Univerzita Karlova v Praze, Přírodovědecká fakulta
Katedra zoologie

Charles University in Prague, Faculty of Science
Department of Zoology

Doktorský studijní program: Zoologie
Ph.D. Study Program: Zoology

Autoreferát disertační práce
Summary of the Ph.D. Thesis

Genomová analýza hybridní zóny myši domácí
Whole-genome analysis of the house mouse hybrid zone

Mgr. Václav Janoušek

Školitel/Supervisor:
Doc. Mgr. Pavel Munclinger, Ph.D.

Praha, 2015

Abstrakt

Dva druhy myší (*Mus musculus musculus*/*Mus musculus domesticus*) tvoří napříč střední Evropou úzkou zónu sekundárního kontaktu, která, jak se předpokládá, je udržována prostřednictvím rovnováhy mezi selekcí hybridů se sníženým fitness a disperzí jedinců. V průběhu mého PhD studia jsem spolu s mými kolegy používal set asi 1400 SNP markerů ke studiu celogenomové introgrese napříč dvěma/třema transekty v hybridní zóně myši domácí. Naším cílem bylo identifikovat genomové oblasti obsahujících geny potenciálně způsobující reprodukční izolaci mezi dvěma poddruhy myši domácí, charakterizovat jejich rozložení v myším genomu a charakterizovat vlastnosti genomu typické pro tyto regiony. Náš výzkum potvrdil na celogenomové škále významnost chromozomu X při evoluci reprodukční izolace a její společný genetický základ mezi dvěma studovanými transekty. V proximální části chromozomu 11 vykazující negativní epistázy v obou transektech jsme identifikovali gen *1700093K21Rik* jako vhodný kandidátní speciační gen. V další studii jsme znovu analyzovali data ze tří transektů ve snaze charakterizovat vlastnosti genomu spojené s oblastmi s rozdílnou introgresí. Oblasti s nízkou mírou introgrese měly tendenci mít vyšší míru genomové diferenciace a naopak nižší míru rekombinace. Tyto oblasti také obsahovaly s vyšší pravděpodobností geny plnící úlohy uvnitř buňky. Genomové oblasti s vysokou mírou introgrese měly naopak tendenci mít nižší míru diferenciace, vyšší míru rekombinace a větší zastoupení genů plnící funkce na buněčné periferii. Na základě našich výsledků se domníváme, že funkční organizace genomu by mohla být důležitým činitelem při druhové divergenci a při evoluci reprodukční izolace.

Abstract

Two house mouse subspecies (*Mus musculus musculus*/*Mus musculus domesticus*) form a narrow zone of secondary contact across Central Europe which is thought to be maintained by a balance between selection against unfit hybrids and dispersion of individuals. During my PhD study my collaborators and I used an array of ~ 1400 SNP markers to study patterns of introgression on a genome-wide scale across two/three house mouse hybrid zone transects. Our aim was to identify genomic regions putatively harboring genes that are involved in reproductive isolation between the two subspecies, to characterize their distribution in mouse genome and to assess genomic features associated with them. We were able to confirm on a genome-wide scale the importance of the X chromosome in the evolution of reproductive isolation and its common genetic basis between the two transects. In the proximal part of chromosome 11 which were exhibiting negative epistasis in both transects we were able to identify a particular gene, *1700093K21Rik,* as a good candidate for the speciation gene. We further re-analyzed the data of three transects to identify genomic regions of high and low introgression and characterized genomic features associated with these regions. We found that genomic regions of low introgression exhibited a higher genomic differentiation and a low rate of recombination. These regions are also more likely to accommodate genes acting in the interior of a cell. On the contrary, genomic regions of high introgression exhibited a lower genomic differentiation, a higher rate of recombination and a higher prevalence of genes acting at cell periphery. We hypothesize that the functional organization of the genome is an important driver of species divergence and in the evolution of reproductive isolation.

