White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/104513/

Version: Accepted Version

Article:

Wielstra, B., Burke, T., Butlin, R.K. et al. (2016) Efficient screening for ‘genetic pollution’ in an anthropogenic crested newt hybrid zone. Conservation Genetics Resources, 8 (4). pp. 553-560. ISSN: 1877-7252

https://doi.org/10.1007/s12686-016-0582-3

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Efficient screening for 'genetic pollution' in an anthropogenic crested newt hybrid zone

B. Wielstra[1,2,3,*], T. Burke[1], R.K. Butlin[1], O. Schaap[4] H.B. Shaffer[3,5], K. Vrieling[4], J.W. Arntzen[2]

[1] Department of Animal and Plant Sciences, University of Sheffield, S10 2TN Sheffield, UK. [2] Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands. [3] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA. [4] Plant Ecology and Phytochemistry, Institute of Biology Leiden, Leiden University, PO Box 9505, 2300 RA Leiden, The Netherlands. [5] La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095, USA.

* Corresponding author:

Department of Animal and Plant Sciences

University of Sheffield

S10 2TN Sheffield

UK

Email: b.wielstra@sheffield.ac.uk

Abstract Genetic admixture between endangered native and non-native invasive species poses a complex conservation problem. Decision makers often need to quickly screen large numbers of individuals and distinguish natives from morphologically similar invading species and their genetically admixed offspring. We describe a protocol using the fast and economical Kompetitive Allele Specific PCR (KASP) technology for genotyping on a large scale. We apply this protocol to a case study of hybridization between a native and an invasive crested newt species. Using previously published data, we designed a panel of ten nuclear and one mitochondrial one diagnostic SNP markers. We observed only minor differences between KASP and next-generation sequencing data previously produced with the Ion Torrent platform. We briefly discuss practical considerations for tackling the insidious conservation problem of genetic admixture between native and invasive species. The KASP genotyping protocol facilitates policy decision making for the crested newt case and is generally applicable to invasive hybridization with endangered taxa.

Keywords Hybridization; Ion Torrent; KASP genotyping; transcriptome; *Triturus*

Introduction

Interspecific hybridization is frequently mediated by human activities through the disruption of reproductive isolation following habitat modification, climate change and the accidental or deliberate spread of non-native and genetically modified species (Crispo et al. 2011; Harrison & Larson 2014; Lowe et al. 2015). Anthropogenically-induced introgressive hybridization has important implications from a conservation perspective, because it results in genetic replacement and hence a loss of biodiversity at the level of the gene (Petit 2004). How to best deal with this difficult conservation problem is a contentious issue (Jackiw et al. 2015; Wayne & Shaffer submitted). Assuming that the goal of conservation biology is to restore native species genetically, resource managers must be able to quickly and reliably identify both the invader and its genetically admixed offspring. Here we run into a practical problem: because hybridizing species are often phenotypically similar, separating out pure natives from invasives, and particularly from admixed individuals, may not be possible based on morphological criteria. Genetic data are generally used to solve this problem (Allendorf et al. 2001; Allendorf et al. 2010).

Based on a panel of single-nucleotide polymorphisms (SNPs), the extent and degree of genetic admixture of a native species and a non-native invasive can be determined (Garvin et al. 2010). Practical considerations aside, total elimination of partially contaminated populations might be undesirable for conservation reasons (the native species is threatened) or legal reasons (the native species is protected). In such instances, the individuals that qualify for protection might be identified by determining genetic admixture (Allendorf et al. 2001). Depending on the degree of genetic admixture that the legislation permits, individuals can be identified for removal or protection on the basis of their genetic composition – in practice turning an abstract management problem into a set of concrete, actionable decisions. Such a management action would entail catching a large number of individuals and holding them in captivity until their genotype has been established and their fate (remove or release) determined. Such a plan requires accurate genotyping and a fast turnaround measured in hours-to-days, not weeks. Although increasing the number of markers studied would result in higher accuracy, the time involved would increase as well. For many applications, a moderate number of SNPs should be sufficient (Buggs 2007; Currat et al. 2008). Data simulation can be used to guide such management decisions, for example to determine the chance of detecting admixed individuals with different numbers of markers, under particular admixture scenarios.

