5 E) and sampled for buccal cell DNA with cotton swabs before release. See Rausch *et al.* (2014) for more details on sampling and DNA extractions.

Wet laboratory

We generated double-digest RADseq libraries following the protocol of Brelsford *et al.* (2016a), adapted from Parchman *et al.* (2012) with some features from Peterson *et al.* (2012). Briefly, we digested genomic DNA with SbfI-HF and MseI restriction enzymes (New England Biolabs) and ligated an individual barcode (a specific sequence of 4–8 bases) to the SbfI cut sites. We then amplified the fragments obtained in four replicate individual-specific PCRs of 20 cycles. PCR products were size-selected to retain fragments of approximately 350–500 bp. The resulting library was single-end sequenced (100 bp) at the Lausanne Genomic Technologies Facility

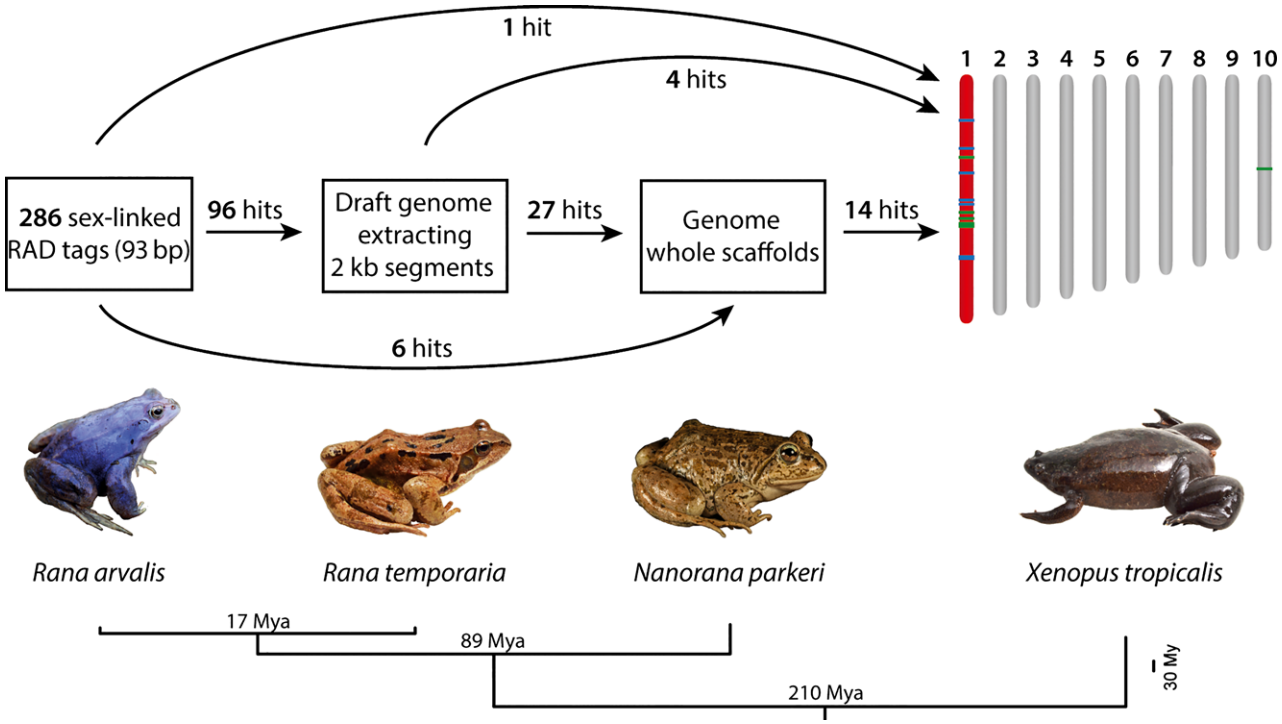

Fig. 1 Summary of the different approaches for mapping *Rana arvalis* sex-linked markers onto the genome of *Xenopus tropicalis*. Each arrow designates a blastn search, with the number of hits given. The sex chromosome is indicated in red on the karyotype of *X. tropicalis*, with the localization of blast hits as horizontal lines. Blue: X-linked markers. Green: Y-specific markers. Numbers at each node in the phylogeny of the different species represent estimated divergence times, taken from Timetree.org. Picture credits: Jan Jezek (*R. arvalis*), Yufan Wang © AmphibiaChina (*Nanorana parkeri*), GL (*Rana temporaria*, *X. tropicalis*).

on one Illumina HiSeq 2500 lane. We obtained 171 357 393 raw Illumina reads.