Úvod

Speciace – proces, kterým dochází ke vzniku nových druhů – patří v současnosti k jednomu z nejvíce zkoumaných evolučních fenoménů. Jedním z cílů studia speciace je pochopení molekulární podstaty vzniku reprodukční izolace mezi nově vznikajícími druhy (Presgraves 2010). Myš domácí představuje z tohoto pohledu velmi vhodný modelový organismus. Dva poddruhy myší domácí (*M. m. musculus*/*M. m. domesticus*), které od sebe začali divergovat zhruba před 500 tisíci lety (Geraldes et al. 2008) na území dnešního Iránu (Duvaux et al. 2011, Hardouin et al. 2015), se setkali ve střední Evropě, kde vznikla zóna sekundárního kontaktu (Boursot et al. 1993). Tyto dva poddruhy byly po většinu času divergence geograficky izolovány (Duvaux et al. 2011), což vedlo k evoluci velmi dobře zdokumentované částečné postzygotické reprodukční izolace způsobené sníženou fertilitou u samců (Forejt & Iványi 1974, Storchová et al. 2004, Britton-Davidian et al. 2005, Good et al. 2008a, 2008b, White et al. 2011, Albrechtová et al. 2012, Turner et al. 2012). Předpokládá se, že tato reprodukční izolace je zodpovědná za udržování hybridní zóny na základně rovnováhy mezi disperzí jedinců s rodičovských populací a selekcí proti hybridům se sníženým fitness (Barton & Hewitt 1985).

Hybridní zóny představují prostředek ke studiu nově vznikajících druhů a pro pochopení procesu speciace (Barton & Hewitt 1985). Teoretické modely (Barton 1979, Barton 1983, Barton & Gale 1993) popisují genový tok mezi dvěma poddruhy jako klinální změnu alelické frekvence napříč hybridní zónou. Strmost této kliny odpovídá síle reprodukční bariéry mezi těmito poddruhy a tedy množství genového toku, ke kterému mezi nimi dochází. Míra genového toku se však liší mezi jednotlivými částmi genomu podle genového obsahu. Oblasti obsahující geny zodpovědné za reprodukční izolaci jsou v hybridní zóně pod selekcí a mají tak genový tok nižší. Neutrální oblasti mají výrazně vyšší šanci, že projdou přes hybridní zónu. Výše genového toku v těchto oblastech však závisí na síle celkové reprodukční bariéry a jejich fyzické vazbě ke genům pod selekcí. Třetí skupinou jsou geny, které poskytují nositelům selekční výhodu i na cizím genetickém pozadí, a mohou tak snadno proniknout přes reprodukční bariéru (Piálek & Barton 1997). Díky těmto vlastnostem představují hybridní zóny velmi efektivní prostředek k identifikaci genů zodpovědných za reprodukční izolaci mezi nově vznikajícími druhy (Rieseberg & Buerkle 2002).

V současnosti vývoj nových sekvenačních technologií značně zjednodušil analýzu velkého množství genomových dat. Během mého PhD studia jsem spolu s mými kolegy využil data vyprodukovaná pomocí sekvenačních technologií nové generace

a výše uvedených vlastností hybridních zón k pochopení vzniku reprodukční izolace u myši domácí.

Cíle práce

- Charakterizovat rozdíly mezi introgresí chromozomu X a autosomů na celogenomové úrovni ve snaze objasnit jejich roli v evoluci reprodukční izolace u myši domácí

- Identifikovat konkrétní genomové oblasti vykazující introgresi odpovídající selekci proti hybridům v hybridní zóně myši domácí a najít kandidátní geny zodpovědné za reprodukční izolaci u myši domácí

- Porovnat introgresi napříč myším genomem s genomickými vlastnostmi jako je míra rekombinace, míra genomické diferenciace a funkční složení genů ve snaze odhalit vztahy těchto vlastností a reprodukční izolace u myši domácí

Materiál a metody

K celogenomové analýze hybridní zóny myši domácí jsem během svého PhD studia využil dataset vyprodukovaný v průběhu několika let na základě spolupráce laboratoří z USA (University of Michigan, University of Arizona) a z České Republiky (Ústav biologie obratlovců, AV ČR, v.v.i., Ústav živočišné fyziologie a genetiky, AV ČR, v.v.i., Karlova Univerzita v Praze). Dataset obsahoval vzorky ze tří transektů hybridní zóny myši domácí. Česko-Bavorský transekt (CZ) obsahoval 895 vzorků, Bavorsko-Rakouský transekt (BV) obsahoval 437 vzorků a Saský transect (SX) obsahoval 322 vzorků. Vzorky byly sbírány v průběhu mnoha let různými výzkumnými týmy. Detaily týkající se jednotlivých vzorků a lokalit pro jednotlivé transekty lze nalézt v Teeter et al. (2008, 2010) a Wang et al. (2011).