We present a case study from the Netherlands for which we possess a tissue bank and background knowledge and which highlights the complexity of management of an invasive hybridizing species (Meilink et

al. 2015). In the Veluwe area in the Netherlands, the native northern crested newt *Triturus cristatus* is being displaced by the introduced Italian crested newt *T. carnifex*; in the process the two species hybridize (Fig. 1). The native species is threatened and protected, but difficult to distinguish morphologically from the invader and genetically admixed offspring are even harder to identify (Bogaerts 2002; Meilink et al. 2015). We outline a protocol employing Kompetitive Allele Specific PCR (KASP) technology that allows for a large numbers of individuals to be accurately genotyped for a modest number of SNPs, rapidly and economically.

Materials and Methods

SNP discovery

Genomic data representing both the native and the invasive species could be used directly for *de novo* SNP discovery, but an ascertainment bias might be introduced if these data did not include multiple individuals from across the range of both parent species (Garvin et al. 2010). However, obtaining genome-scale data for many individuals is often not feasible either technically or economically. A workaround is to identify potential variable sequence regions from limited genomic data, sequence these particular regions for a larger set of individuals, and use the resulting dataset for SNP discovery (Wielstra et al. 2014).

Based on a single *Triturus* transcriptome, Wielstra et al. (2014) sequenced 52 short (*ca* 140 bp) nuclear markers positioned in 3'UTR regions of protein-coding genes, and used an Ion Torrent next-generation sequencing protocol to obtain sequence data for three individuals from four populations across the range of both *T. carnifex* and *T. cristatus* (Fig. 1). We used these sequence data for SNP discovery. We focussed on a subset of 15 nuclear markers for which *T. carnifex* and *T. cristatus* had species diagnostic allele variants. Additionally, Wielstra et al. (2013) used Sanger sequencing to obtain sequence data (658 bp) for one mtDNA marker from across the range of *T. carnifex* and *T. cristatus* and we included this marker for SNP discovery. We determined diagnostic SNPs by checking the sequence alignments by eye in MacClade 4.08 (Maddison & Maddison 2005).

Sampling

DNA was extracted with the Qiagen Dneasy 96 Blood & Tissue Kit and the resulting DNA extract was used at a 1:49 dilution. We genotyped 156 crested newts in total to test our protocol (Fig. 1; Table S1). We sampled the 12 individuals of both parental species on which marker design was based and the 132 individuals studied by Meilink et al. (2015) from 11 populations from the Veluwe in the Netherlands (12 individuals per population).

The Veluwe area is positioned well within the native range of *T. cristatus* and *T. carnifex* has been introduced there locally, presumably in the late 1970s, near the village of Vaassen (Bogaerts 2002). Meilink et al. (2015) demonstrated that the 11 sampled populations ranged from pure native (*T. cristatus*) via different degrees of genetic admixture to pure invasive (*T. carnifex*).

SNP validation

Genotyping was conducted at the SNP genotyping facility of the Institute of Biology, Leiden University, using the Kompetitive Allele-Specific PCR (KASP) genotyping system (LGC genomics, UK). The KASP technology encompasses fluorescence-based genotyping (Semagn et al. 2014). The SNP variant present in each individual (both variants in the case of a heterozygote) is determined in uniplex assays, based on two allele-specific primers with a final base complementary to one of the two potential SNPs, that also possess a unique tail sequence. Different fluorescently labelled primers present in the KASP master mix correspond to each tail sequence and are activated when incorporated during subsequent PCR cycles, with further cycling causing signal intensity to increase. No library preparation, DNA sequencing and bioinformatics is involved. The KASP technology is economical and accurate compared to other SNP genotyping technologies (Semagn et al. 2014). We designed primers using the Kraken software (LGC genomics, UK) and ordered them from Integrated DNA Technologies.

The PCR mix was made according to the manufacturer's instructions. The LGC genomics KASP mix 4.0 was used and PCRs were carried out in 1536 wells plates with a reaction volume of 1 ul in a hydrocycler. The PCR consisted of an initial step of 15 min at 94 ºC, followed by 10 cycles of 20s at 94 ºC and 60s at 61 ºC, followed by 26 cycles of 20s at 94 C and 60s at 55 ºC. After these 36 cycles fluorescence was measured on a Pherastar Plate reader. PCR was continued and after 39, 42 and 45 cycles fluorescence was measured again to follow the trajectory of all samples. Genotypes were automatically called using the Kraken software, visually inspected and occasionally manually corrected.