Filtering and SNP calling

Raw Illumina reads were quality-checked with FASTQC v0.10.1 (Andrews 2010) and demultiplexed using the *process_radtags* module of STACKS (v1.30; Catchen *et al.* 2013) with the *filter_illumina* option. We then used the STACKS *denovo_map.pl* pipeline to assemble our reads into RAD tags (which refers to the sequenced portion of DNA that flanks a SbfI restriction site) and call SNPs. We used STACKS default parameters. Using VCFTOOLS v0.1.12b (Danecek *et al.* 2011), we filtered our loci for a minimum depth of 7, a minimum minor allele frequency of 0.05 and presence in at least 75% of individuals. These parameters were a good compromise between data quality and quantity (Fig. S1, Supporting Information).

Finding sex-linked markers

We tested three complementary approaches to identify sex-linked markers, respectively, based on (i) sex differences in allele frequencies, (ii) sex differences in heterozygosity and (iii) sex-limited occurrence. The rationale behind each approach is outlined below, and custom R scripts were used to parse STACKS output files accordingly.

1 Because X-linked alleles in male-heterogametic systems should occur in two copies in females but only one in males, we expect a 0.5 frequency difference between females and males. The same rationale applies to female-heterogametic systems, where Z-linked allele should occur in two copies in males but only one in females, leading to a minus 0.5 frequency difference (i.e. with opposite sign). We considered a SNP as sex-linked if one allele had a frequency ≥0.95 in one sex (the homogametic one) and a frequency difference >0.4 between sexes (thereby allowing for a few sequencing errors, recombination events or sex reversals). Sex-specific allelic frequencies were computed in VCFTOOLS v0.1.12b (Danecek *et al.* 2011).
2 Excess heterozygosity is expected in the heterogametic sex (XY males or ZW females) if the X and Y chromosomes (or Z and W) are fixed for different alleles. We therefore searched for loci where at least two-thirds of the individuals from one sex were heterozygous, while those from the other sex were homozygous (allowing

for one single heterozygote). Although this method is expected to return some of the same loci as the first approach, apparent homozygosity might also result from hemizygosity in the heterogametic sex, as should occur if the Y (or W) alleles are absent or sufficiently diverged from their gametolog to be assigned to another locus. Hence, this second approach might also identify X-limited or Z-limited markers. The two alternatives can be tested a posteriori by checking the depth of coverage, which should be equal between sexes in the first case, but halved in the hemizygous sex in the second case.

3 Strong divergences between gametologs, or mutations in the restriction site, might also result in Y- or W-limited markers. We parsed the unfiltered data set in R to look for loci that were present in at least half of individuals from one sex, but absent in the other.

The number of individuals necessary to confidently identify the sex chromosome was estimated through resampling, by analysing for heterozygosity differences (approach 2) all combinations of 4, 8, 12, 16, 20 and all individuals of either sex, randomly sampled from the total data set (10 replicates for each combination).

Mapping sex-linked markers

Sex chromosomes were identified by mapping the candidate sex-linked markers onto the *X. tropicalis* genome, which is the best assembled amphibian genome available so far (Hellsten *et al.* 2010). Comparative studies have consistently shown extensive synteny across anuran families, except for occasional differences in chromosome numbers resulting from fusions or fissions (Hotz *et al.* 1997; Brelsford *et al.* 2013, 2016a,b). As the direct mapping of short reads (93 bp) was bound to meet little success (given the 210 My divergence between *Rana* and *Xenopus*; Gomez-Mestre *et al.* 2012), we devised a stepwise strategy (Fig. 1) by first mapping our sex-linked RAD tags to a draft *R. temporaria* genome (Brelsford *et al.* 2016b), with ~17 My divergence from *R. arvalis*, using BLASTN (v.2.2.26). Blast hits were only retained if their *e*-value was at least five orders of magnitude lower than the second best hit, using a custom python script (Purcell *et al.* 2014). We then extracted 1 kb of the *R. temporaria* scaffold each side of the hit and blasted these 2 kb sequences to the *X. tropicalis* genome. Preliminary tests had revealed that longer sequences from the *R. temporaria* scaffolds (10 kb) provided more hits but also more false positives, due to poor assembly and repeated sequences. We also tested a three-step procedure by blasting first the *R. temporaria* sequences onto the genome of *Nanorana parkeri* (Sun *et al.* 2015), with a ~89 My divergence time (Gomez-Mestre *et al.* 2012), as an additional intermediate step, before blasting *N. parkeri* scaffolds to the *X. tropicalis* genome (Fig. 1). We also tried to blast the *R. arvalis* RAD tags directly to the *N. parkeri* and *X. tropicalis* genomes (Fig. 1).