K analýze hybridní zóny byly využity diagnostické 'Single Nucleotide Polymorphisms' (SNPs) detailně popsané Wangem et al. (2011), k jejichž vývoji byla využita beta verze 'Mouse Diversity Array' (Yang et al. 2009) a genotypy pro vzorky divokých myší z území Evropy (Yang et al. 2011). Jako diagnostické byly považovány SNP markery rozdílně fixované mezi sedmi vzorky *M. m. musculus* a devíti vzorky *M. m. domesticus* (Yang et al. 2011). Finální set diagnostických markerů sestával z asi 1400 SNP markerů rovnoměrné distribuovaných napříč myším genomem a hlavní část genotypů z hybridní zóny myši domácí byla publikována ve Wang et al. (2011).

Introgrese napříč hybridní zónou byla analyzována pomocí dvou metod genomických klin – likelihood přístup (Gompert & Buerkle 2009, 2010) a Bayesiánský přístup (Gompert & Buerkle 2011, 2012). Genomické kliny popisují pravděpodobnost vyskytu genotypu/alely daného původu pro daný locus vzhledem průměrnému genotypu/alele pro daného jedince. Průměrný genotyp/alela je reprezentována hybridním indexem (Gompert & Buerkle 2009).

Pro další analýzy byla využita data volně dostupná v bioinformatických on-line databázích: Mouse Diversity Array (Yang et al. 2009, 2011), myší genetická mapa (Shifman et al. 2006, Cox et al. 2009), Gene Ontology (Ashburner et al. 2000).

Výsledky a diskuze

Během svého PhD studia jsem spolu s mými kolegy analyzoval introgresi asi 1400 SNP markerů v hybridní zóně myši domácí. V první fázi jsme využívali data ze dvou transektů (CZ, BV), u kterých Wang et al. (2011) prokázal dostatečný počet vzorků v centrálních populacích. Naším cílem byla charakterizace celogenomové architektury reprodukční izolace mezi dvěma poddruhy myši domácí (*M. m. musculus*/*M. m. domesticus*). Nejprve jsme charakterizovali detailní introgresi na tomto chromozomu porovnáváním introgrese mezi páry markerů. Cílem bylo nalezení větších úseků (bloků), které vykazují podobnou introgresi na chromozomu X. Dále byla analyzována introgrese markerů uvnitř bloků vzhledem celogenomickému pozadí. Lokalizace nalezených bloků pokrývajících značnou část chromozomu X se mezi transekty shodovala a markery tvořící tyto bloky vykazovaly v obou transektech negativní epistázi. Výsledky naší analýzy tak potvrdily společnou reprodukční bariéru na chromozomu X mezi oběma transekty. Naše výsledky korespondovaly s výsledky předchozích studií, které přisuzovaly chromozomu X vyšší míru selekce v hybridní zóně než autozomům (Tucker et al. 1992, Dod et al. 1993, Macholán et al. 2007). Největší z námi nalezených bloků v centrální části chromozomu X odpovídá výsledkům starších prací využívající nižší počet markerů a analýzu geografických klin (Payseur et al. 2004, Dufková et al. 2011). X vázáné markery byly dále využity k identifikaci SNP markerů vykazující negativní epistázi na autozomech. Jejich rozsah činil v obou transektech 20-25 %. Jejich malý překryv mezi transekty (pouhých 5 %) byl také v kontrastu s chromozomem X a poukazoval spíše na geografickou variabilitu v reprodukční izolaci. Náš nález odpovída současné speciační teorii tzv. 'the large X-effect', která přisuzuje chromozomu X významně větší roli při formování reprodukční izolace než autozomům (Coyne & Orr 1989, Charlesworth et al. 1987).

Na základě našich předchozích výsledků jsme vybrali oblast v proximální části chromozomu 11 vykazující negativní epistázy v obou transektech, a u které byla také v minulosti prokázána asociace této oblasti se samčí hybridní sterilitou (Oka et al. 2007). Výstupem detailní analýza této oblasti bylo nalezení genu *1700093K21Rik*, který vykazoval charakteristiky typické pro kandidátní speciační geny. Cílem této studie bylo ukázat efektivnost kombinace dat z hybridní zóny a veřejně dostupných myších genomických zdrojů.

