USE OF THESES

This copy is supplied for purposes of private study and research only. Passages from the thesis may not be copied or closely paraphrased without the written consent of the author.
Advantages of chasmogamy and cleistogamy in a perennial Glycine clandestina Wendl. (Fabaceae)

Katarzyna Anna Hempel

July 2004

A thesis submitted for the degree of Doctor of Philosophy of
The Australian National University
For Jim
I verify that this thesis contains original work, great majority of which I carried out independently. I collaborated on the production of the paper describing the transfer of microsatellites with my supervisor, Dr. Rod Peakall. He provided advice on sequencing and performed the analysis of the sequences. I also created all of the images included in thesis apart from the drawings on the title page of Chapter 1, which come from copyright-free sources.

Katarzyna Hempel
Acknowledgments

First and foremost, I would like to thank my supervisory panel for volunteering their time to the good cause of completion of this work. I would like to thank my supervisor, Dr. Rod Peakall, for the continuing support, in particular his assistance with the population genetics aspect of my thesis. I am also very grateful for the enthusiastic support I always received from Dr. Mike Ramsey, which amounted to a substantial assistance in design and analysis of experiments in Chapter 3 and 4. I would also like to thank other members of my supervisory panel, Dr. Julian Ash and Dr. Tony Brown for their regular comments.

I am also grateful to the friends who braved the scorching heat of summer days in the field, for assisting in my fieldwork and, more importantly, for keeping me company: Kathy Tsang, Simon Mockler and Stuart Dennis. Andy Leigh is a saint for saving me from insanity of doing all of the pollen counts by myself. I also owe lots more beers to Inspector Simon Gilmore, Dan Ebert, Trish Hogbin, Marita Sydes, Lyn Cook, Siegfried Krauss, Tristan Armstrong, and Lolly Grasso for the laughs in the lab and the hints on how to make the PCR magic work. The beautiful SEM pictures of the plants sexy bits could not have happened without the wonderful assistance from Roger Heady and the electron microscopy unit at the RSBS, ANU. My inbreeding depression experiment would not have been the same without the help from the Plant Growth Facility staff and without flirting with Jesus.

My data analysis was facilitated by advisors from the Statistical Unit, ANU who were very skilful at clouding my understanding and Dr. Julian Ash and Dr. Mike Ramsey who effortlessly unclouded it again; Dr. Rod Peakall (and a little of Peter Smouse) teaching me everything there is to know about analysis of population differentiation; Dr. Kermit Ritland promptly providing his mating system analysis programs and advising on their use.

I would like to thank all the colleagues who by providing me with encouraging comments on my work presented at conferences and in written drafts managed to keep up my motivation and faith in the meaning of it all. I am also especially thankful for the patience of the courageous souls who gave feedback on my early (and subsequent)
drafts including my supervisory panel, Dr. Lyn Cook, Dr. Adrienne Nicotra, Dr. Scott Keogh, Dr. Shaun Cunningham, and Dr. James Procter.

The survival of the darkest hours would have been a struggle without my parents standing by me with a plate of delicious cookies and unbounded faith in my abilities. I will never forget all of my fellow dance fiends, Phoenicians, and the lively Canberra night- and day-life for providing timely relief from the grind. And finally, I would like to voluntarily give up innumerable c mocks to my muse and inspiration, jp, without whom this thesis would have remained unfinished.
Abstract

Dimorphic strategies for sexual reproduction are common in plants and their evolutionary dynamics is one of the central themes in plant evolutionary biology. True cleistogamy, or CLOP, is one such strategy where individuals use two types of flowers for reproduction: normal chasmogamous (or OP) flowers, which are open for insect pollination and so able to outcross; and reduced, closed cleistogamous (or CL) flowers specifically modified for self-pollination. This thesis examines the relative advantages of cleistogamous and chasmogamous reproduction in an Australian perennial species, *Glycine clandestina*. It focuses on inbreeding depression, which has the potential to favour chasmogamy, and resource and pollinator limitation as factors that may favour cleistogamy. The investigation described in this work integrated field and glasshouse experiments, reproductive biology and population genetics. It also involved the sourcing of informative microsatellite markers by cross-transfer from the congeneric crop, *G. max* (soybean).