PCR Validation

Our GBS approach pointed to a male-heterogametic system, with a sex locus mapping to chromosome 1 (see Results). To validate these conclusions, we analysed the same sample with PCR primers developed by Ma *et al.* (2016) to investigate *Dmrt1* polymorphism in *R. temporaria*, a candidate sex determination gene mapping to chromosome 1 throughout amphibians (Brelsford *et al.* 2013). The same primer pairs and PCR conditions were used as described in their Appendix S2, Table S2 (Supporting Information) (http://onlinelibrary.wiley.com/store/10.1002/ece3.2209/asset/supinfo/ece32209-sup-0002-AppendixS2.doc?v=1&s=12ad4a427fc2207f5138cdb3171ff2e72f7ddb18); cross-amplifying PCR products were analysed for sex differences in allelic frequencies.

Results

After filtering, we obtained a total of 63 784 SNPs located on 33 073 different RAD tags, as compared to 411 446 RAD tags for the unfiltered data set. The approach based on frequency differences (i) identified 43 X-linked alleles, located on 37 RAD tags. By contrast, no single marker matched a Z-linked pattern, a highly significant difference ($\chi^2 = 37$, d.f. = 1, $P = 1.2e^{-09}$). The approach based on heterozygosity differences (ii) identified 127 sex-linked SNPs with an XY pattern, located on 115 RAD tags, of which 22 also belonged to the set obtained using approach (i). Two SNPs located on the same RAD tag displayed a pattern compatible with a ZW system, being heterozygous in the majority of females, but homozygous in all males. This, however, was due to male hemizygosity: although average coverage did not differ between sexes genomewide (being 12.19 in males and 12.91 in females), female coverage was significantly higher at this locus (36.4 vs. 21.1 in males; Welch two-sample *t*-test: $t = 2.76$, d.f. = 29.8, $P = 0.01$), but not different from twice the male value ($t = 0.82$, d.f. = 47.2, $P = 0.41$), pointing to an X-limited marker. Finally, the approach based on sex-limited occurrence (iii) identified 155 male-limited RAD tags (nine of which polymorphic), and six female-limited RAD tags, out of the 411 446 unfiltered RAD tags. Together, these three methods identified a total of 292 putatively sex-linked RAD tags, 98% of which indicate an XY system. From our power analysis, 12 individuals of each sex seem enough to confidently identify the system (Fig. S2, Supporting Information), providing an average of ~80 SNPs in

favour of an XY system (as opposed to 127 for the total data set) vs. ~4 SNPs in favour of a ZW system (as opposed to 2 with the total data set). For lower sample sizes, false positives increase in numbers and blur the picture.

Direct blasting of the 286 XY-indicative reads to the *X. tropicalis* genome provided one single hit (Fig. 1). A two-step procedure using *R. temporaria* as an intermediate produced 96 hits on the *R. temporaria* genome, of which four (including the one obtained by direct blast) could be mapped onto the *X. tropicalis* genome. The alternative two-step procedure using *N. parkeri* only as an intermediate (instead of *R. temporaria*) produced six hits on the *N. parkeri* genome, of which one could be mapped on the *X. tropicalis* genome. The three-step procedure was, by far, the most successful: out of the 96 hits on the *R. temporaria* genome, 27 could be mapped on an *N. parkeri* scaffold, of which 14 met a correspondence on the *X. tropicalis* genome. Accounting for redundancies (same hits identified through different procedures), we obtained 15 unique hits out of the 286 putatively sex-linked RAD tags. Out of these 15 hits, 14 were located on *X. tropicalis* scaffold 1, which corresponds to the largest pair of chromosomes in Ranidae (Miura 2007). They were not clustered at a single position on this chromosome, but instead scattered all along its length, except for the tips (Fig. 1). The only exception mapped to scaffold 10.