Recentní studie Renaut et al. (2013) poukázala na možný význam funkční architektury genomu při druhové divergenci několika druhů slunečnic. Ve finální studii jsme se zaměřili na genomické vlastnosti související s funkční architekturou genomu v oblastech s nízkou a vysokou introgresí. Výsledkem bylo nalezení statisticky signifikantního rozdílu pro míru genomické diferenciace, míru

rekombinace a funkční složení genů mezi oblastmi s různou mírou introgrese. Genomové oblasti vykazující nízkou míru introgrese byly charakteristické vyšší mírou diferenciace, nižší mírou rekombinace a větším zastoupením genů plnících úlohy uvnitř buňky. Z hlediska molekulární funkce byly tyto geny častěji asociovány s funkcemi týkajícími se regulace transkripce. Genomové oblasti s vysokou mírou introgrese naopak vykazovaly nízkou genomovou diferenciaci, vyšší míru rekombinace a vyšší zastoupení genů mající funci na buněčné periferii. Molekulární funkce těchto genů byla spojená s transdukcí signálu v buňce naznačující receptorovou aktivitu. Olfaktorické a vomeronasální receptory byly často zastoupené mezi těmito geny. Rozdíl v aktivitě genů uvnitř nebo na povrchu buňky poukazuje na významnou roli funkčního uspořádání genomu při evoluci reprodukční izolace a při druhové divergenci.

Závěry

- Celogenomová analýza hybridní zóny mezi dvěma poddruhy myši domácí (*M. m. musculus*/*M. m. domesticus*) ukázala na významný vliv pohlavního chromozomu X ve formování reprodukční izolace u myši domácí napříč transekty.
- Identifikovali jsme řadu autozomálních oblastí, které vykazovali známky negativní epistáze. V každém transektu bylo takto identifikováno mezi 20 – 25 % markerů, jejich překryv mezi transekty byl však výrazně nižší (~ 5 %), což poukazovalo spíše na geografickou variabilitu v reprodukční izolaci na autozomech.
- Analýza oblasti v proximální části chromozomu 11 vykazující negativní epistázi pomocí veřejně dostupných genomických zdroji umožnila identifikaci genu *1700093K21Rik* významně korelujícího s reprodukční izolací mezi dvěma poddruhy myši domácí
- Analýza genomických vlastností (míra genomické diferenciace, míra rekombinace a funkční složení genů) asociovaných s různou mírou introgrese na autozomech naznačila možný vliv funkční organizace genomu při druhové divergenci a formování reprodukční izolace

Introduction

Speciation is a process by which new species arise and it is currently one of the most researched fields of evolutionary biology. One of the aims of speciation research is to understand the molecular basis of reproductive isolation (Presgraves 2010). From this perspective the house mouse is a good model species. Two house mouse subspecies (*M. m. musculus*/*M. m. domesticus*) started to diverge approximately 500,000 years ago (Geraldes et al. 2008) in the region of present-day Iran (Duvaux et al. 2011, Hardouin et al. 2015). They subsequently met in Central Europe where they formed a zone of secondary contact (Boursot et al. 1993). These two subspecies were geographically isolated for most of the time of divergence (Duvaux et al. 2011) which led to the evolution of well characterized partial postzygotic reproductive isolation, with males having lower fertility (Forejt & Iványi 1974, Storchová et al. 2004, Britton-Davidian et al. 2005, Good et al. 2008a, 2008b, Turner et al. 2011, White et al. 2011, Albrechtová et al. 2012). It is assumed that the reproductive isolation is responsible for maintaining the house mouse hybrid zone based on an equilibrium between individual dispersion and selection against hybrids with low fitness (Barton & Hewitt 1985).

Hybrid zones are a tool to study newly arising species and to understand the speciation process (Barton & Hewitt 1985). According to theoretical models (Barton 1979, Barton 1983, Barton & Gale 1993) clinal change in allele frequency across hybrid zones describes gene flow between two subspecies. The steepness of a cline is proportional to the strength of the reproductive barrier between the two subspecies and reflects the amount of gene flow. However, the patterns of introgression are supposed to vary across the genome according to the actual gene content. Loci involved in reproductive isolation are under selection in hybrid zones and as such are expected to have low or no gene flow. Gene flow on neutral loci is generally higher, though still impeded, depending on the strength of the barrier and also on it's linkage to the barrier loci. The third group – loci bearing selective advantage to the host even on an alien genetic background – may easily escape the influence of the barrier (Piálek & Barton 1997). Given the predicted differences in the way various genes introgress across hybrid zones these interfaces between the (sub)species provide a very efficient tool to identify genes putatively involved in reproductive isolation and also those which readily cross species boundaries (Rieseberg & Buerkle 2002).

Current developments in sequencing technologies made analysis of large amount of genetic characters much easier. In my PhD study my collaborators and I employ a large number of SNP markers spanning the entire mouse genome to explore

patterns of introgression across the house mouse hybrid zone in an effort to better understand the molecular basis of reproductive isolation and its evolution.

