

## Mode of reproduction in *Arabidopsis suecica*

TORBJÖRN SÄLL<sup>1</sup>, CHRISTINA LIND-HALLDÉN<sup>2</sup>, MATTIAS JAKOBSSON<sup>1</sup> and  
CHRISTER HALLDÉN<sup>3</sup>

<sup>1</sup>Department of Cell and Organism Biology, Genetics, Lund University, Lund, Sweden

<sup>2</sup>Department of Mathematics and Natural Sciences, Kristianstad University, Kristianstad, Sweden

<sup>3</sup>Department of Clinical Chemistry, Malmö University Hospital, Malmö, Sweden

Säll, T., Lind-Halldén, C., Jakobsson, M. and Halldén, C. 2004. Mode of reproduction in *Arabidopsis suecica*. — *Hereditas* 141: 313–317. Lund, Sweden. ISSN 0018-0661. Received December 19, 2003. Accepted November 19, 2004

The breeding system of *Arabidopsis suecica* was investigated through genetic analysis of microsatellite segregation patterns in five controlled crosses as well as in 16 single-mother families collected in the wild. Analysis of single and two-locus segregations in the F<sub>2</sub> generation following a cross clearly shows that *A. suecica* reproduces sexually. The single-mother families show a high level of homozygosity corroborating earlier results indicating a high level of inbreeding. The high level of individual homozygosity is due both to a high level of selfing and to the underlying population structure.

Torbjörn Säll, Genetics, Department of Cell and Organism Biology, Lund University, Sölvegatan 29, SE-223 62 Lund, Sweden.  
E-mail: torbjorn.sall@cob.lu.se

*Arabidopsis suecica* (2n = 26) is an allopolyploid with *A. thaliana*, (2n = 10) and *A. (Cardaminopsis) arenosa* (2n = 16, 32) as its parental species (HYLANDER 1957; MUMMENHOF and HURKA 1994; PRICE et al. 1994; O'KANE et al. 1996). It has recently attracted interest since one of its parents, *A. thaliana*, is the leading model organism in plant biology (CHEN et al. 1998, 2004; COMAI et al. 2000). Thus *A. suecica* is an obvious candidate model species for polyploidy research. It shares several features with its parent *A. thaliana*: it has a short generation time and it can be kept at a limited size. In addition, *A. suecica* has a limited overall genome size of approximately 280 Mb (unpubl. results) and has a single origin less than 60 000 years ago, with respect to both the chloroplast (SÄLL et al. 2003) and the nucleus (unpubl.).

Whether *A. suecica* is used for evolutionary or functional investigations, it is of considerable interest to extend our knowledge of its reproductive system and population structure. The parents of *A. suecica* have widely different breeding systems. *A. thaliana* is a selfer with all the 'classical' correlates that normally accompany this breeding system; high individual homozygosity, a relatively high level of linkage disequilibrium, a low within-population variation and a relatively high between-population differentiation (BERGELSON et al. 1997, HAGENBLAD and NORDBORG 2002). On the other hand *A. arenosa* is, according to our observations in the greenhouse, self-incompatible and thus a strict outbreeder. Active pollination between individuals is, in our experience, necessary to achieve seed formation. In addition, *A. arenosa* produces a distinct scent and its flowers are clearly visible also at a distance. Both these

features are typical of insect pollinated cross-breeders (RICHARDS 1986). *A. thaliana* and *A. suecica* differ from *A. arenosa* in both these traits: neither has a scent and the flowers are much less conspicuous.

*A. suecica* has not so far been investigated with respect to its breeding system. Our experience shows that it is fully self-fertile when grown in isolation (unpubl. results). A limited investigation of variation and population structure using RAPDs (LIND-HALLDÉN et al. 2002) indicated a low level of variation in *A. suecica*. This is compatible with its single origin and short time of existence. The limited variation was mostly distributed between populations which is typical for effective inbreeders. However, our observations so far do not definitely exclude apomixis which also tends to show low within- and relatively large between-population variation (LOVELESS and HAMRICK 1984). The a priori probability of a single species in the family Brassicaceae being apomictic may be low; still, apomixis occurs in *Arabis holeboellii* (BÖCHER 1951; SHARBEL and MITCHELL-OLDS 2001), and for future interpretation of the evolution of *A. suecica* the difference between any sexual system and apomixis is crucial. In this study we investigate controlled crosses as well as single mother families in order to establish the mode of reproduction and the population structure of *A. suecica*.