All results thus converged to the conclusions that *R. arvalis* has a genetic sex determination system with male heterogamety and that the sex locus maps to chromosome pair 1. To validate these conclusions, we analysed this sample with primer pairs developed by Ma *et al.* (2016) to investigate *Dmrt1* polymorphism in *R. temporaria*. Three of the five primer pairs (*Dmrt1-1*, *Dmrt1-2* and *Dmrt1-5*) cross-amplified PCR products in *R. arvalis*, but only *Dmrt1-1* showed informative length polymorphism. This revealed complete sex linkage (Fig. S3, Supporting Information), with allele *342* found in all of the 30 males but in none of the 21 females ($\chi^2 = 31.03$, $P = 2.6 \times 10^{-8}$).

Discussion

Our results provide strong evidence that the population of *R. arvalis* under study has a strictly genetic sex determination system with male heterogamety (supported by 286 informative RAD tags out of 292), with the sex locus lying on the largest pair of chromosomes (supported by 14 hits out of 15). Both conclusions are validated by our independent PCR-based analysis. We discuss below the methodological and biological insights gained by our approach.

Regarding heterogamety, all three strategies implemented to find sex-linked markers reached the same conclusion. The approach based on frequency distribution (i) was less powerful than the two others: 43 SNPs located on 37 RAD tags presented a twofold higher frequency in females than in males (as expected from X linkage), while none showed a reverse pattern. The approach based on heterozygosity differences (ii), combined with coverage analysis, conferred a higher power: 115 RAD tags (22 of which shared with the first approach) showed a male-biased heterozygosity, pointing to an XY system. A single RAD tag showed a female-biased heterozygosity suggestive of a ZW system, but coverage analysis revealed this locus to be actually X-limited and thus hemizygous in males. The third approach, which searched the unfiltered data set for sex-limited occurrence (iii), was the most powerful in terms of absolute numbers of putatively sex-linked markers, but also provided more false positives: 155 RAD tags were male-limited, and six female-limited.

Altogether, 98% of putative sex-linked markers pointed to male heterogamety; the remaining 2% (six female-specific markers) are to be interpreted as false positives. Our results thus add the moor frog to the list of Ranidae that display male heterogamety, thereby increasing support for the idea that XY systems are much more common than ZW among true frogs, being maintained throughout the radiation despite high rate of sex chromosome turnovers. Whatever the ultimate reason for this high level of conservation, it confirms that estimates of sex chromosome turnovers only based on changes in the patterns of heterogamety are likely to overlook high numbers of transitions.

Gamble & Zarkower (2014) implemented an approach akin to our third strategy to search for Y- or W-limited markers among seven males and ten females of *A. carolinensis*. This strategy met limited success in their case, resulting in a single male-specific marker, presumably because the sex locus in *Anolis* maps to a microchromosome (while chromosome 1 in frogs is very large). A similar approach was used by Gamble *et al.* (2015) to identify heterogamety in 12 species of geckos; more markers in general were found to meet the criteria, with, however, more false positives (male-specific markers found in ZW systems and vice versa), mostly due to low sample sizes. In both studies, sex-specific markers were validated using PCRs on additional individuals, motivated by the low numbers of loci identified and/or frequent false positives. Male heterogamety was similarly confirmed by PCR in our case, even though this validation step was actually not really required, given the high number of markers obtained overall, the very low proportion of false positives and the convergence of all approaches. We recommend that several strategies be applied in parallel to detect sex-linked markers, not only to allow for cross-validation, but also because the relative

power of each strategy is likely to be species specific; the number of X- or Y-limited markers, for example, is expected to increase with the amount of sex chromosome differentiation.

Our mapping strategies met limited success overall, with only 15 markers successfully mapped out of 286 (i.e. 5%). However, all but one were on the same scaffold, providing high confidence regarding the identity of the sex chromosome (furthermore validated by PCR). The only exception (a male-specific RAD tag mapping to chromosome 10) might have occurred for several reasons. First, this locus could be falsely male specific, a likely possibility given our findings of falsely female-specific RAD tags. Second, it could be a mapping error, caused by alignment to the wrong copy of a repetitive sequence in the genome of *X. tropicalis* or one of the intermediate species. Third, Y-linked sequences are expected to evolve faster than X-linked ones because of their smaller effective population sizes, and might thus be harder to map (although in our case, Y- and X-linked RAD tags were mapped in similar proportions, 5.1% and 5.3%, respectively). Finally, a region of *X. tropicalis* chromosome 10 might have been translocated onto *R. arvalis* chromosome 1 (although this might not seem parsimonious given the highly conserved synteny in amphibians).

