## Paper outline

**Background and questions:**

In many closely related bird species, higher interspecific divergence and lower intraspecific diversity is observed on the Z chromosome compared to autosomes. This can be explained by (i) lower gene flow on the Z chromosome, (ii) lower effective population size of the Z chromosome (and thus stronger effect of genetic drift) and/or (iii) stronger positive selection acting on the Z chromosome.

Previously, we brought evidence on lower levels of gene flow on the Z chromosome compared to autosomes, suggesting the important role of the Z chromosome in reproductive isolation between the nightingales (Storchová et al. 2010). This study was, however, based on limited number of loci (4 Z-linked, 8 autosomal).

Data from other bird species (6 galloanserae species) (Wright et al. 2015) suggests that selection is less effective on the Z chromosome (MK test) and that faster Z evolution in birds is primarily due to genetic drift.

Here we will perform detailed analysis of genetic polymorphism and divergence along the Z chromosome and autosomes in two nightingale species to elucidate selective forces responsible for higher genetic divergence of the Z chromosome compared to autosomes and to identify genomic regions responsible for reproductive isolation.

**Questions and hypotheses:**

1. What are the levels of genetic divergence and gene flow along the Z chromosome.   
2. Is there any evidence for positive selection on the Z chromosome of nightingales?

3. Is there any functional enrichment among genes lying in island of high differentiation?

**Data, analyses and results:**

1. Analysis of gene flow along the Z chromosome using intronic sequences from 12 Z-linked loci

Data: Intronic sequences of 12 loci evenly scattered along the whole Z chromosome from cca 20 allopatric individuals of both nightingale species and 1 individual of bluethroat (outgroup).

Using IM analysis, we estimated locus specific migration rates for 12 loci. We found evidence for low levels of gene flow in 4 out of 12 Z-linked loci (*PPWD1*, *TG853*, *TG1015* a *VLDLR7*). This gene flow was strongly asymmetric (more intense from Thrush Nightingale to the Common Nightingale). Interestingly, our previous data (based on analysis of 8 loci) suggested that on autosomes gene flow was more intense in the opposite direction. Why should be the patterns of gene flow opposite on the Z chromosome and autosomes?

Using Tajima’s D and Fu and Li’s D tests, we found evidence of positive selection at one Z-linked locus (*TG401)* in the Thrush Nightingale and one Z-linked locus (*TG1925)* in the Common Nightingale. No evidence for positive selection was found in HKA test (in neither of both species, even when only 8 loci without gene flow were included into the analysis). Nevertheless, *TG401* and *TG1925* showed the highest deviations from expected polymorphism and divergence values.

Absence of gene flow along the large part of the Z chromosome suggests that there could be multiple speciation genes along the Z chromosome. Alternatively, one or a few speciation loci could cause low levels of gene flow along the majority of the Z chromosome length if the recombination between the species is suppressed on this chromosome, for example by the presence of chromosomal inversion(s). To test this hypothesis, we genotyped about 200 sympatric individuals for 5 species specific SNP markers lying in different highly divergent regions of the Z chromosome and searched for recombinant genotypes. Only three recombinant genotypes were found among approximately 200 individuals. In all three cases, there was an introgression of a single always different SNP marker. Thus although the recombination on the Z chromosome seems to be rare, it’s occurrence in sympatric population together with observed interspecific gene flow in 3 Z-linked loci supports the scenario of multiple independent speciation loci scattered along the Z chromosome.

2. Analysis of genetic divergence along the Z chromosome and autosomes using 454 transcriptome data

Data: 454 sequences of liver transcriptome of 8 allopatric individuals of the Common Nightingale, 7 individuals of the Thrush Nightingale and 1 individual of Bluethroat (outgroup).

Show levels of differentiation along the Z chromosome and autosomes.

Identification of islands of differentiation and GO analysis of genes within the islands.