We visualized the KASP genotyping data for the nuclear markers using the R package HIest (Fitzpatrick 2012), which determines the genomic composition of hybrids based on ancestry (the fraction of alleles derived from each parental species) and heterozygosity (the fraction of loci heterozygous for alleles from each parental species).

Comparison with Ion Torrent dataset

The nuclear markers in which the targeted SNPs were positioned were previously sequenced using Ion Torrent next-generation sequencing for the same individuals (Wielstra et al. 2014; Meilink et al. 2015). We compared the results of the two genotyping strategies for each marker–individual combination to identify potential differences between the KASP genotyping and Ion Torrent protocols. Similarly, mtDNA sequences were previously obtained by Sanger Sequencing (Wielstra et al. 2013; Meilink et al. 2015) and compared with the KASP genotyping results.

Data simulation

We used HybridLab (Nielsen et al. 2006) to visualize the detectability of alien alleles under a scenario of repeated backcrossing to a native species using different numbers of diagnostic nuclear markers. We modeled 1000 offspring each for 10 generations of backcrossing for 10, 50, 100 and 500 diagnostic nuclear markers.

Data accessibility

The Ion Torrent next-generation sequence data of Wielstra et al. (2014) used for SNP discovery are available from Dryad Digital Repository entry http://dx.doi.org/10.5061/dryad.36775. Sampling details and a summary of genotyping results of the KASP genotyping and Ion Torrent next-generation sequencing protocol are provided in Table S1. Sequence alignments and Kraken input for all markers are shown in Table S2. Raw output of the KASP genotyping protocol is provided in Table S3. GenBank Accession numbers for mtDNA haplotypes can be found in Table S4. Plotted output of the KASP genotyping protocol is visualized in Fig. S1.

Results

For the mitochondrial marker an assay was successfully designed and all individuals could be genotyped (Fig. 1B). Allocation to species mtDNA type was identical to results previously obtained by Sanger sequencing (Table S1).

Of 15 candidate nuclear markers, one was dropped because no single diagnostic SNP could be identified for primer design. For the remaining 14 diagnostic nuclear markers assays could be designed (see Table S2 for details). Genotypes for two of these markers (amot and ibtk) could not be unambiguously called (Table S3; Fig. S1). A further two nuclear SNP markers (cnppd and col18) showed a single instance of heterozygosity in a parental individual (Table S3; Fig. S1), suggesting either a genotyping error or that the markers were not fully diagnostic. These four markers were excluded from further analysis.

For the ten remaining nuclear markers, 1472 out of 1560 (94.4%) attempted SNP calls provided unambiguous results (Table S1). For the remaining 88 attempted SNP calls, in three instances the PCR failed, in 31 instances no distinction could be made with confidence between heterozygotes and homozygous natives (i.e. the presence of 0 or 1 alien alleles) and in 54 instances we could not distinguish between heterozygous and homozygous invasives (1 or 2 alien alleles). Hence, 54 + 1472 = 1526 (97.8%) of attempted SNP calls provided reliable information on the presence of alien alleles (Table S1). Ancestry and heterozygosity for parental and hybrid populations based on the KASP genotyping of the nuclear SNPs (excluding SNP calls in which homozygotes and heterozygotes could not be distinguished) are visualized in Fig. 2.

For the Ion Torrent protocol for 1497 out of 1560 (96.0%) marker-individual combinations two native and/or invasive alleles could be genotyped unambiguously. Ten marker-individual combinations had no data. An issue particular to the Ion Torrent protocol was that new alleles that were not previously identified in the parental species were found and hence could not be allocated to either species. For ten marker-individual combinations neither allele could be allocated to species, for 21 only one native allele was identified, and for 22 the presence of one alien allele could be determined. Hence, for 1497 + 22 = 1519 (97.4%) marker-individual combinations we obtained reliable information on the presence of alien alleles (Table S1).