The most powerful mapping strategy appeared to be the one involving several intermediate steps. Adding first the genomes of *R. temporaria* (with 17 My divergence from *R. arvalis*), and then of *N. parkeri* (89 My divergence) as intermediate steps between *R. arvalis* reads and the *X. tropicalis* genome (210 My divergence), greatly improved our mapping yield, allowing identification of 14 out of our 15 hits. Hence, if no good genome from a closely related species is available, we strongly recommend passing through intermediate, more closely related genomes, before reaching the well-assembled target genome. As our *R. temporaria* genome is only an early draft, this approach is worth trying even if the intermediate genomes are not of great quality.

Interestingly, the 14 markers mapping to *X. tropicalis* chromosome 1 were not clustered, but scattered all along the chromosome except for distal segments; combined with the absence of cytogenetic differentiation, this result falls in line with the patterns of sex chromosome evolution in frogs. As chromosomes only recombine in their distal segments in males (Brelsford *et al.* 2016a,b), sex chromosomes start differentiating over most of their length as soon as the Y becomes male-limited (i.e. no evolutionary strata are expected). However, XY similarity is regularly reset by occasional XY recombination or new turnovers, so that chromosomes have little time to degenerate, and many markers still have a gametolog. Accordingly, Brelsford *et al.* (2016a) identified the sex chromosome in a *H. arborea* family through the large excess of heterozygosity in the male, spread all over chromosome 1 except for distal segments. By the same token, it is expected that ZW systems will be more difficult to identify with our approach: given that female frogs recombine all along their chromosomes, ZW differentiation should be limited to the immediate surroundings of the sex locus. On the positive side, successful mapping of a sex-linked marker in a ZW system should directly provide relevant information on the precise localization of the sex determination locus.

The fact that *X. tropicalis* has 10 pairs of chromosomes, while most Ranidae have 13 pairs, and *R. arvalis* 12 pairs, implies that a few fissions and fusions have occurred despite the overall strong synteny. However, none has apparently affected chromosome pair 1. Indeed, if the sex locus in *R. arvalis* had been on a chromosome resulting from a fusion (relative to the *X. tropicalis* karyotype), then sex-linked markers would have been spread over two different *X. tropicalis* chromosomes. If, conversely, this sex locus had been on a chromosome resulting from a fission, then sex-linked markers would have mapped to one single arm of a *X. tropicalis* chromosome, which is clearly not the case either (Fig. 1).

The sex chromosome pair identified here (*Xt1*) is the same as in several species of Hylidae and Bufonidae (Brelsford *et al.* 2013), and belongs to the set of five pairs recurrently co-opted by different species of Ranidae. Its role in sex determination has been established in particular in *Lithobates clamitans*, *Lithobates sphenocephalus*, *Lithobates berlandieri*, *Rana japonica* (Sumida & Nishioka 2000; Miura 2007) and *R. temporaria* (Brelsford *et al.* 2013). Given the high rate of turnover that characterizes this family, it does not necessarily represent an ancestral pair, particularly because these six species are widely scattered across the Ranidae phylogeny (Pyron & Wiens 2011). In addition, several of them show some polymorphism: in *R. japonica*, different chromosomes (corresponding to *Xt1* and *Xt3*) determine sex depending on populations (Sumida & Nishioka 1994). In *R. temporaria*, sex determination is not genetic in some populations (Brelsford *et al.* 2016b), while either *Xt1* or both *Xt1* and *Xt2* segregate with sex in other populations (Rodrigues *et al.* 2016). The number of independent co-options of *Xt1* might be estimated via proper phylogenetic analyses. The identification and analysis of candidate sex-determining genes might also shed light on this issue: alleles at genes from a shared ancestral sex determination region should cluster by gametologs, while a reverse pattern (clustering by species) is expected for all sex-linked genes if sharing of sex chromosomes results from independent co-option (Brelsford *et al.* 2016c).

Conclusion and perspectives

Our approach successfully established the patterns of heterogamety and identity of sex chromosomes in one species of Ranidae, based on noninvasive samples and despite limited availability of genomic resources. This approach can be readily extended to other lineages from this species-rich radiation, where sex determination systems show particularly high diversity and strong dynamics. Combined with phylogenetic analyses, these developments have the potential to provide new insights onto the rate and mechanisms of sex chromosome turnover in Ranidae. More generally, our approach provides operational rules that can be readily applied to a wide range of species with undifferentiated sex chromosomes, even in cases when biological samples and genetic resources are limited.