### MATERIAL AND METHODS

#### *Plant material and experimental design*

*Crosses.* — We investigated the segregation ratios in five different crosses. Each cross was made between a

single maternal and a single paternal plant. The maternal parents in the crosses were offspring from a single plant, S361. The paternal parents were all offspring of two single plants S300/10 (four parents) and S290/11 (one parent). The crosses were the following S361:1 × S300/10:1, S361:2 × S300/10:2, S361:3 × S300/10:3, S361:4 × S300/10:4 and S361:5 × S290/11:1. The F<sub>1</sub> parents were grown simultaneously in a greenhouse and were isolated before flowering. The F<sub>2</sub> seeds were harvested at maturity and later sown in the greenhouse. A total of 15 offspring were grown and harvested per cross. The origins of the plants are shown in Table 1.

*Single-mother families.* — Seeds from 16 families were sown in the glass house. Each of these families originated from a single field collected maternal plant which had been naturally pollinated in the wild prior to collection. A total of seven offspring were grown and harvested per family. The origins of the maternal plants are shown in Table 1.

#### Microsatellite analysis

DNA was isolated from young greenhouse grown plants using a Plant DNeasy kit from Qiagen according to the manufacturer's instructions. DNA integrity was checked by agarose gel electrophoresis and DNA concentration was determined using fluorometry with Pico Green (Molecular Probes). Microsatellite loci were identified in the *A. thaliana* genome using Tandem Repeat Finder (BENSON 1999) and primers designed with Oligo (Molecular Biology Insights). Primer sequences were subsequently analysed for specificity using BLAST (GenBank). PCR-assays were optimized with regard to annealing temperature and primer concentrations. PCR reactions contained 5

ng template DNA, 2 mM MgCl<sub>2</sub>, 100 μM dNTP (Amersham/Pharmacia), 0.75 units AmpliTaq Gold polymerase (Applied Biosystems), 10 mM Tris-HCl, pH 8.3, 50 nM KCl and primer concentrations as given in Table 2 in a final volume of 25 μl. The DNA fragments were amplified on PTC-100 thermocyclers (MJ Research) using an initial denaturation step at 95°C for 9 min, followed by 30 cycles of 96°C for 1 min, the annealing temperature given in Table 2 for 1 min and 72°C for 2 min. A final extension step of 72°C for 10 min was then made. PCR products were resolved using capillary electrophoresis run on ABI 3100 or 3730 sequencers employing GeneMapper v. 3.0 software (Applied Biosystems). The allele sizes of the microsatellite markers were determined in relation to a size marker, GeneScan-500 LIZ (Applied Biosystems, USA).

## RESULTS

### Crosses

A screen for polymorphisms among the parents showed that these differed at five loci. All paternal plants showed the same allele for all loci. Since all maternal plants also showed the same allele for all loci, the results in the analyses below could be combined from the five crosses. A total of 73 F<sub>2</sub> offspring were typed. Due to missing values the number of offspring typed for the different loci varied between 61 and 73. The distributions of genotypes at the different loci are shown in Table 3. In no case are the segregation ratios significantly different from 1:2:1, which is the expected distribution for a single co-dominant locus under sexual reproduction. The two-locus combinations were subsequently tested for independent segregation. No locus-pair show any indication of dependence: the P-values varied between 0.09 and 0.90. This is expected if the loci are unlinked and reproduction is strictly sexual.

### Single-mother families

A total of 16 families, each representing a different population, were typed for nine microsatellites that were shown to be variable in *A. suecica*. Seven offspring were typed in each family. A total of 19 missing values occurred leaving a total of 989 offspring locus combinations. The number of detected alleles and the levels of heterozygosity per locus are presented in Table 4. As seen, the level of variation at the selected nine loci was quite high. The estimates of the total level of expected heterozygosity varied between 0.40 and 0.79, with an average of 0.53. In contrast, the overall level of individual heterozygosity was very low. Of the 144 (9 × 16) locus-family

Table 1. Description of the 19 *A. suecica* accessions used in this study.