Genotyping efforts with the KASP genotyping and Ion Torrent protocol provided different calls for 46 out of 1560 (2.9%) genotyping efforts (Table S1). For 17 incongruent genotyping efforts there was disagreement whether zero or one alien alleles were present. For the remaining 29 genotyping efforts both protocols agreed upon the presence of alien alleles, but disagreed on whether one or two were present. Agreement between the two datasets on the presence of alien alleles using Cohen's kappa (the number of genotypes consistently called as with or without alien alleles by both protocols, divided by the total number of genotypes called with both protocols), ranged from 0.97 to 1 for individual markers, and for the total dataset Cohen's kappa was 0.99 (Table S1). The differences between genotype methods occured in 12 out of 156 (7.7%) individuals, for which Cohen's kappa ranges from 0.7-0.9 (Table S1). For all these individuals the presence of alien alleles was established unambiguously for some, but not all of the markers. At the level of the total dataset, the KASP genotyping and Ion Torrent protocols agreed that for 128 individuals alien alleles were present whereas for 27 individuals neither protocol could confirm the presence of alien alleles; for one individual the presence of a single alien allele identified with the Ion Torrent protocol could not be confidently established with the KASP genotyping protocol.

Results of the data simulation are presented in Fig. 3 and illustrate that as the number of diagnostic markers increases, alien alleles can be detected over more generations of backcrossing with the native species. For example, the number of generations of backcrossing for which at least 95% of individuals are correctly identified as containing at least one alien allele is one for 10 diagnostic markers, four for 50 markers, five for 100 markers and seven for 500 markers.

Discussion

Performance of the KASP genotyping protocol

A genotyping method aiming to identify alleles derived from an invasive species should minimize both false negatives (missing alien alleles when they are really present) and false positives (claiming an allele is non-native when it is actually native). Failing to identify and remove a proportion of the genetically admixed individuals, or inadvertently removing pure natives, would both compromise the management goal of restoring native genetic integrity. However, effective management may also require rapid genetic testing and swift decisions, which may come with high cost or error rates. The KASP genotyping strategy performed well compared to a more expensive and time-consuming Ion Torrent next-generation sequencing pipeline (Wielstra et al. 2014; Meilink et al. 2015). In spite of minor differences between the two protocols, at the level of the total dataset agreement is essentially perfect and only a single, slightly admixed individual yielded equivocal results.

A complication we encountered with KASP genotyping in some of the attempted SNP calls was that it was not always possible to distinguish heterozygotes from homozygotes. Although polyploidy and paralogy could underlie this pattern we consider this unlikely because: (1) polyploids are rare in salamanders and have virtually never been recorded in *Triturus* newts (Borkin et al. 1996; Litvinchuk et al. 2001; Mable et al. 2011), (2) the initial marker design phase excluded multicopy genes (Wielstra et al. 2014), and (3) the distribution of unclear SNP calls was not linked to any particular individuals or markers. An alternative explanation is that ambiguous SNP calls are a consequence of the large genome size of our study system (c. 20-30 Gb in the genus *Triturus*; Gregory 2016). The larger the genome, the lower the concentration of the targeted SNPs per unit of DNA, which may compromise SNP calling. Another project on *Triturus* using DNA extract diluted 1:9 rather than 1:49 yielded more consistently interpretable results (BW, unpublished data). This suggest that our starting amount of DNA was on the low side and we predict elevating the starting DNA levels would increase genotyping performance.

We consider KASP genotyping a reliable and, when a large number of individuals has to be genotyped for a moderate number of SNPs, preferable tool for the detection of genetic admixture between native and invasive species. Compared to the Ion Torrent and other next-generation protocols the KASP genotyping protocol has some major advantages (Semagn et al. 2014). KASP genotyping is: (1) fast (at the SNP genotyping facility of the Institute of Biology, Leiden University, 100,000 datapoints can be genotyped in a single day, so for ten markers 10,000 individuals can be genotyped), (2) relatively affordable if a small number of SNPs is sufficient (at about 0.30 € per SNP, a study with ten markers amounts to 3.0 € per individual), (3) simple (PCRs are conducted per individual per marker using an automated pipeline and genotypes are automatically called and available in a user-friendly format), and (4) scalable (SNP calls can be repeated or extra SNPs can easily be genotyped at a later stage).

The rate-limiting step in our protocol is DNA extraction. If there is no requirement to keep extracted DNA after an individual has been genotyped, a fast and cheap but destructive Chelex extraction could be used (including labor and plastics about 0.50 € per individual, a twenty-fold decrease in costs compared to the extraction method used in the present study; we estimate c. 1000 individuals could be extracted in a day). Another *Triturus* project using the Chelex method yields high quality genotyping results with the KASP platform (J.W. Arntzen, unpublished data).