A key element of the experimental design was the examination of two distinctive morphotypes – the high elevation broad leaf morphotype (BRO) and the low elevation narrow leaf morphotype (NAR). Despite expectations, no difference in reproductive strategy or population genetic structure was observed. The two classic advantages of cleistogamy - more reliable and economic seed production as compared to chasmogamy were confirmed in *G. clandestina*, however. It was also revealed that in the perennial *G. clandestina* cleistogamy functions similarly to many annuals – as a conditional reproductive strategy when resources are likely to limit chasmogamous seed production in small plants. Interestingly, in larger *G. clandestina* plants cleistogamy was also regulated to compensate for OP fruit set failure, forming a typically perennial conditional strategy. In contrast to the benefits of cleistogamous seed production, there was little evidence for advantages to chasmogamy from avoidance of inbreeding depression. Not only was the inbreeding similar in CL and OP seeds due to biparental inbreeding and selfing in chasmogamous flowers, but also the lack of detectable inbreeding depression at the juvenile stage suggested that deleterious alleles had been partially purged from populations. Moreover, the selfed cleistogamous progeny was of greater quality than the outcrossed chasmogamous progeny at this stage due to greater CL seed mass. Similarly, substantial inbreeding depression is often hard to show in
other cleistogamous species. This emphasises that the additional gains via male function in OP flowers may have a more important role than those of the female function in shaping the reproductive strategy of true cleistogamy. Importantly, findings in *G. clandestina* provide new evidence for the potential importance of the role of male function in promoting increase in investment in chasmogamy with plant size.

This study has contributed to our understanding of cleistogamy in perennial species and the patterns of genetic differentiation between and within populations, which may influence the relative advantage to OP from avoidance of inbreeding. It also demonstrates that *G. clandestina* is a study system where the male function hypothesis for OP can be profitably explored, and provides informative genetic markers for its further investigation.
# Table of contents

Acknowledgments ................................................................. IV
Abstract ................................................................................ VI
Table of contents .................................................................... 8

## Chapter 1: Introduction

**What is CLOP?** .................................................................... 14
**Evolutionary dynamics of CLOP** ........................................... 15
  - CLOP as a dimorphic reproductive strategy
  - Unique opportunities for fitness gain via cleistogamy and chasmogamy
  - CLOP as a reproductive strategy for heterogenous environments

**The scope of CLOP studies** .................................................. 19
**Study species and populations** .............................................. 20
  - Study species – *Glycine clandestina*
  - *G. clandestina*'s reproductive strategy
  - Microsatellites in *G. clandestina*
  - Study populations

**Aims and thesis structure** ..................................................... 24

## Chapter 2: Cleistogamy as a more reliable and cost effective seed source than chasmogamy in *Glycine clandestina*

**Introduction** ....................................................................... 33

**Materials and Methods** ..................................................... 38
  - Individual investment in cleistogamy and chasmogamy
  - Relative efficiencies of CL and OP seed production
  - Pollinator limitation in OP flowers

**Results** .............................................................................. 46
  - Individual investment in cleistogamy and chasmogamy
  - Cleistogamy as a more efficient source of seed
  - Inefficient pollination mechanism in OP flowers

**Discussion** ......................................................................... 53
  - Function of cleistogamy as a more reliable source of seed
  - Function of cleistogamy as a more economical source of seed

**Conclusions** .................................................................... 65

## Chapter 3: Limited inter-population variation in chasmogamous inbreeding and small-scale genetic structure in *Glycine clandestina* – implications for the advantage to chasmogamy from avoidance of inbreeding depression

**Introduction** ....................................................................... 83

**Materials and Methods** ..................................................... 87
  - Molecular analyses
  - Genetic structure at three spatial scales
Self-fertilisation rates within populations
Statistical analyses

Results
Genetic structure at three spatial scales
Inbreeding and selfing in populations

Discussion

Conclusions

Chapter 4: Non-genetic effects on selfed seed from cleistogamous flowers cause better seedling performance than outcrossing in chasmogamous flowers of *Glycine clandestina*

Introduction

Materials and Methods
Sampling and design
Glasshouse experiment
Determination of mating type in OP progeny
Statistical analysis

Results
Reproductive mode affects seed mass and seedling size
Reproductive mode affects the relationship between seed mass and seedling size

Discussion
Seedling performance enhanced by greater cleistogamous seed mass
Outcrossing does not improve seedling performance
Evolutionary implications for CLOP in *G. clandestina*
Comparison with other CLOP species

Conclusions

Chapter 5: Importance of male function in favouring chasmogamy in *Glycine clandestina* – discussion and new hypotheses

Sexual selection via male function in plants
Male function hypothesis for chasmogamy in CLOP
Evidence for the male function hypothesis for chasmogamy in *G. clandestina*
Male function hypothesis for chasmogamy in other CLOP and other plants
Conclusions

Chapter 6: Summary and future directions

References

Appendix: Cross-species amplification from crop soybean *Glycine max* provides informative microsatellite markers for the study of inbreeding wild relatives
Chapter 1

Introduction

Viola sp.

open flowers

cleistogamous flowers

soyabean

Impatiens capensis

peanut
What is CLOP?

The term cleistogamy was derived from Greek (*kleistos* = closed, and *gamein* = marriage) by Khun in 1867 to describe the production of flowers that self in the bud, in addition to or instead of normal open or chasmogamous flowers (Uphof 1938; Lord 1981). Based on her developmental studies and a comprehensive review of others, Lord (1981) refined this definition. Now true cleistogamy is understood as a real floral dimorphism, with the “two types of flower arising within an individual or a species as a result of the divergent developmental pathways” (Lord 1981, p.422). True cleistogamy has been described in over 148 angiosperm species (Lord 1981). In these species, cleistogamous flowers are a modified form of open flowers that result from a change in the timing of developmental events (heterochrony) rather than just a failure of anthesis (flower opening) or pollination in bud followed by anthesis (Lord and Hill 1987). In this thesis these plants and their strategy will be referred to as CLOP, the two modes of reproduction as cleistogamy and chasmogamy, and cleistogamous or chasmogamous as CL or OP respectively.