Aims of the Thesis

- To characterize genome-wide differences in introgression between X chromosome and autosomes and to elucidate their role in the evolution of reproductive isolation in the house mouse

- To identify specific genomic regions under the selection against hybrids in the house mouse hybrid zone and search for candidate genes putatively involved in the reproductive isolation in the house mouse

- Compare the introgression along the house mouse genome with genomic correlates such as the rate of recombination, the genomic differentiation and the functional composition of genes in an effort to reveal relationships between these features and the reproductive isolation in the house mouse

Materials & Methods

The house mouse hybrid zone dataset was produced over the course of several years as part of a collaborative effort between several laboratories from the USA (University of Michigan, University of Arizona) and the Czech Republic (Institute of Vertebrate Biology, AS CR, Institute of Animal Physiology and Genetics, AS CR, and Charles University in Prague). The samples were from three house mouse hybrid zone transects – the Czech-Bavarian (CZ), Bavarian-Austrian (BV) and Saxony (SX) transect – and included 895, 437 and 322 mice samples. Samples were collected throughout many years by different research groups. The details of sampling can be found in Teeter et al. (2008, 2010) and Wang et al. (2011).

The diagnostic single nucleotide polymorphisms (SNPs) used in our analysis are fully described in Wang et al. (2011). They were identified by using the beta version of the Mouse Diversity Array (Yang et al. 2009) and from genotypes for wild occurring european mouse samples provided by Yang et al. (2011). The SNP markers were considered to be diagnostic when differentially fixed between seven *M. m. musculus* and nine *M. m. domesticus* mice samples (Yang et al. 2011). The final set of ~ 1400 SNPs along with the genotype data from the house mouse hybrid zone was published by Wang et al. (2011).

To analyze the introgression across hybrid zones we used two methods of analysis for genomic clines: the likelihood (Gompert & Buerkle, 2009, 2010) and the Bayesian (Gompert & Buerkle 2011) approach. Genomic clines describe the probability of genotype/allele occurrence of a given ancestry for a locus with respect to the average genotype/allele for a given individual. The average genotype/allele is represented by the hybrid index (Gompert & Buerkle 2009).

Furthermore, we used mouse genomic data from publicly available resources: Mouse Diversity Array (Yang et al. 2009, 2011), Mouse Genetic Map (Shifman et al. 2006, Cox et al. 2009) and Gene Ontology (Ashburner et al. 2000).

Results and Discussion

During my PhD study my colleagues and I used ~ 1400 SNP markers to analyze introgression in the house mouse hybrid zone. In the first phase of our research we used data from two transects (CZ, BV) which were shown to have a sufficient number of samples in central populations (Wang et al. 2011). Our aim was to characterize the whole genome architecture of reproductive isolation between the two house mouse subspecies (*M. m. musculus*/*M. m. domesticus*). Firstly, we characterized detailed introgression on the X chromosome, making a pairwise comparison between markers. The aim of this approach was to find X-linked blocks which had the same introgression and assess introgression for individual markers within these blocks. We found considerable overlap between transects for X-linked blocks. These blocks comprise a large portion of the chromosome X and markers from within the blocks exhibited negative epistasis in both transects. Our results were in accordance with previous findings which suggested a stronger selection on the X chromosome with respect to autosomes (Tucker et al. 1992, Dod et al. 1993, Macholán et al. 2007). The largest of the blocks we found in the central portion of the chromosome X corresponded to the previous results findings using geographic clines (Payseur et al. 2004, Dufková et al. 2011). The X-linked markers were used to identify SNP markers under negative epistasis on autosomes. Only 20 – 25 % of autosomal markers exhibited negative epistasis in the two transects, however, the overlap comprised only ~ 5 %, suggesting that geographic variability might be in play. Our finding corresponds to the current speciation theory – the large X-effect – which assumes a large influence of the X chromosome in the evolution of reproductive isolation (Coyne & Orr 1986, Charlesworth et al. 1987).

Based on our previous results we selected a genomic region in the proximal part of chromosome 11, which exhibited negative epistasis in both transects, and which was previously associated with male hybrid sterility (Oka et al. 2007). Detailed analysis of this genomic region revealed gene *1700093K21Rik,* which exhibited typical features for speciation genes. The aim of this study was to show the value of combining hybrid zone data and publicly available genomic resources in searching for speciation genes.

Recently, Renaut et al. (2013) pointed out the possible importance of the functional architecture of the genome in species divergence among several species of sunflowers. In the final phase of our research we focused on genomic features putatively associated with the functional architecture of the genome in the genomic regions of low and high introgression. We found a statistically significant difference for genomic differentiation, rate of recombination and functional composition of genes between genomic regions of varying degree of introgression.