Implementing the KASP genotyping protocol

If anthropogenic hybrids are considered a threat to native biodiversity and are of low conservation value, eradicating them could be regarded as an ethical conservation strategy (Jackiw et al. 2015). For situations where the invasive species and their hybrids closely resemble native endangered species, molecular data are often pertinent in deciding which populations are to be protected, which are eligible for elimination and, in a more complicated scenario, which are partially contaminated and need pure native individuals to be filtered out (Rhymer & Simberloff 1996). The threshold that should be applied for marking a population or individual as 'worthy to protect' versus 'to be eliminated' is eventually a juridical decision (Fitzpatrick et al. 2015). Data simulation can help to determine the number of markers required to reach a particular threshold with a certain probability under definable assumptions about breeding probabilities (for example, all individuals mate randomly with respect to genotype) and survival likelihoods (essentially that native and alien alleles are neutral with respect to fitness). Additionally, molecular techniques are crucial in monitoring the effectiveness of conservation actions, and each case study may suggest different genetic thresholds, with different goals, to

achieve a specified set of management outcomes (Wayne & Shaffer submitted). We agree with Fitzpatrick and colleagues (2015) that tackling the complex management of genetic admixture between native and invasive species requires interactions between scientists, legislators and conservationists and ideally that all parties involved have an understanding of genetic results.

After the introduction of an invasive hybridizing species, introgression is (fortunately) expected to be biased towards the expanding invasive species (Buggs 2007; Currat et al. 2008). In line with this prediction, with the relatively low number of markers used in this study we still manage to detect unambiguous genetic admixture between native and invasive species throughout most affected populations. Yet, a potential difficulty in recognizing genetically admixed individuals, particularly at the advancing frontier of an invasion, is that repeated backcrossing to the native species would cause invasive DNA to become diluted (Fig. 3). With ten nuclear markers, as used in the present study, the power of detecting alien alleles rapidly decreases under a scenario of backcrossing to the native species (Fig. 3). Our genetic data highlight some individuals that would benefit from being studied with a larger number of nuclear markers: in populations 8 and 10 (Figs. 1 and 2) individuals that are unambiguously genetically admixed are found in syntopy with individuals for which no alien alleles were observed. When using the KASP genotyping protocol, we suggest a two-step approach, with an initial scan using a small number of nuclear markers (about ten) to identify strongly admixed populations, and a subsequent higher resolution screening employing a larger number of markers for a (presumably much smaller) subset of populations that is highlighted as mildly admixed. This strategy quickly identifies the most problematic populations, although it may not lead to immediate conservation actions in populations with low levels of admixture or where a smaller number of highly selected alien alleles exists (Fitzpatrick & Shaffer 2007; Fitzpatrick et al. 2009).

The native, threatened *T. cristatus* is a European wildlife icon that is threatened by invasive hybridization with *T. carnifex* (Meilink et al. 2015). The Netherlands has an (inter)national responsibility to protect *T. cristatus* and the availability of the KASP protocol should facilitate policymakers as they formulate management decisions for the Dutch *Triturus* situation. The molecular toolkit introduced here is directly applicable to other *T. carnifex* introductions that have occurred in the range of *T. cristatus* (Fig. 1); these are known from Geneva, France/Switzerland (Arntzen 2001; Dufresnes et al. 2016); Munich, Germany (Maletzky et al. 2008); and Surrey, England (Brede et al. 2000; Brede 2015). It also may prove useful in other situations where advancing introgression is occurring, and where the front-line defence may include removing admixed individuals based on their genotypic composition (Searcy et al. 2016).

Nowadays, genomics plays a crucial role in biological invasion studies (Rius et al. 2015) and genetic screening is a key component of management actions concerning invasive hybridizers (Chown et al. 2015; Shaffer et al. 2015). KASP genotyping is an efficient way to obtain genetic data: it can be used to genotype thousands of individuals rapidly, at a reasonable cost (Semagn et al. 2014). Provided that genomic data for SNP discovery are available, the KASP genotyping protocol can easily be adapted to other cases of invasive hybridizing species. With the identification of invasive hybridizing species and their genetically admixed offspring no longer an empirical challenge, the main remaining hurdle in tackling the difficult conservation issue of anthropogenic hybridization is drafting clear, implementable policy alternatives (Wayne & Shaffer submitted).