A true CL flower is characterised by reductions in corolla size and either size or number of stamens and pollen grains (androecium) compared to OP flower. It also lacks pollinator rewards and attractants (nectar, scent) that are usually present in OP flowers. In addition to the reduced size of the organs, the developmental modification leads to autonomous self-pollination without the need for anthesis, which causes the CL flower to look like an immature OP bud. Pollination in CL flowers is achieved by anthers dehiscing to release the pollen onto a receptive stigma, or more commonly by pollen germinating within undehisced anthers and pollen tubes penetrating the anther wall, followed by stigma or style to reach the ovaries (Lord 1984; Lord and Hill 1987). In contrast, effective pollination in OP flowers usually requires mediation of a vector (either animal or wind) in pollen transfer. The divergent pollination mechanism in the two flower types result in different types of mating - CL flowers are obligately self-fertilised whereas OP flowers allow for matings with other plants in the population (outcrossing).

Reviews of cleistogamy have found that in the great majority of CLOP species both floral forms occur on the same individuals at some point in their life cycle, despite that
expression of cleistogamy varies greatly among species (Darwin 1897; Uphof 1938; Lord 1981; Campbell et al. 1983; Schoen and Lloyd 1984; Cheplick 1987; Plitmann 1995). Species differ in the proportion of CL flowers, their relative positions and the timing of their production. Sometimes the differences in flower morphology extend to the fruit or seed derived from the two flower types. An extreme example of such fruit dimorphism is amphicarpy, a rare CLOP strategy, where CL fruits are produced below or near the soil surface and contain larger seeds while OP fruits with much smaller seeds are produced aerially (Cheplick 1987).

The ecological correlates of CLOP are also diverse. Life histories include annual and perennial herbs in equal proportions (Campbell et al. 1983; Plitmann 1995), OP pollination occurs by animal (birds and insects) or wind vectors (Schoen and Lloyd 1984), habitats range from rainforests to deserts (Lord 1981; Campbell et al. 1983; Cheplick 1987), and seed dispersal mechanisms include mymerchochorous (Le Corff 1996), ballistic (Schmitt et al. 1985) or gravitational (Berg and Redbo-Torstensson 1999). This diversity in ecology is accompanied by the wide phylogenetic base of CLOP species, spanning both dicots and monocots and at least 28 families (Lord 1981). This indicates that CLOP has evolved independently many times.

Evolutionary dynamics of CLOP

CLOP as a dimorphic reproductive strategy

Dimorphic cleistogamy is a type of a dimorphic strategy, where two morphologically distinct structures fulfil the same function of reproduction (Harper 1977; Lloyd 1983; Willson 1983). Dimorphic strategies are unique to plants, whose modality allows for the required organ redundancy (Lloyd 1983; Cox 1988). Because dimorphic strategies of sexual reproduction are common in plants their evolutionary dynamics constitutes one of the central themes in plant evolutionary biology (e.g. simultaneous hermaphroditism Charnov 1982; mixing of self-fertilisation and outcrossing Lande and Schemske 1985; diaspore dimorphism Venable 1985).

The interest in an explanation for adaptive value of CLOP was sparked by Darwin’s treatment in his seminal work on floral dimorphisms ‘The different forms of flowers on plants of the same species’ (1897). The early ideas put forward by Darwin and others were later incorporated in a formal, evolutionarily stable strategy (EES) model for
CLOP by Schoen and Lloyd (1984). This model has since provided a framework for investigating the adaptive value of CLOP. Below, I briefly review the factors that in theory are considered central in maintaining the cleistogamy-chasmogamy balance.

Unique opportunities for fitness gain via cleistogamy and chasmogamy

Theoretically, in order for the dimorphic reproductive strategy to be evolutionarily stable, the two types of flowers should provide non-interchangeable opportunities for individual fitness gains either at the stage of progeny production or progeny performance (Lloyd 1983; Schoen and Lloyd 1984). Cleistogamy has been universally viewed to increase individual fitness by providing reproductive assurance and increase in fecundity (number of progeny), while OP reproduction is considered to provide progeny of better quality (e.g. Darwin 1897; Waller 1980; Schoen and Lloyd 1984; Le Corff 1993; Berg 2000b).

