

Diss. ETH N° 20391

***Silene latifolia* (Caryophyllaceae)**
sex chromosome evolution

A dissertation submitted to

ETH ZURICH

for the degree of

DOCTOR OF SCIENCES

presented by

Nicolas Blavet

Master of Science in

Mathematical and Informatical Analysis of Life

born March 6th, 1983

citizen of France

accepted on the recommendation of

Prof. Dr. Alex Widmer, examiner

Prof. Dr. Paul Schmid-Hempel, co-examiner

Dr. Gabriel Marais, co-examiner

2012

Summary

Sex chromosomes evolved more than a hundred million years ago in mammals and birds but much more recently in some flowering plant lineages. Studies on the evolution of sex chromosomes are important to understand sex determination mechanisms and are of general evolutionary interest because sex chromosomes have evolved independently numerous times in different lineages of animals, fungi and plants. Flowering plants are particularly suited for investigating sex chromosome evolution because sex chromosomes emerged relatively recently in different groups and because of the presence of closely related species lacking sex chromosomes. In this thesis, I contribute to the understanding of sex chromosome evolution by analyzing the recently emerged sex chromosomes of the white campion, *Silene latifolia*.

First, I investigated transcriptomic data coming from eight individuals representing five closely related Caryophyllaceae species, including four *Silene* and one *Dianthus* species (Chapter I). I found about 74000 genes in the studied species with about 1400 genes specific to the *Silene* genus. Moreover the detection of thousands of single nucleotide polymorphisms (SNPs) provides new molecular resources for linkage mapping and population genetic analyses. This study showed the value of comparative transcriptome analyses based on next generation sequencing data for the characterization of genetic variation in non-model species.

In a second step, I analyzed a large piece of genomic DNA located in a pseudoautosomal region (PAR) of plant sex chromosomes (Chapter II). Comparison of *S. latifolia* and *S. vulgaris* homologous bacterial artificial chromosome (BAC) sequences identified new pseudoautosomal genes in *S. latifolia*. These genes are conserved in size and linear arrangement in both species, which indicates small-scale gene collinearity between the PAR region of sex chromosomes in *S. latifolia* and the

corresponding autosomal region in *S. vulgaris*. Contrary to expectations, I found no increase in GC and GC3 content of the pseudoautosomal genes, in contrast to the situation in mammals, but I found evidence for a moderate size increase of the *S. latifolia* PAR compared to the *S. vulgaris* autosomes. This increase is similar to the size difference observed between *S. latifolia* and *S. vulgaris* autosomes and can not explain the specific size increase seen in the *S. latifolia* sex chromosomes.

In a third study, I focused on the non-recombining part of the *S. latifolia* sex chromosomes (Chapter III). Similarly to Chapter II, I compared BAC sequences from *S. latifolia* and *S. vulgaris*, identified new sex-linked genes and found conserved small-scale collinearity between both *S. latifolia* X and Y chromosomes and the corresponding *S. vulgaris* autosomal region for genes located in the oldest evolutionary stratum. The absence of pseudogenes on the Y chromosome BACs and the accumulation of transposable elements were notable as they indicate that *S. latifolia* is rarely losing genes but is accumulating transposable elements on the Y chromosome. An analysis of the distribution of genes with reduced expression of the allele on the Y chromosome further revealed that Y gene degeneration is a random process in *S. latifolia*.

This thesis provides new insights to the understanding of the early stages of sex chromosome evolution in plants and shows the value of next generation sequencing technologies for the study of genomic and transcriptomic variation in formerly largely intractable non-model organisms.

Résumé

L'évolution des chromosomes sexuels a débuté il y a des centaines de millions d'années chez les mammifères et les oiseaux mais est apparue plus récemment dans certains taxons de plantes à fleurs. L'analyse de l'évolution des chromosomes sexuels est importante pour la compréhension des mécanismes de la détermination sexuelle et est d'intérêt général pour l'évolution, car les chromosomes sexuels ont évolué indépendamment à plusieurs reprises dans des lignées différentes d'animaux, de champignons et de plantes. Les plantes à fleurs sont particulièrement intéressantes pour la recherche sur l'évolution de ces chromosomes du fait de la récente émergence de ceux-ci et de la présence d'espèces proches étant dépourvues de chromosome sexuel. J'ai contribué par cette thèse à la compréhension de l'évolution des chromosomes sexuels, en analysant ceux, récemment apparus, du compagnon blanc *Silene latifolia*.

Premièrement, j'ai étudié les données transcriptomique provenant de huit individus représentant cinq Caryophyllacées phylogénétiquement proches dont quatre espèces de *Silene* et une espèce de *Dianthus* (Chapitre I). Nous avons trouvé environ 74000 gènes dans les espèces étudiées et environ 1400 gènes spécifiques du genre *Silene*. De plus, la détection de milliers de site de polymorphisme nucléotidique va apporter de nouvelles ressources moléculaires pour la cartographie génétique et les analyses de génétique des populations. J'ai également mis en évidence l'intérêt des analyses comparatives de transcriptomes d'espèces non modèles à grande échelle en utilisant les techniques de séquençage de nouvelle génération.

En deuxième étape, j'ai analysé un large fragment d'ADN génomique provenant de la région pseudoautosomale de chromosome sexuel d'une plante (Chapitre II). La comparaison de séquences homologues de chromosomes artificiels de bactéries (BAC) venant de *S. latifolia* et *S. vulgaris* a révélé de nouveaux gènes

pseudoautosomaux dont la taille et l'arrangement linéaire est conservé entre les deux espèces, ce qui indique une colinéarité à petite échelle entre le chromosome X et la région autosomal correspondante. Contrairement aux prévisions, je n'ai trouvé aucune augmentation des taux de GC et GC3 sur les gènes pseudoautosomaux de *S. latifolia*, à la différence des mammifères, mais j'ai mis en évidence une augmentation modérée de la taille de la région pseudoautosomal de *S. latifolia* comparé à la séquence autosomal de *S. vulgaris*. Cette augmentation est similaire à la différence existant entre les autosomes des deux espèces et ne peut pas expliquer l'augmentation spécifique de la taille des chromosomes sexuels de *S. latifolia*.

Dans la troisième étude, je me suis concentré sur la partie non recombinante des chromosomes sexuels de *S. latifolia* (Chapitre III). De même qu'au Chapitre II, nous avons comparé des séquences de BAC provenant de *S. latifolia* et *S. vulgaris*. J'ai identifié de nouveaux gènes liés aux sexes et j'ai trouvé une colinéarité entre des gènes situés dans la plus vieille strate évolutive des chromosomes X et Y de *S. latifolia* et de l'autosome correspondant de *S. vulgaris*. L'absence de pseudogène et l'accumulation d'éléments transposables, m'ont permis d'indiquer que les gènes du chromosome Y de *S. latifolia* sont rarement perdus et que ce dernier accumule des éléments transposables. Une analyse portant sur l'expression réduite des allèles du chromosome Y ont révélé que la dégénération des gènes Y de *S. latifolia* est un processus aléatoire.

Cette thèse apporte de nouveaux éléments pour la compréhension des premiers stades de l'évolution des chromosomes sexuels de plantes et montre l'importance des techniques de séquençage de nouvelle génération pour l'étude des variations génomiques et transcriptomiques d'organismes non modèles.