Genomic regions of low introgression tend to have higher genomic differentiation, a lower rate of recombination and a higher prevalence of genes functioning in the interior of a cell. Regulation of transcription was an important molecular function associated with these genes. On the contrary, genomic regions of high introgression tend to have a lower genomic differentiation, a higher rate of recombination and a higher prevalence of genes functioning at the periphery of a cell. The most common molecular function was signal transduction, which suggests receptor activity. Olfactoric and vomeronasal receptor genes were the most prominent group among these genes. The difference in the localization of gene activity points to the importance of functional organization of genome in the evolution of reproductive isolation and species divergence.

Conclusions

- Genome-wide analysis of the hybrid zone between two house mouse subspecies (*M. m. musculs*/*M. m. domesticus*) pointed out to the importance of the sex-chromosome X in the formation of reproductive isolation among house mouse hybrid zone transects.
- We identified 20 – 25 % autosomal SNP markers exhibiting negative epistasis. The overlap between the two transects was around 5 % which suggests geographic variability in reproductive isolation on autosomes.
- Detailed analysis of the genomic regions in the proximal part of chromosome 11 (which exhibited negative epistasis in both transects) and the use of publicly available mouse genomic resources revealed gene *1700093K21Rik* correlating with reproductive isolation between the two house mouse subspecies.
- Analysis of genomic features (genomic differentiation, rate of recombination and functional composition of genes) associated with varying degrees of introgression on autosomes suggested that the functional organization of the genome might be an important driver of species divergence and in the evolution of reproductive isolation.

Použitá literatura/References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, 25:25-29.

Albrechtová J, Albrecht T, Baird SJ, Macholán M, Rudolfsen G, Munclinger P, Tucker PK, Piálek J (2012). Sperm-related phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. *Proceedings of Biology Sciences*, 279:4803-4810.

Barton, NH (1979). Gene flow past a cline. *Heredity*, 43:333-339.

Barton NH (1983) Multilocus clines. *Evolution*, 37:454-471.

Barton NH, Hewitt GM (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, 16:113-148.

Barton NH, Gale KS (1993). Genetic analysis of hybrid zones. In: Hybrid Zones and the Evolutionary Process (ed. Harrison RG), pp. 13–45. Oxford University Press, Oxford, UK.

Boursot P, Auffray JC, Britton-Davidian J, Bonhomme F (1993). The evolution of house mice. *Annual Review of Ecology and Systematics*, 24:119-152.

Britton-Davidian J, Fel-Clair F, Lopez J, Alibert P, Boursot P (2005). Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. *Biological Journal of the Linnean Society*, 84:379-393.

Charlesworth B, Coyne JA, Barton NH (1987). The relative rates of evolution of sex-chromosomes and autosomes. *American Naturalist*, 130:113-146.

Cox A, Ackert-Bicknell CL, Dumont BL, Ding Y, Bell JT, Brockmann GA, Wergedal JE, Bult C, Paigen B, Flint J, et al. (2009). A new standard genetic map for the laboratory mouse. *Genetics*, 182:1335-1344.

Coyne JA, Orr HA (1989). Patterns of speciation in *Drosophila*. *Evolution*, 43:362-381.

Dod B, Jermiin LS, Boursot P, Chapman VH, Nielsen JT, Bonhomme F (1993). Counterselection on sex-chromosomes in the *Mus musculus* European hybrid zone. *Journal of Evolutionary Biology*, 6:529-546.

Dufková P, Macholán M, Piálek J (2011). Inference of selection and stochastic effects in the house mouse hybrid zone. *Evolution*, 65(4):993-1010.

Duvaux L, Belkhir K, Boulesteix M, Boursot P (2011). Isolation and gene flow: inferring the speciation history of European house mice. *Molecular Ecology*, 20:5248-5264.

Forejt J, Iványi P (1974). Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genetical Research*, 24:189-206.

Geraldes A, Basset P, Gibson B et al. (2008). Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Molecular Ecology*, 17:5349-5363.

Gompert Z, Buerkle CA (2009). A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, 18:1207-1224.

Gompert Z, Buerkle CA (2010). INTROGRESS: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, 10:378-384.

Gompert Z, Buerkle CA (2011). Bayesian estimation of genomic clines. *Molecular Ecology*, 20:2111-2127.

Gompert Z, Buerkle CA (2012). bgc: software for Bayesian estimation of genomic clines. *Molecular Ecology Resources*, 12:1168-1176.