Fitness gains via cleistogamy

Reliable production of seed

Modifications typical of CL flowers are expected to assure autonomous self-pollination (Lord 1981, 1984), while OP flowers are expected to be reliant on the vagaries of vectors like wind or animals for pollination. As insufficient pollination is a major factor limiting seed set of chasmogamous flowers in plants (Bierzychudek 1981; Burd 1994; Larson and Barrett 2000), the different pollination mechanisms in CL and OP flowers are likely to result in greater fertilisation certainty in CL flowers and thus greater seed production. For the advantage to cleistogamy to be based on high fertilisation certainty in CL flowers, OP flowers are expected to lack the ability for autonomous pollination and to be vulnerable to pollinator limitation (Schoen and Lloyd 1984).

Cost-effective production of seed

What makes cleistogamy unique is that greater fertilisation certainty is associated with a reduced material investment in flowering because of the reduction or elimination of structures associated with pollinator attraction (ancillary costs: corolla, nectar) or those directly involved with male function (male costs: pollen, anthers) (reviewed in Darwin 1897; Uphof 1938; Lord 1981; Campbell et al. 1983; Plitmann 1995). Darwin hypothesised that the lower material costs of CL flowering should make it a more cost effective mode of reproduction than chasmogamy (Darwin 1897). Thus cleistogamy is expected to overcome another factor commonly limiting seed set in plants, the resource
shortage (Stephenson 1981; Willson and Burley 1983). The overall resource saving is still a common assumption used in explanations of adaptive value of cleistogamy in CLOP species (Jasieniuk and Lechowicz 1987; Waller 1988; Mattila and Salonen 1995; Berg and Redbo-Torstensson 1998; Sun 1999).

The assumption of the overall resource saving on reproduction via cleistogamy compared to chasmogamy may not be valid. Schoen and Lloyd (1984) argued that for cleistogamy to evolve, an enhancement of fitness gains via female function is required. This may be expressed as greater seed number per plant or fecundity, or larger seeds with greater viability or progeny quality. This is because, unlike many other forms of selfing, cleistogamy does not allow for the export of pollen so it is not expected to be favoured by the 50% automatic transmission advantage that normally offsets the effects of inbreeding depression in other selfers (Lloyd 1979; Damgaard et al. 1992; Jarne and Charlesworth 1993). In CLOP it is the increased investment in female function that replaces the transmission advantage to balance out gains from outcrossing and avoidance of inbreeding in chasmogamy (see below)(Schoen and Lloyd 1984).

Increasing fecundity through female function or seed production is recognised as costly in plants (Harper 1977; Charnov 1982). Theoretically, it is the saving on the costs of flowering that covers the increased expense of seed production via cleistogamy (Schoen and Lloyd 1984). Therefore, CL reproduction will only be cheaper overall if the savings on flowering exceed the cost of increase in female function (Campbell et al. 1983).

*Fitness gains via chasmogamy – greater seed quality*

Darwin (1897) was the first to point out that a principal source of fitness gains from OP reproduction is the production of outcrossed seed. He considered outcrossing to produce superior offspring because they avoid inbreeding depression (ID) – the depression of fitness associated with inbreeding (Darwin 1876). Inbreeding depression is now a well-established empirical phenomenon (Charlesworth and Charlesworth 1987b; Thornhill 1993; Husband and Schemske 1996; Byers and Waller 1999), and it continues to be included as a central force in favour of outcrossing in evolutionary models for the plant mixed mating strategies (Lande and Schemske 1985; Campbell 1986; Uyenoyama 1986; Charlesworth and Charlesworth 1987b; Holsinger 1988; Holsinger 1991; Damgaard et al. 1992; Jame and Charlesworth 1993; Barrett 1995; Ronfort and Couvet 1995; Sakai 1995; Barrett and Harder 1996).
In CLOP the role of ID in providing an advantage to chasmogamy can be influenced by several factors. The magnitude of the difference in quality of seed between chasmogamy and cleistogamy due to ID is dependent on the difference in the type of mating allowed by CL and OP flowers. Any selfing or inbreeding in OP flowers will diminish this difference and hence the superiority of OP seed (Schoen and Lloyd 1984; Masuda et al. 2001). The enhancement of the female function leading to increase in CL seed quality, predicted in CLOP species, may also diminish the relative advantage to chasmogamy based on progeny quality (Schoen and Lloyd 1984).

The long-term interest in ID has resulted in much theoretical and empirical work and it is now appreciated that it is a complex phenomenon and there is much debate about its importance in habitual selfers like CLOP. One theoretical prediction important for the role of ID is that its magnitude will coevolve with the level of self-fertilisation (Lande and Schemske 1985). This relies on the assumption that inbreeding depression is due to genetic load of recessive mutations, which is expressed in inbred homozygous individuals. Extended periods of inbreeding are expected to reduce ID, because the rise in homozygosity exposes deleterious recessive mutations of large effect to selection, increasing the effectiveness of selection against them to remove or purge them from inbred populations (Lande and Schemske 1985; Charlesworth and Charlesworth 1987b). While such purging is supported by some empirical evidence, other work shows that the nature and the degree of purging may depend on many factors including the genetic basis of ID, mutation rates as well as the dominance and selection coefficients of the underlying mutations (Schemske and Lande 1985; Charlesworth et al. 1990; Husband and Schemske 1996; Byers and Waller 1999; Crnokrak and Barrett 2002).