* mean, max Fst per contig -> histogram, xy plot (mean x max)
* windowed Fst in zebra finch coordinates (+ xy plot n vars vs aggregate)
  + use all snps in bucktes per contig (min-max exon in contig)
* GO for top contigs
* GO for genes in whole islands

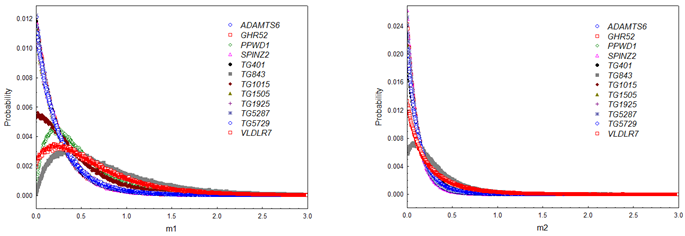
3. McDonald-Kreitman test using 454 transcriptome data (separate publication?).

Is positive selection more frequent on the Z chromosome compared to autosomes.

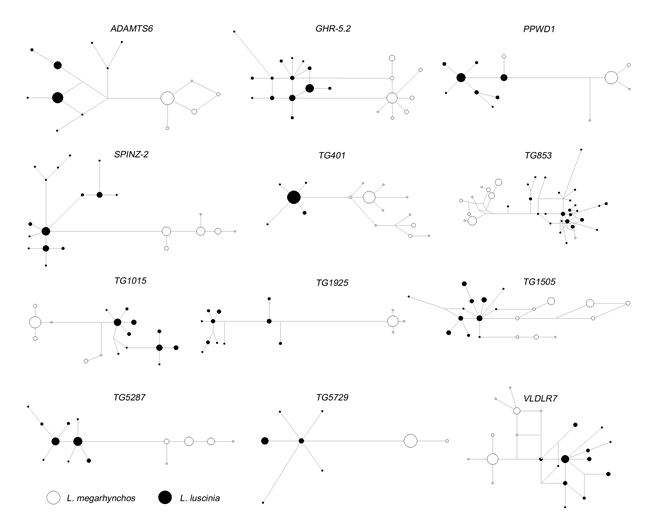
* MK test?

**Conclusions:**

**Figure 1** Posterior probability distribution of locus specific migration rates for 12 Z-linked loci. Migration from the Thrush Nightingale into the Common Nightingale (m1), migration from the Common Nightingale into the Thrush Nightingale (m2).



**Figure 2.** Genealogical networks of 12 Z-linked genes.



**Table 1.** Tajima’s and Fu and Li’s D tests. Significant (p < 0.05) values are in bold and marked with \*.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Lokus** | **Druh1** | **Tajima's D** | | **Fu and Li's D** | |
| ADAMTS6 | SO | -0,43180 |  | 0,89169 |  |
|  | ST | -0,46257 |  | -0,12666 |  |
| GHR-5.2 | SO | -0,86293 |  | 0,58682 |  |
|  | ST | -1,11683 |  | -1,55077 |  |
| PPWD1 | SO | -1,54526 |  | -0,38571 |  |
|  | ST | -0,51646 |  | -0,06560 |  |
| SPINZ-2 | SO | 0,43920 |  | -0,54412 |  |
|  | ST | -1,01856 |  | -1,41373 |  |
| TG401 | SO | -1,09664 |  | -2,30551 |  |
|  | ST | **-1,81947** | **\*** | **-2,69383** | **\*** |
| TG853 | SO | 0,02873 |  | -2,01498 |  |
|  | ST | -1,50162 |  | -1,80568 |  |
| TG1015 | SO | -0,63331 |  | 0,81290 |  |
|  | ST | 0,18939 |  | -0,09269 |  |
| TG1505 | SO | 1,67716 |  | 0,94766 |  |
|  | ST | -1,06522 |  | -1,37685 |  |
| TG1925 | SO | -1,73075 |  | **-2,71054** | **\*** |
|  | ST | -0,00454 |  | 0,21443 |  |
| TG5287 | SO | -0,19239 |  | -1,07927 |  |
|  | ST | -1,15899 |  | -1,47069 |  |
| TG5729 | SO | -0,83603 |  | 0,56369 |  |
|  | ST | -1,26506 |  | -2,35568 |  |
| VLDRL7 | SO | -0,56542 |  | -1,23423 |  |
|  | ST | -0,90388 |  | -0,59055 |  |