Acknowledgements

BW conducted part of this work as a Newton International Fellow and is currently supported by a Marie Skłodowska-Curie Individual Global Fellowship (EU project 655487). TB is a Leverhulme Fellow. JWA is supported by NWO (ALW project 824.14.014) and HBS receives support from the NSF (DEB 1239961). Reptile, Amphibian and Fish Conservation Netherlands (RAVON) contributed the tissue samples and the distribution data shown in the inset of Figure 1.

Compliance to ethical standards

The authors declare that they have no conflict of interest. Non-invasive tissue sampling was conducted with legal consent (WoD RAVON 12.01).

References

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet, 11, 697-709.
Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. Trends Ecol Evol, 16, 613-622.
Arntzen JW (2001) Genetic variation in the Italian crested newt, *Triturus carnifex*, and the origin of a non-native population north of the Alps. Biodivers Conserv, 10, 971-987.
Bogaerts S (2002) Italian crested newts, *Triturus carnifex*, on the Veluwe, Netherlands. Z Feldherpetol, 9, 217-226.
Borkin LJ, Litvinchuk SN, Rosanov JM (1996) Spontaneous triploidy in the crested newt, *Triturus cristatus* (Salamandridae). Russ J Herpetol, 3, 152-156.
Brede EG (2015) Beam Brook revisited: a molecular study of a historically introduced non-native amphibian (*Triturus carnifex*) and its limited introgression into native UK *Triturus cristatus* populations. Amphibia-Reptilia, 36, 287-299.
Brede EG, Thorpe RS, Arntzen JW, Langton TES (2000) A morphometric study of a hybrid newt population (*Triturus cristatus*/*T. carnifex*): Beam Brook Nurseries, Surrey, U.K. Biol J Linn Soc, 70, 685-695.
Buggs RJA (2007) Empirical study of hybrid zone movement. Heredity, 99, 301-312.

Chown SL, Hodgins KA, Griffin PC, Oakeshott JG, Byrne M, Hoffmann AA (2015) Biological invasions, climate change and genomics. Evol Appl, 8, 23-46.
Crispo E, Moore J-S, Lee-Yaw JA, Gray SM, Haller BC (2011) Broken barriers: human-induced changes to gene flow and introgression in animals. Bioessays, 33, 508-518.
Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. Evolution, 62, 1908-1920.
Dufresnes C, Pellet J, Bettinelli-Riccardi S, Thiébaud J, Perrin N, Fumagalli L (2016) Massive genetic introgression in threatened northern crested newts (*Triturus cristatus*) by an invasive congener (*T. carnifex*) in Western Switzerland. Conserv Genet.
Fitzpatrick BM (2012) Estimating ancestry and heterozygosity of hybrids using molecular markers. BMC Evol Biol, 12, 131.
Fitzpatrick BM, Johnson JR, Kump DK, Shaffer HB, Smith JJ, Voss SR (2009) Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. BMC Evol Biol, 9, 176.
Fitzpatrick BM, Ryan ME, Johnson JR, Corush J, Carter ET (2015) Hybridization and the species problem in conservation. Curr Zool, 61, 206-216.
Fitzpatrick BM, Shaffer HB (2007) Hybrid vigor between native and introduced salamanders raises new challenges for conservation. Proc Natl Acad Sci USA, 104, 15793-15798.
Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. Mol Ecol Resour, 10, 915-934.
Gregory TR (2016) Animal Genome Size Database. http://www.genomesize.com. Accessed 3 January 2016
Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. J Hered, 105, 795-809.
Jackiw RN, Mandil G, Hager HA (2015) A framework to guide the conservation of species hybrids based on ethical and ecological considerations. Conserv Biol, 29, 1040-1051.
Litvinchuk SN, Rosanov JM, Borkin LJ (2001) Natural autotriploidy in the Danube newt, *Triturus dobrogicus* (Salamandridae). Russ J Herpetol, 8, 74-76.
Lowe WH, Muhlfeld CC, Allendorf FW (2015) Spatial sorting promotes the spread of maladaptive hybridization. Trends Ecol Evol, 30, 456-462.
Mable BK, Alexandrou MA, Taylor MI (2011) Genome duplication in amphibians and fish: an extended synthesis. J Zool, 284, 151-182.
Maddison DR, Maddison WP (2005) MacClade: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.
Maletzky A, Mikulíček P, Franzen M, Goldschmid A, Gruber H-J, Horák A, Kyek M (2008) Hybridization and introgression between two species of crested newts (*Triturus cristatus* and *T. carnifex*) along contact zones in Germany and Austria: morphological and molecular data. Herpetol J, 18, 1-15.
Meilink WRM, Arntzen JW, van Delft JJCW, Wielstra B (2015) Genetic pollution of a threatened native crested newt species through hybridization with an invasive congener in the Netherlands. Biol Concerv, 184, 145-153.
Nielsen EE, Bach LA, Kotlicki P (2006) Hybridlab (version 1.0): a program for generating simulated hybrids from population samples. Mol Ecol Notes, 6, 971-973.
Petit RJ (2004) Biological invasions at the gene level. Divers Distrib, 10, 159-165.
Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annu Rev Ecol Syst, 27, 83-109.
Rius M, Bourne S, Hornsby HG, Chapman MA (2015) Applications of next-generation sequencing to the study of biological invasions. Curr Zool, 61, 488–504.
Searcy CA, Rollins HB, Shaffer HB (2016) Ecological equivalency as a tool for endangered species management. Ecological Applications, 26, 94-103.
Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breeding, 33, 1-14.
Shaffer HB, Gidiş M, McCartney-Melstad E, Neal KM, Oyamaguchi HM, Tellez M, Toffelmier EM (2015) Conservation genetics and genomics of amphibians and reptiles. Annu Rev Anim Biosci, 3, 113-138.
Wayne RK, Shaffer HB (submitted) Hybridization and endangered species protection in the molecular era.
Wielstra B, Crnobrnja-Isailović J, Litvinchuk SN, Reijnen BT, Skidmore AK, Sotiropoulis K, Toxopeus AG, Tzankov N, Vukov T, Arntzen JW (2013) Tracing glacial refugia of *Triturus* newts based on mitochondrial DNA phylogeography and species distribution modeling. Front Zool, 10, 13.
Wielstra B, Duijm E, Lagler P, Lammers Y, Meilink WRM, Ziermann JM, Arntzen JW (2014) Parallel tagged amplicon sequencing of transcriptome-based genetic markers for *Triturus* newts with the Ion Torrent next-generation sequencing platform. Mol Ecol Resour, 14, 1080-1089.