Good JM, Dean MD, Nachman MW (2008a). A complex genetic basis to X- linked hybrid male sterility between two species of house mice. *Genetics*, 179:2213-2228.

Good JM, Handel MA, Nachman MW (2008b). Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution*, 62:50-65.

Hardouin EA, Orth A, Teschke M, Darvish J, Tautz D, Bonhomme F (2015). Eurasian house mouse (*Mus musculus* L.) differentiation at microsatellite loci identifies the Iranian plateau as a phylogeographic hotspot. *BMC Evolutionary Biology*, 15:26.

Macholán M, Munclinger P, Sugerková M et al. (2007). Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution*, 61:746-771.

Oka A, Aoto T, Totsuka Y, Takahashi R, Ueda M, Mita A, Sakurai-Yamatani N, Yamamoto H, Kuriki S, Takagi N, Moriwaki K, Shiroishi T (2007). Disruption of genetic interaction between two autosomal regions and the X chromosome causes reproductive isolation between mouse strains derived from different subspecies. *Genetics*, 175:185-197

Payseur BA, Krenz JG, Nachman MW (2004). Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution*, 58:2064-2078.

Piálek J, Barton NH (1997). The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics*, 45(2):493-504.

Presgraves DC. 2010. The molecular evolutionary basis of species formation. *Nature Review Genetics*, 11:175-180.

Rieseberg LH, Buerkle CA (2002). Genetic mapping in hybrid zones. *American Naturalist*, 159:S36-S50.

Shifman S, Bell JT, Copley RR, Taylor MS, Williams RW, Mott R, Flint J (2006). A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biology*, 4:e395.

Storchová R, Gregorová S, Buckiová D, Kyselová V, Divina P, Forejt J (2004). Genetic analysis of X-linked hybrid sterility in the house mouse. *Mammalian Genome*, 15:515-524.

Teeter KC, Payseur BA, Harris LW, Bakewell MA, Thibodeau LM, O'Brien JE, Krenz JG, Sans-Fuentes MA, Nachman MW, Tucker PK (2008). Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Research*, 18:67-76.

Teeter KC, Thibodeau LM, Gompert Z, Buerkle CA, Nachman MW, Tucker PK (2010). The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution*, 64:472-485.

Tucker PK, Sage RD, Warner J, Wilson AC, Eicher EM (1992). Abrupt cline for sex-chromosomes in a hybrid zone between two species of mice. *Evolution*, 46:1146-1163.

Turner LM, Schwahn DJ, Harr B (2012). Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. *Evolution*, 66:443-458.

Wang L, Luzynski K, Pool JE, Janoušek V, Dufková P, Vyskočilová MM, Teeter KC, Nachman MW, Munclinger P, Macholán M, Piálek J, Tucker PK (2011). Measures of linkage disequilibrium among neighbouring SNPs indicate asymmetries across the house mouse hybrid zone. *Molecular Ecology*, 20(14):2985-3000.

White MA, Steffy B, Wiltshire T, Payseur BA (2011). Genetic dissection of a key reproductive barrier between nascent species of house mice. *Genetics*, 189:289-304.

Yang H, Ding YM, Hutchins LN et al. (2009). A customized and versatile high-density genotyping array for the mouse. *Nature Methods*, 6:663-666.

Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, Welsh CE, Bonhomme F, Yu AH, Nachman MW, Piálek J, et al. (2011). Subspecific origin and haplotype diversity in the laboratory mouse. *Nature Genetics*, 43:648-655.

Curriculum Vitae

Václav Janoušek
Phone: (+420) 777 219 401
Email: vaclav.janousek@natur.cuni.cz, vjanousk@gmail.com

Department of Zoology
Faculty of Science
Charles University in Prague
Viničná 7, 128 00 Prague
Czech Republic

Institute of Vertebrate Biology
Academy of Sciences of the Czech Republic
Květná 8, 603 65 Brno
Czech Republic

Education:

Charles University in Prague

- Ph.D., Genetics of wild animals (2010 – expected completion in 2015)
- M.Sc., Genetics of wild animals (2008 – 2010)
- B.Sc., General Biology (2005 – 2008)

Research Experience:

Charles University in Prague (CZ), University of Michigan (USA), 2010 – 2015

- Doctoral thesis research conducted with Dr. Pavel Munclinger (Charles University in Prague) and Dr. Priscilla K. Tucker (University of Michigan)
- *Genome-wide analysis of the house mouse hybrid zone*
- Awarded with grant from the Grant Agency of the Charles University in Prague (No. 6421/2012), 2012 – 2014