On a more specific basis, our identification of sex chromosomes in *R. arvalis* opens the way to the development of a variety of sex-linked markers, which will prove useful in population genetics analyses requiring information on individual sex, or aimed at drawing inferences on sex-biased patterns (e.g. sex-biased dispersal). Sex-diagnostic markers are also crucial for investigations on the potential sex-reversing effects of pesticides through endocrine disruption, for which amphibians are becoming a model system (e.g. Kloas 2002; Hayes *et al*. 2010; Lambert *et al*. 2015). Given the strong synteny that characterizes amphibians, many sex-linked gene-based markers can be developed based on *X. tropicalis* scaffold 1, along the line followed by Brelsford *et al*. (2013). Our PCR validation based on a *Dmrt1* primer pair designed for *R. temporaria* also shows that markers developed for related species might already reveal useful in this context.

Of particular interest might be the analysis of sequence data from candidate sex determination genes such as *Dmrt1* or *Amh*. Indications for a direct involvement of *Dmrt1* in sex determination have been gathered from several species of *Hyla* (Brelsford *et al*. 2016c) as well as from *R. temporaria* (Ma *et al*. 2016). Further characterization of this gene in the several species of Ranidae that use this same sex chromosome will allow testing whether this situation results from several independent co-options, or inheritance of an ancestral sex determination system.

Acknowledgements

Thanks to Eva Ringler and Marc Sztatecsny for help in obtaining the samples, Daniel L. Jeffries for providing a script for data analysis and for insightful comments on the manuscript, and Jan Jezek and Yufan Wang (Amphibiachina.org) for frog pictures. The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. The Swiss National Science Foundation provided financial support (grants 31003B_147091 and CRSII3_147625 to NP).

References

Andrews S (2010) FASTQC: *A Quality Control Tool for Highthroughput Sequence Data*. Babraham Bioinformatics, Cambridge. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bachtrog D, Mank JE, Peichel CL *et al*. (2014) Sex determination: why so many ways of doing it? *PLoS Biology*, **12**, e1001899.

Berset-Brändli L, Jaquiéry J, Dubey S, Perrin N (2006) A sex-specific marker reveals male heterogamety in European tree frogs. *Molecular Biology and Evolution*, **23**, 1104–1106.

Beukeboom L, Perrin N (2014) *The Evolution of Sex Determination*, 240 pp. Oxford Univ. Press, Oxford.

Brelsford A, Stöck M, Betto-Colliard C *et al*. (2013) Homologous sex chromosomes in three deeply divergent anuran species. *Evolution*, **67**, 2434–2440.

Brelsford A, Dufresnes C, Perrin N (2016a) High-density sex-specific linkage maps of a European tree frog (*Hyla arborea*) identify the sex chromosome without information on offspring sex. *Heredity*, **116**, 177–181.

Brelsford A, Rodrigues N, Perrin N (2016b) High-density linkage maps fail to detect any genetic component to sex determination in a *Rana temporaria* family. *Journal of Evolutionary Biology*, **29**, 220–225.

Brelsford A, Dufresnes C, Perrin N (2016c) Trans-species variation in *Dmrt1* is associated with sex determination in four European tree-frog species. *Evolution*, **70**, 840–847.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) STACKS: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Danecek P, Auton A, Abecasis G *et al*. (2011) The variant call format and VCFTOOLS. *Bioinformatics*, **27**, 2156–2158.

Dufresnes C, Borzée A, Horn A *et al*. (2015) Sex chromosome homomorphy in Palearctic tree frogs results from both turnovers and X–Y recombination. *Molecular Biology and Evolution*, **32**, 2328–2337.

Gamble T, Zarkower D (2014) Identification of sex-specific molecular markers using restriction site associated DNA sequencing. *Molecular Ecology Resources*, **14**, 902–913.

Gamble T, Coryell J, Ezaz T, Lynch J, Scantlebury D, Zarkower D (2015) Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Molecular Biology and Evolution*, **32**, 1296–1309.

Gomez-Mestre I, Pyron RA, Wiens JJ (2012) Phylogenetic analyses reveal unexpected patterns in the evolution of reproductive modes in frogs. *Evolution*, **66**, 3687–3700.

Green DM, Borkin LJ (1983) Evolutionary relationships of Eastern Palearctic Brown Frogs, genus *Rana*: paraphyly of the 24-chromosome species group and the significance of chromosome number change. *Zoological Journal of the Linnean Society*, **109**, 1–25.