Figure legends

Fig. 1 Natural distribution and sampling scheme of two crested newt species that show anthropogenically induced hybridization. Panel a: The range of the northern crested newt *Triturus cristatus* (in red) and the range of the Italian crested newt *T. carnifex* (in blue). The natural populations of both species used for SNP discovery are denoted with black dots and the known introductions of *T. carnifex* into the range of *T. cristatus* are denoted with white dots. The inset shows the area lifted out in panel b. Panel b: Sampling of crested newt populations for one of the introduction sites, the Veluwe in the Netherlands. Each pie reflects a population of 12 individuals and pie slices are coloured either red (native *T. cristatus* mtDNA) or blue (invasive *T. carnifex* mtDNA). The presumed introduction site is indicated with a black star. Sampling details are provided in Table S1

Fig. 2 Results from the HIest analysis of populations from the Veluwe in the Netherlands, where anthropogenically induced hybridization between two crested newt species occurs, based on KASP genotyped nuclear SNPs. Results for populations 1-3 are combined and results for populations 4-11 are shown for each population. The lower left corner of each triangle corresponds to a pure native genotype, the lower right corner to a pure invasive genotype, and the upper corner to a pure F1 hybrid. Individuals with identical genotypes overlap and cannot be discerned. Population numbers correspond to Fig. 1B

Fig. 3 Detecting genes of an invasive hybridizer under a scenario of repeated backcrossing to a native species. After an initial hybridization event individuals are heterozygous native/invasive for all markers (F1). Over ten generations of backcrossing to the native species (B1-10) the proportion and detectability of invasive alleles continually decreases, but the ability to detect at least some non-native ancestry decreases more slowly as more markers are studied. The dotted line indicates the 95% probability level for detecting at least one alien allele in a sampled individual with random mating and no fitness differences between native and non-native alleles

Supplementary data

Table S1 Sampling details, summary of output and comparison of genotyping methods
Table S2 Sequence alignments and Kraken input
Table S3 Raw KASP genotyping output
Table S4 GenBank Accession numbers for mtDNA
Figure S1 Plotted KASP genotyping output