Name	Location	Name	Location
S61	Vännäs (SW)	S241	Gålsjö (SW)
S71	Söder Nyåker (SW)	S290 <sup>a</sup>	Ede (SW) <sup>b</sup>
S81	Nordmaling (SW)	S300 <sup>a,c</sup>	Sörfjärda (SW) <sup>b</sup>
S92	Västanbäck (SW)	S340	Kotka (FI) <sup>d</sup>
S133	Strömsbruk (SW)	S351	Iisalmi (FI) <sup>d</sup>
S163	Ytterhogdal (SW)	S361 <sup>a</sup>	Hanko (FI) <sup>d</sup>
S171	Los (SW)	S420	Ramsberg (SW)
S182	Våxnan (SW)	S520	Artjärvi (FI) <sup>c</sup>
S231	Olofsfors (SW)	S700	Ulricehamn (SW)

<sup>a</sup>Accessions used in crosses. <sup>b</sup>Svante Holm, Mitthögskolan, <sup>c</sup>S301 is a different accession collected at the same location as S300. <sup>d</sup>Outi Savolainen, Oulo University, <sup>e</sup>Arrto Kurrto, Helsinki University. SW = Sweden, FI = Finland. When not otherwise noted the accessions have been collected by the authors.

Table 2. Primer sequences used to amplify the analyzed sequences. The 3'-positions of the primer sequences in the genome of *A. thaliana* (TIGR 4.0 version of the *Arabidopsis* genome) and the annealing temperatures used for the PCR amplification are shown.

Primer name	Primer sequence	Annealing temperature (°C)	Primer concentration (F/R)	Chromosome	3'-position
1F	5'-(HEX)TGAAGGAGCCGATCAG-3'	56	400/400	1	3224440
1R	5'-(HEX)TACATTCGGTTTGGTTGATAA-3'	56	400/400	1	3224632
8F	5'-(HEX)CTAGAAACCGAAATTGTTAC-3'	59	400/400	1	22612217
8R	5'-(HEX)CCCAATTTACAGTATCGTATT-3'	59	400/400	1	22612061
12F	5'-(HEX)CTTCTCTCTCCGACTTCC-3'	59	50/50	2	8685448
12R	5'-(HEX)CAAGAACTCAAACCGTTTAT-3'	59	50/50	2	8685657
21F	5'-(HEX)CTCTTCGTTTCCCATGAG-3	59	50/50	2	18321697
21R	5'-(HEX)CTCCACAAAATATGACCTTA-3'	59	50/50	2	18322005
34F	5'-(HEX)ATAGAGCCTATAATGAGCGT-3'	59	400/400	3	19302301
34R	5'-(HEX)AAAATAAATTGTAACCCTTTG-3'	59	400/400	3	19302548
36F	5'-(HEX)GCAGCAGAAAGACAGTCTC-3'	59	50/300	4	21377070
36R	5'-(HEX)AGCCTAGAGTTTGTGGTACAC-3'	59	50/300	4	21376923
40F	5'-(HEX)JGTCAAAGACCATAACAATGA-3'	59	300/50	4	748695
40R	5'-(HEX)AATTTCACTGATTCATCGCTA-3'	59	300/50	4	748844
52F	5'-(HEX)CAATATGTCATTTTGTCAACC-3'	56	400/400	5	12273386
52R	5'-(HEX)ACCAGTGGCCTAAGTAGATTA-3'	56	400/400	5	12273528
59F	5'-(HEX)AGAAGTTTGTAAATGTCGTT-3'	56	50/50	5	21424815
59R	5'-(HEX)AAATCAGATTTATTGGT CATA-3'	56	50/50	5	21425023

combinations 136 were completely uniform and only eight were variable. Four of these were classified as originating from a heterozygous mother. Thus the heterozygosity per locus among mothers was 0.028 (4/144). The remaining four variable families appeared to originate from outcrossing of a homozygous mother plant. These were found in three individual offspring, two in family S182 and one in S520, which was heterozygous for two loci. Thus, the relative frequency of detected outcrossing was 0.027 (3/112). In all there were 10 cases where an individual offspring was heterozygous at a locus which yields an average heterozygosity of 0.01 per locus among the individual offspring. Thus the estimate of  $H_O$ , the observed proportion of heterozygosity, was 0.028 among the maternal plants and 0.010 among the offspring. Given that no selection or drift occurs, these estimates should be the same. The observed discrepancy is due to the fact that the families where the mother was

heterozygous showed fewer heterozygote offspring than expected.