Hayes TB, Khoury V, Narayan A *et al*. (2010) Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proceedings of the National Academy of Sciences*, **107**, 4612–4617.

Hellsten U, Harland RM, Gilchrist MJ *et al*. (2010) The genome of the western clawed frog *Xenopus tropicalis*. *Science*, **328**, 633–636.

Hillis DM, Green DM (1990) Evolutionary changes of heterogametic sex in the phylogenetic history of amphibians. *Journal of Evolutionary Biology*, **3**, 49–64.

Hotz H, Uzzell T, Berger L (1997) Linkage groups of protein-coding genes in Western Palearctic water frogs reveal extensive evolutionary conservation. *Genetics*, **147**, 255–270.

Kloas W (2002) Amphibians as a model for the study of endocrine disruptors. *International Review of Cytology*, **216**, 1–57.

Lambert MR, Giller GSJ, Barber LB, Fitzgerald KC, Skelly DK (2015) Suburbanization, estrogen contamination, and sex ratio in wild amphibian populations. *Proceedings of the National Academy of Sciences*, **112**, 11881–11886.

Ma WJ, Rodrigues N, Sermier R, Brelsford A, Perrin N (2016) *Dmrt1* polymorphism covaries with sex-determination patterns in *Rana temporaria*. *Ecology and Evolution*, **6**, 5107–5117.

Miura I (2007) An evolutionary witness: the frog *Rana rugosa* underwent change of heterogametic sex from XY male to ZW female. *Sexual Development*, **1**, 323–331.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle C (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, **21**, 2991–3005.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, **7**, e37135.

Purcell J, Brelsford A, Wurm Y, Perrin N, Chapuisat M (2014) Convergent genetic architecture underlies social organization in ants. *Current Biology*, **24**, 2728–2732.

Pyron RA, Wiens JJ (2011) A large-scale phylogeny of amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution*, **61**, 543–583.

Rausch AM, Sztatecsny M, Jehle R, Ringler E, Hödl W (2014) Male body size and parental relatedness but not nuptial colouration influence paternity success during scramble competition in *Rana arvalis*. *Behaviour*, **151**, 1869–1884.

Rodrigues R, Vuille Y, Brelsford A, Merilä J, Perrin N (2016) The genetic contribution to sex determination and number of sex chromosomes vary among populations of common frogs (*Rana temporaria*). *Heredity*, **117**, 25–32.

Schartl M (2004) Sex chromosome evolution in non-mammalian vertebrates. *Current Opinion in Genetics & Development*, **14**, 634–641.

Sumida M, Nishioka M (1994) Geographic variability of sex-linked loci in the Japanese Brown Frog, *Rana japonica*. *Scientific Report of the Laboratory for Amphibian Biology*, **13**, 173–195.

Sumida M, Nishioka M (2000) Sex-linked genes and linkage maps in amphibians. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **126**, 257–270.

Sun YB, Xiong ZJ, Xiang XY *et al.* (2015) Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of Sciences*, **112**, E1257–E1262.

Voituron Y, Barré H, Ramløv H, Douady CJ (2009) Freeze tolerance evolution among anurans: frequency and timing of appearance. *Cryobiology*, **58**, 241–247.

Volff JN, Nanda I, Schmid M, Schartl M (2007) Governing sex determination in fish: regulatory putsches and ephemeral dictators. *Sexual Development*, **1**, 85–99.

N.P. and A.B. designed the study. A.R. collected the samples. R.S. and A.B. involved in laboratory work. G.L. and A.B. analysed the data. N.P., G.L. and A.B. wrote the study with input from all authors.

Data accessibility

Raw Illumina GBS reads have been deposited at NCBI Sequence Read Archive, BioProject PRJNA339473. STACKS outputs have been archived on Dryad, doi: 10.5061/dryad.n5r52. FASTA sequence of sex-linked RAD tags is available as Supporting Information.

Supporting Information

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

Table S1 Mapping of sex-linked reads onto *Xenopus tropicalis* genome.

Table S2 Fasta file of all sex-linked RADtags. IDs as in Table S1 (Supporting Information).

Figure S1 Relationship between filtering parameters and the number of retained loci.

Figure S2 Relationship between samples size and the number of retained loci.

Figure S3 Sex-specific allelic frequencies at *Dmrt1-1*.