The effects of ID on the advantage to outcrossing in selfers may be further complicated by the strong population genetic structure or genetic differentiation that is characteristic of selfers. Habitual inbreeding leads to the reduction in effective population size (Pollack 1987) and the reduction of the effective recombination rates (Narrain 1966). The former enhances effects of genetic drift and isolation by distance (Wright 1978) and the latter, the effects of background selection and genetic hitchhiking (Charlesworth et al. 1997; Nordborg and Donnelly 1997). The result is an enhancement of genetic differentiation at multiple scales. In selfers most of the genetic variation is found between populations (reviewed by Hamrick and Godt 1996), and within populations a
structure develops of genetically distinct patches or subpopulations consisting of related individuals (small scale population genetic structure) (reviewed by Heywood 1991).

The most important consequence of enhanced genetic structure is its effect on the relative inbreeding of outcrossed progeny. If outcrossing happens within subpopulations it is likely to result in mating with relatives and biparental inbreeding (inbreeding in outcrossed progeny due to consanguineous matings), thus diminishing the ID-based advantage to outcrossing (Uyenoyama 1986). However, an opposite effect is expected when outcrossing occurs between different inbred subpopulations, because heterozygosity of such crosses will be enhanced due to genetic drift tending to fix different random subsets of alleles in neighbouring populations (Waller 1993; Ronfort and Couvet 1995; Whitlock 2000; Theodorou and Couvet 2002). On the other hand if the inter-patch crosses disrupt genetic complexes responsible for local adaptation, the progeny may suffer from outbreeding depression (Waser 1993).

**CLOP as a reproductive strategy for heterogenous environments**

Central to the study of mixed strategies is the search for factors that maintain the alternative structures in the evolutionarily stable balance instead of one of the structures being driven to fixation. The reproductive strategy of CLOP is thought to be a case of a strategy where the mixing of alternative structures allows for optimisation of fitness in heterogenous environments (Schemske 1978; Waller 1980; Schoen and Lloyd 1984). In such strategies each type of structure is designed to gain fitness via optimisation of reproductive output or progeny performance in a different environment (Lloyd 1983; Venable 1985).

For such a strategy to be adaptive the individuals are expected to assess environmental conditions and adjust the allocation to the two structures to match relative incidence of the relevant environments (multiple strategy), or switch between them when the environment changes (conditional strategy) (Lloyd 1983). This is consistent with empirical evidence in CLOP that the variation in the proportion of cleistogamy between and within populations is often environmentally determined (Schemske 1978; Waller 1980; Clay 1982a; Wilken 1982; Bell and Quinn 1987) or that the two types of flowers are produced at different locations and different times within plants (Schemske 1978; Jasieniuk and Lechowicz 1987; Mattila and Salonen 1995).
Because of the perceived cost effectiveness and fertilisation reliability of the CL flowers it has been proposed that cleistogamy acts to assure reproduction or increase fecundity under ‘harsh conditions’, when seed production in OP flowers would otherwise be limited by resources or pollinators (Darwin 1897; Uphof, 1938; Darwin, 1897; Schoen and Lloyd, 1984; Campbell, 1983). Otherwise OP flowers would be favoured because they are expected to produce seed of better quality due to effects of ID.

The scope of CLOP studies

Secondary features of CLOP reproductive strategies, and ecological and phylogenetic diversity of CLOP species have lead to the recognition that the relevant selection pressures may vary considerably among the species. Thus our understanding of the evolutionary dynamics of the strategy benefits from the investigation of a broad range of CLOP species (Schoen and Lloyd 1984; Lord 1981; Campbell et al. 1983). To date CLOP has been well studied in some groups, e.g. grasses where wind pollination is important (Campbell et al. 1983), and amphicarps where divergent CL and OP seed morphology and dispersal are important (Cheplick 1987). Most substantial body of work exists in the annual balsams (Impatiens: Balsamineae), representative of animal pollinated species with aerial cleistogamy (e.g. Schemske 1978; Waller 1984; Antlfinger 1986; Schmitt and Ehrhardt 1987; Masuda and Yahara 1994; Stewart 1994; Gross et al. 1998; Lu 2002).

Currently there is an interest in extending the study of the evolutionary dynamic of CLOP to the perennial systems in recognition that despite their numerical prevalence among CLOP species we know little about them (e.g. Le Corff 1996; Berg 2000b; Culley 2002). Theoretically, perennial CLOP may be expected to differ from annuals because any increase in fecundity provided by CL reproduction can be discounted by the trade offs with future reproduction, considered important in perennial plants (Harper 1977). There is already some evidence that suggests different responses to resource and pollinator limitation compared with annuals (Clay 1982a; Jasieniuk and Lechowicz 1987; Mattila and Salonen 1995; Redbo-Torstensson and Berg 1995; Berg and Redbo-Torstensson 1998). However, most of the investigations to date focus only on two genera with very similar ecology, Viola and Oxalis, which are both Northern Hemisphere temperate understorey herbs of deciduous forests (Jasieniuk and Lechowicz 1987; Mattila and Salonen 1995; Redbo-Torstensson and Berg 1995; Berg and Redbo-
Torstensson 1998, 1999; Berg 2000a, b; Culley 2000; Culley and Wolfe 2001; Culley 2002).