The contrast between the overall level of expected heterozygosity,  $H_T$ , and the proportion of heterozygous individuals,  $H_O$ , may be due to a high level of selfing or a population structure with a high degree of isolation between subpopulations, or a combination of these. The classical parameterization of these entities are the  $F$  statistics introduced by WRIGHT (1951) and extensions of these NEI (1973), where the overall deviation from random mating  $F_{IT}$  is partitioned into one intra- and one inter-population component,  $F_{IS}$  and  $F_{ST}$  respectively. If the individual increase in homozygosity within subpopulations is due to selfing

Table 3. Single locus segregation at the five loci.  $M$  = maternal allele,  $P$  = paternal allele.

Locus	Genotype			$\chi^2$	p
	MM	MP	PP		
1	15	31	24	3.22	0.07
8	19	36	18	0.04	0.84
34	23	32	15	2.34	0.13
52	20	36	12	2.12	0.14
59	21	24	16	3.20	0.07

Table 4. Level of variation in the 16 single-mother families.  $N$  is the number of alleles,  $H_T$  is the overall heterozygosity,  $H_{OO}$  is the proportion of heterozygous offspring and  $H_{OM}$  is the proportion of heterozygous mothers.

Locus	N	$H_T$	$H_{OO}$	$H_{OM}$
1	2	0.43	0	0
8	2	0.46	0.02	1/16
12	4	0.54	0	0
21	3	0.40	0	0
34	3	0.45	0.01	0
36	5	0.79	0.04	1/16
40	3	0.54	0.01	1/16
52	4	0.66	0.01	0
59	3	0.53	0.01	1/16
Average	3.2	0.53	0.028	0.028

only, the relation between  $F_{IS}$  and the selfing rate,  $S$ , is  $F_{IS} = S/(2 - S)$ , (HEDRICK 1999). This leads to the relation  $S = (H_T - H_O - H_T \times F_{ST}) / (H_T - H_O/2 - H_T \times F_{ST})$ . The present data indicate that  $H_O$  is in the range 0.01–0.03. The average of  $H_T$  for the different loci is 0.53 as stated above. Due to the fact that we have only sampled one family per population it is not possible to obtain a good estimate of  $F_{ST}$  directly from data. Instead we have calculated  $S$  for a range of  $F_{ST}$  values, given  $H_O = 0.01$  and  $H_O = 0.03$ . The results show that the level of actual selfing is high, over a broad range of  $F_{ST}$  (Table 5).

## DISCUSSION

The major aim of this study was to investigate the breeding system of *A. suecica*. Specifically we wanted to determine whether *A. suecica* is apomictic or sexual. The results show clearly that *A. suecica* is sexual. This is shown by the fact that crosses can be readily made and that the single and two locus segregations in the  $F_2$  generation are fully compatible with sexual reproduction. In addition, the offspring in the single-mother families are predominantly homozygous which corroborate this conclusion since apomicts are generally highly heterozygous (HUGHES and RICHARDS 1988; GORNALL 1999). Thus *A. suecica* reproduces sexually with a high level of individual homozygosity, also for variable loci. The analysis of the pattern of variation among the single-mother families shows that it is most likely that selfing occurs at a high level. In LIND-HALLDÉN et al. (2002) it was estimated that some 80% of the variation in *A. suecica* was distributed between populations. This is a very high level also for known selfers. HAMRICK and GODT (1990) found an average proportion of between-population variation of 51% when compiling data

from a number of allozyme studies. The high level of individual homozygosity in *A. suecica* is thus clearly due to both a high level of selfing as well as an additional effect of population structure with a high degree of between-population variation.