Charles University in Prague (CZ), University of Arizona (USA), 2012 – present

- Project in collaboration with Dr. Robert C. Karn and Dr. Christina M. Laukaitis from the University of Arizona
- *Role of retrotransposons in gene family expansions in mammals*

Charles University in Prague (CZ), Université Montpellier II (FR), 2007 – 2010

- Master thesis project conducted with Dr. Pavel Munclinger (Charles University in Prague) and Dr. Pierre Boursot (Université Monpellier II)
- *Analysis of recently active retrotransposons in the mouse genome*

Other Work/Research Experience:

Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, 2012 – 2015

- NextGen Project: The aim of the project is to promote next-generation sequencing technologies and provide support for the NGS data analysis
- Data analyst

Institute of Archeology, Academy of Sciences of the Czech Republic, 2008 – 2009

- Project: *Analysis of Alu polymorphisms in human populations of Sahel*
- Data analyst

Honors:

Awards & Nominations:

- Nominated for the John Maynard Smith Prize given by the European Society for Evolutionary Biology (2015)

Scholarships:

- Awarded with travelling scholarship from the Mobility Fund by the Charles University in Prague (2012)

Societies:

- Society for Molecular Biology and Evolution member
- European Society for Evolutionary Biology member

Teaching Experience:

University Teaching Experience:

- Introduction to Bioinformatics (as part of graduate course Genetical Methods in Zoology provided by the Department of Zoology, Charles University in Prague), Fall 2012, Fall 2013

Other Teaching Experience:

- Natural Sciences (teaching at elementary school, Jesenice, Prague), September 2008 – June 2009

Important Lectures:

Biological Thursdays (series of lectures on the evolution at the Charles University in Prague) (November 2014)
Retrotransposons as a source of evolutionary novelties (invited lecture)
Prague, Czech Republic

NGS Workshop (October 2014)
Manipulation of the NGS data in the UNIX system – in collaboration with Libor Mořkovský (invited lecture)
Nové Hrady, Czech Republic

Mouse Biology Meeting (May 2014)
Genomic features associated with regions of high and low introgression in the house mouse hybrid zone (lecture)
Plön, Germany

Current methods on speciation research: combining population and quantitative genetic approaches (August 2012)
Case study: Genomic clines in search of speciation genes in *Mus musculus* (invited lecture)
Prague, Czech Republic

Important Workshops (participating):

Workshop on Molecular Evolution, evomics.org (February, 2015)
Český Krumlov, Czech Republic

Workshop on Genomics, evomics.org (January 2015)
Český Krumlov, Czech Republic

Workshop on Evolutionary Genetics, The Institute of Science and Technology in Austria (September 2010)
Klosterneuburg, Wien, Austria

List of Peer-Reviewed Publications

Janoušek V, Munclinger P, Wang L, Teeter KC, Tucker PK (2015). Functional organization of the genome may shape the species boundary in the house mouse. *Molecular Biology and Evolution*, 32(5):1208-1220. DOI: 10.1093/molbev/msv011. (IF_{2014} = 14.308)

Kass DH, **Janoušek V**, Wang L, Tucker PK (2014). The uncharacterized gene 1700093K21Rik and flanking regions are correlated with reproductive isolation in the house mouse, *Mus musculus*. *Mammalian Genome*, 25(5-6):223-234. DOI: 10.1007/s00335-014-9506-2 (IF_{2014} = 2.883)

Janoušek V, Karn RC, Laukaitis CM (2013). The role of retrotransposons in gene family expansions: insights from the mouse Abp gene family. *BMC Evolutionary Biology*, 13:107. DOI:10.1186/1471-2148-13-107 (IF_{2014} = 3.41)

Janoušek V, Wang L, Luzynski K, Dufková P, Vyskočilová MM, Nachman MW, Munclinger P, Macholán M, Piálek J, Tucker PK (2012). Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. *Molecular Ecology*, 21(12):3032-3047. DOI: 10.1111/j.1365-294X.2012.05583.x. (IF_{2014} = 5.84)

Wang L, Luzynski K, Pool J, **Janoušek V**, Dufková P, Vyskočilová M, Teeter KC, Nachman MW, Munclinger P, Macholán M, Piálek J, Tucker PK (2011). Measures of linkage disequilibrium among neighboring SNPs indicate asymmetries across the house mouse hybrid zone. *Molecular Ecology*, 20(14):2985-3000. DOI:10.1111/j.1365-294X.2011.05148.x (IF_{2014} = 5.84)