It has long been appreciated, at least in theory, that the interaction between selfing rate and population genetic structure can affect the role of ID. Therefore, the study of population substructure and mating patterns is central for testing predictions regarding mating system evolution (Uyenoyama 1986; Charlesworth and Charlesworth 1987b; Waller 1993; Ronfort and Couvet 1995). The recent development of high-resolution markers has fuelled an increased interest in the estimates of population genetic structure in conjunction with selfing rates, in order to describe genetic differentiation patterns at multiple scales in highly selfing plants (e.g. Todokoro et al. 1995; Innan et al. 1997; Kuittinen et al. 1997; Bonnin et al. 2001). Such estimates are absent in CLOP species.

The lack of sufficient variability at allozyme loci in CLOP species curtails their use in assessment of population genetic structure or as molecular ecology tools (e.g. Waller and Knight 1989a; Cole and Biesboer 1992; Sun 1999). Microsatellites, or short sequence repeats (SSRs), are a newer class of molecular markers that overcome this limitation in selfers (e.g. Kuittinen et al. 1997; Viard et al. 1997; Bonnin et al. 2001). Microsatellite loci consist of stretches of DNA with tandem repeats of a short nucleotide motif, and variation is expressed as repeat number differences among alleles. They have the same desirable attributes as population genetic markers as allozymes (codominance, Mendelian inheritance, presumed neutrality) but with an added advantage of increased variability (Jarne and Lagoda 1996; Cruzan 1998).

**Study species and populations**

**Study species – Glycine clandestina**

Glycine clandestina (Fabaceae: Phaseolae) is a perennial twining herb with a woody taproot (Fig. 1.1)(Tindale 1991). It belongs to a group of >17 perennial Glycine species mostly endemic to Australia (subgenus Glycine), and closely related to the annual wild and cultivated soybeans (G. soja and G. max, subgenus Soja)(Hymowitz and Singh 1987). G. clandestina is widespread in SE Australia from the coastal to subalpine areas (Fig. 1.2), and is common in grasslands and understoreys of open Eucalyptus woodlands and forests (Tindale 1991; Pfeil et al. 2001). Within the geographic range of G. clandestina a number of distinct leaf morphotypes can be identified, indicating ecotypic
or genetic differentiation among its populations (Pfeil et al. 2001). As a perennial and a legume *G. clandestina* represents two groups of CLOP species in which the evolutionary dynamics are little studied although they are well represented among CLOP.

**G. clandestina's reproductive strategy**

In taxonomic descriptions of *G. clandestina* and its close relatives, the presence of both cleistogamous flowers and chasmogamous flowers on the same individual have been recorded (Tindale 1991), suggesting that the species employs a dimorphic reproductive strategy of CLOP. However, at the start of this study, little was known about this reproductive strategy in the perennial *Glycine*.

The differences between the two types flowers in size and appearance of *Glycine clandestina* are consistent with the other CLOP species (Fig. 1.3). OP flowers are typically pea-like, with structure specialised for bee-pollination (Fig. 1.3A,B). The pollinators recorded during this study included the exotic honeybee, *Apis mellifera* L. (Hymenoptera: Apidae) and three native Australian bees, *Lasioglossum* sp. (Hymenoptera: Halictidae), *Leioproctus* sp. (Hymenoptera: Colletidae), *Trichocolletes* sp. (Hymenoptera: Colletidae) (pers.obs., identified by Kim Pullen Identification and Advice Service, Australian National Insect Collection CSIRO Entomology). The flowers have a showy mauve corolla 5-10 mm across, with the standard petal acting as an attractant, keel to protect the staminal column, and the wings together with the keel acting as a landing platform for the pollinators. They also produce nectar as a reward (pers. comm. Reid G. Palmer, Iowa State University, U.S. and Brown, A.H.D, Plant Industry, CSIRO, Canberra, Australia). In contrast, CL flowers reach at most 3 mm, their petals are reduced in size, green and remain enclosed within the calyx at maturity (Tindale 1991) (Fig. 1.3C). Only when the ovary starts to enlarge after fertilisation, the style and then the developing fruit emerge from between the calyx lobes (Fig. 1.3D).