The observation that *A. suecica* is self-compatible means that the self-incompatibility of *A. arenosa* has broken down in the formation of the allopolyploid while it survived the formation of the autopolyploid form of *A. arenosa*. The exact nature of the self-incompatibility (SI) system of *A. arenosa* has not been determined but it can safely be assumed to function as a multiallelic sporophytic system, which is the only SI-system found in Brassicaceae (RICHARDS 1986). More specifically, *A. lyrata*, which is closely related to *A. arenosa* (YANG et al. 1999), is known to possess this kind of SI-system (MABLE et al. 2003). *A. arenosa* is tetraploid in the majority of its range (MESICEK 1970) and all our observations are based on tetraploid accessions. Since the SI system is functional in *A. arenosa*, polyploidy per se does not impair its function as may be the case with gametophytic systems. Thus, the  $S$  locus of *A. suecica* was probably silenced through mutations.

The breeding system has a profound effect on the evolution of a species. Polyploids like *A. suecica* are believed to gradually approach a diploid genome organisation (OHNO 1970; SOLTIS and SOLTIS 1993; SOLTIS and SOLTIS 1995; WENDEL 2000). This process occurs through the fixation of inactive alleles at one of the two homeologous loci for a specific gene, transforming that locus into a pseudogene. Non-functional alleles are normally recessive to active wild-type alleles (WRIGHT 1968). The rate of fixation of a recessive allele under selection differs between asexual and sexual species also when a single locus is considered. Sexually reproducing species will, through segregation, produce homozygotes for the recessive allele which, can only be formed by recurrent mutation in asexuals. In particular, among sexuals, inbreeding further enhances the rate of selection for recessive alleles (HEDRICK 1999).

Another aspect of genome evolution is the process of building up neutral or near-neutral variation. The rate of increase of neutral variation does not differ between sexual and asexual species with respect to single loci but when multiple loci are considered they differ. When new mutations occur they will always be in linkage disequilibrium with all variable loci. Asexuals retain these disequilibria, the level of which is unrelated to the positions of the loci. In sexual species the disequilibrium is broken down according to well known patterns, however more slowly in inbreeders than in outbreeders. In our opinion, it is of greater

Table 5. Level of selfing,  $S$ , calculated for different levels of  $F_{st}$ . For details, see text.

$F_{st}$	$H_O$	
	0.01	0.03
0	0.99	0.97
0.1	0.99	0.97
0.2	0.99	0.96
0.3	0.98	0.96
0.4	0.98	0.95
0.5	0.98	0.94
0.6	0.98	0.92
0.7	0.97	0.90
0.8	0.95	0.84
0.9	0.80	0.60
1.0	–	–

interest to study this process under sexual reproduction than in asexuals. Thus *A. suecica* is potentially very well suited for investigations of how variation is built up in the genome.

*Acknowledgements* – We thank S. Holm, A. Kurrto and O. Savolainen for help with material. We thank Prof. Bengt Olle Bengtsson, Dr. Alf Ceplitis and Margaret Sykes for comments on the manuscript. This work was supported by the Crafoord foundation, the Erik-Philip Sörensen foundation, the Tryggers foundation and Magnus Bergvalls foundation.