The two types of flowers are produced in distinct inflorescences in leaf axils: chasmogamous flowers are borne in 4-18 flowered racemes, while CL flowers are sessile and solitary or rarely in pairs (Tindale 1991). In *G. clandestina* there is some anecdotal evidence that CL and OP inflorescences also differ in timing of production and position on the plant, with CL flowers and fruit arising earlier and at lower leaf axils than OP flowers (Tindale 1991) (Tony A.H. Brown, pers. comm.). These
differences would place the two types of flowers in different environments within a plant and within a growth season, which suggests that they may be adaptive under distinct environmental conditions and employed as a conditional strategy in this species.

Seeds produced in CL flowers of perennial *Glycine* arise from self-fertilisation (Brown et al. 1986; Hempel and Peakall 2003) while OP seeds arise from mixed outcrossing and selfing (Brown et al. 1986; Schoen and Brown 1991). One study of mating in OP flowers using allozymes revealed that at least 50% of seed was derived from self-fertilisation in two populations of *G. clandestina* (Schoen and Brown 1991). Moreover, differences were detected between the populations in the magnitude of selfing and the patterns of selfing within fruit. The higher OP selfing rate (89%) was associated with high proportion of whole-flower or correlated selfing, while in the population where OP selfing rate was lower (53%), OP flowers experienced a mixture of selfing and outcrossing. As the former population was placed at a high elevation of 1500 m in a sub-alpine zone and the latter at 750 m in a temperate area, the authors attributed these differences in OP mating to differences in pollinator availability in the two environments. Schoen and Brown (1991) hypothesised that *G. clandestina* flowers are capable of autonomous self-pollination after the vector-mediated pollination fails (delayed pollination sensu Lloyd and Schoen 1992). They proposed that the high elevation population experienced more selfing and more whole-flower selfing because of inefficient pollinator service in this environment and thus higher rates of delayed and whole-flower selfing. In contrast, in the population at a lower elevation pollinator service was adequate so no delayed selfing occurred but rather selfing resulted from pollinator mediated pollen transfer within flowers or between flowers of the same plant (geitonogamous selfing).

The two populations in Schoen and Brown’s study (1991) also represent two distinct leaf morphotypes of *G. clandestina*. The high elevation population was of the broad leaf morphotype (BRO) of the species, which is characteristic of higher parts of the Great Dividing Range in New South Wales and Victoria, the low elevation population belonged to the narrow leaf morphotype (NAR) typical of Canberra area, Australian Capital Territory (Pfeil et al. 2001). Thus the patterns ascribed to differences in elevation could equally have been produced by ecotypic differentiation between the two morphotypes. The apparent variation in OP mating patterns in conjunction with
elevation and morphological differentiation in *G. clandestina* implies that populations may differ in their reproductive strategy. This creates a possibility of comparative study between populations that could yield additional insights into selective forces shaping the balance between cleistogamy and chasmogamy in this species.

**Microsatellites in *G. clandestina***

The widespread application of microsatellites in wild species is limited by a requirement for species-specific primers (Ashley and Dow 1994), the development of which remains a costly and lengthy process despite the continuing improvements in its efficiency (Zane et al. 2002). Transfer of primers (cross-species amplification) can offer an alternative to the *de novo* development in plants (Peakall et al. 1998, Rossetto, 2001). *G. clandestina* is related to the commercial crop, *G. max* (soybean), where a large set of microsatellites has been developed for genome mapping (Mughan et al. 1995; Diwan and Cregan 1997; Cregan et al. 1999). Thus *G. clandestina* is a good candidate for development of microsatellite markers through transfer of primers. The development of microsatellites via transfer from soybean constituted an important part of this study. For the sake of preserving the flow this material was excluded from the main body of the thesis and the publication describing it is included as the Appendix 1 and should be read as a part of the thesis.

**Study populations**

The populations for this study were selected to maximise the chance of capturing the local variation in the CLOP strategy suggested by Schoen and Brown’s study (1991). The sampling included the two local morphotypes and a range of elevations. The populations were sampled over a geographic range of approximately 35km, within the Canberra region of the Australian Capital Territory, Australia (see Table 1.1, Fig. 1.2). Four of the populations represented the BRO morphotype (FR, TE, B2, B1) and three of the populations represented the NAR morphotype (TA, AR, MA) (Fig. 1.2). The BRO populations were selected to span the range of elevations where the morphotype can be found (760-1260 m). The NAR populations only occurred at elevations corresponding to lower end of the range for BRO (630-780 m) (Table 1.1). All of the populations were located within open *Eucalyptus* forests or woodlands.
Aims and thesis structure

The overall aim of this thesis was to investigate the mostly likely factors that maintain the evolutionary balance between cleistogamy and chasmogamy in *G. clandestina*’s mixed reproductive strategy. Factors examined included inbreeding depression which has the potential to favour chasmogamy, and resource and pollinator limitation as factors that may favour cleistogamy. The effects of these factors on fitness gains via female function (seed production) in the two types of flowers are most commonly evoked in evolutionary explanations of cleistogamy.