## REFERENCES

- Benson, G. 1999. Tandem repeat finder: a program to analyse DNA sequence. – *Nucleic Acids Res.* 27: 573–580.
- Bergelson, J., Stahl, E., Dudek, S. et al. 1997. Genetic variation within and among populations of *Arabidopsis thaliana*. – *Genetics* 148: 1311–1323.
- Böcher, T. W. 1951. Cytological and embryological studies in the amphipomictic *Arabis holboellii* complex. – *Kong. Danske Vidensk. Selsk.* 6: 1–59.
- Chen, J. Z., Comai, L. and Pikaard, C. S. 1998. Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in *Arabidopsis* allotetraploids. – *Proc. Natl Acad. Sci. USA* 95: 14891–14896.
- Chen, J. Z., Wang, J., Tian, L. et al. 2004. The development of *Arabidopsis* model system for genome – wide analysis of polyploidy effects. – *Biol. J. Linn. Soc.* 82: 689–700.
- Comai, L., Tyagi, A. P., Winter, K. et al. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. – *Plant Cell* 12: 1551–1567.
- Gornall, R. J. 1999. Population genetic structure in agamosperous plants. – In: Hollingworths, P. M., Bateman, R. M. and Gornall, R. J. (eds), *Molecular systematics and plant evolution*. Taylor and Francis, p. 118–138.
- Hagenblad, J. and Nordborg, M. 2002. Sequence variation and haplotype structure surrounding the flowering time locus FRI in *Arabidopsis thaliana*. – *Genetics* 161: 289–298.
- Hamrick, J. L. and Godt, M. J. W. 1990. Allozyme diversity in plant species. – In: Brown, H. D., Clegg, T. M., Kahler, A. L. et al. (eds), *Plant population genetics, breeding and genetic resources*. Sinauer.
- Hedrick, P. W. 1999. *Genetics of populations*, 2nd ed. – Jones and Bartlett, Sudbury Mass.
- Hughes, J. and Richards, A. J. 1988. The genetic structure of populations of sexual and asexual *Taraxacum* (dandelions). – *Heredity* 60: 161–171.
- Hylander, N. 1957. *Cardaminopsis suecica* (Fr.) Hiit., a northern amphidiploid species. – *Bull. Jardin Bot. Bruxelles.* 27: 591–604.
- Lind-Halldén, C., Säll, T. and Halldén, C. 2002. Genetic variation in *Arabidopsis suecica* and its parental species *A. arenosa* and *A. thaliana*. – *Hereditas* 136: 45–50.
- Loveless, M. D. and Hamrick, J. L. 1984. Ecological determinants of genetic structure in plant populations. – *Annu. Rev. Ecol. Syst.* 15: 65–95.
- Mable, B. K., Schierup, M. H. and Charlesworth, D. 2003. Estimating the number, frequency, and dominance of S-alleles in a natural population of *Arabidopsis lyrata* (Brassicaceae) with sporophytic control of self-incompatibility. – *Heredity* 90: 422–431.
- Mesicek, J. 1970. Chromosome counts in *Cardaminopsis arenosa* Agg. (Cruciferae). – *Preslia* 42: 225–248.
- Mummenhof, K. and Hurka, H. 1994. Subunit polypeptide composition of Rubisco and the origin of allopolyploid *Arabidopsis suecica* (Brassicaceae). – *Biochem. Syst. Ecol.* 22: 807–812.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. – *Proc. Natl Acad. Sci. USA* 70: 3321–3323.
- Ohno, S. 1970. *Evolution by gene duplication*. – Springer Verlag.
- O’Kane, S. L., Schaal, B. A. and Al-Shehbaz, I. A. 1996. The origin of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. – *Syst. Bot.* 21: 559–566.
- Price, R. A., Al-Shehbaz, I. A. and Palmer, J. D. 1994. Systematic relationships of *Arabidopsis*: a molecular and morphological perspective. – In: Meyerowitz, E. and Somerville, C. (eds), *Arabidopsis*. Cold Spring Harbor Laboratory Press, p. 7–19.
- Richards, A. J. 1986. *Plant breeding systems*. – Gore Allen & Unwin, London.
- Sharbel, T. F. and Mitchell-Olds, T. 2001. Recurrent polyploidy origins and chloroplast phylogeography in the *Arabis holboellii* complex (Brassicaceae). – *Heredity* 87: 59–68.
- Soltis, D. E. and Soltis, P. S. 1993. Molecular data and the dynamic nature of polyploidy. – *Crit. Rev. Plant Sci.* 12: 243–273.
- Soltis, D. E. and Soltis, P. S. 1995. The dynamic nature of polyploid genomes. – *Proc. Natl Acad. Sci. USA* 92: 8089–8091.
- Säll, T., Jakobsson, M., Lind-Hallden, C. et al. 2003. Chloroplast DNA indicate a single origin of the allopolyploid *Arabidopsis suecica*. – *J. Evol. Biol.* 16: 1019–1029.
- Wendel, J. K. 2000. Genome evolution in polyploids. – *Plant Mol. Biol.* 42: 225–249.
- Wright, S. 1951. The genetical structure of populations. – *Ann. Eugen.* 15: 323–354.
- Wright, S. 1968. – In: *Evolution and genetics of populations*, Vol. 1. The Univ. of Chicago Press.
- Yang, Y.-W., Lai, K.-N., Tai, P.-Y. et al. 1999. Molecular phylogenetic studies of *Brassica*, *Rorippa*, *Arabidopsis* and allied genera based on the internal transcribed spacer region of 18S-25S rDNA. – *Mol. Phylogeny Evol.* 13: 455–462.