Our understanding of the cleistogamous reproductive strategy, to date, is primarily based on annuals. Therefore, this thesis fills an important gap in our knowledge by focussing on a perennial system. This study also took advantage of the availability of highly polymorphic markers sourced by cross-transfer from the congeneric crop, *G. max* (soybean). This has enabled the first comprehensive study of population genetic structure and mating patterns in a cleistogamous species. The investigation integrated field and glasshouse experiments, reproductive biology and population genetics, within a comparative framework facilitated by the availability of the two distinct leaf morphotypes.

The thesis is divided into 6 chapters and an Appendix. In addition to this introductory chapter, it includes three data chapters (Chapter 2, 3, 4) and two discussion chapters (Chapter 5 and 6). Each of the data chapters includes their own introduction and detailed discussion of the results together with their evolutionary interpretation. Figures and Tables are located at the end of each chapter. Appendix 1 contains the publication describing the transfer of microsatellite primers used in this study together with the assessment of indicators of successful transfer. The specific objectives for each chapter are summarised below.

In Chapter 2, I examined whether the two postulated advantages of cleistogamy over chasmogamy, the greater economy and reliability of seed production, determine the function of cleistogamy in *G. clandestina* and whether the *G. clandestina* individuals employ cleistogamy in a reproductive strategy similar to other perennials. This was achieved by a study of reproductive patterns and field experiments in four natural populations across the two morphotypes and a range of elevations. In order to determine
if cleistogamy functions as a cost saving mode of reproduction under resource limitation the following questions were addressed:

1. How does the relative investment in cleistogamy change in relation to plant size?
2. Is cleistogamy more cost efficient at flowering as well as fruiting compared to chasmogamy?

In order to determine if cleistogamy functions as a more reliable means of seed production under pollinator limitation the following questions were addressed:

1. How does the investment in cleistogamy change in response to chasmogamous fruit set?
2. Does morphology of pollination in CL and OP flowers support autonomous self-pollination in both flower types?
3. Are OP flowers able to autonomously initiate fruit in the field and is fruit initiation pollinator limited?
4. Are CL flowers more reliable at producing seed and fruit than OP flowers?

The study described in Chapter 3 had two objectives: firstly, to examine the patterns of selfing and associated genetic structure of *Glycine clandestina* populations; and secondly to determine if there are any differences between *G. clandestina* populations of BRO and NAR morphotypes that could potentially affect the relative advantage to chasmogamous reproduction derived from inbreeding avoidance. In order to achieve this, I analysed population genetic patterns at 11 microsatellite loci among and within seven *G. clandestina* populations and combined this with mating pattern estimates within four representative populations of both morphotypes. The results of this survey also aided the selection of representative populations for study in Chapter 2 and 4. The following specific questions were addressed:

1. What are the patterns of genetic differentiation at three scales: between morphotypes, among populations and within populations?
2. What is the level of inbreeding in mature plants and does it correspond to direct estimates of total selfing combining CL- and OP-selfing?
3. What is the level of inbreeding in OP flowers, including selfing and biparental inbreeding?
4. Is there evidence for inter-morphotype and inter-population differences in inbreeding, OP mating patterns, and small-scale spatial genetic structure?
In Chapter 4, I investigated the advantage to chasmogamy derived from avoidance of inbreeding depression (ID) and the advantage to cleistogamy derived from greater CL seed mass in *G. clandestina* in relation to early seedling performance. This was examined in a glasshouse experiment comparing the performance of three types of progeny: selfed chasmogamous, outcrossed chasmogamous and selfed cleistogamous. Naturally pollinated seed from one population was used and the selves and outcrosses were identified *a posteriori* based on their microsatellite genotypes. The study addressed two specific questions:

1. What are the effects of the reproductive mode (cleistogamy vs chasmogamy) and ID (selfing vs. outcrossing) on early performance indicators in progeny, including seed size, early growth, date of emergence and seedling size?

2. Does reproductive mode and ID affect a relationship between the seedling size as the juvenile fitness indicator and the seed size, early growth and date of emergence.

In Chapter 5, I discuss the potential for gains from male function to favour chasmogamy in addition to more traditionally recognised effects of inbreeding depression on the female function in the two flower types. This aspect has been neglected in study of cleistogamy, despite recognition of its theoretical importance by some authors and despite the growing support for the role of male function in evolution of other reproductive behaviours in plants. The results of this study indicate that the male function in *G. clandestina* is an important consideration and future research directions exploring this aspect are proposed.

Chapter 6 provides a general summary and conclusions.
Table 1.1  Location of populations of *G. clandestina* included in the study. See also Fig. 1.3.

<table>
<thead>
<tr>
<th>Morphotype</th>
<th>Population</th>
<th>Location</th>
<th>Altitude (m)</th>
<th>Geographic coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NAR</strong> (narrow leaf)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Canberra Nature Reserves: Aranda Bushland</td>
<td>630</td>
<td>35°16'00&quot;S, 149°04'50&quot;E</td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>Mt Taylor</td>
<td>720</td>
<td>35°22'25&quot;S, 149°04'50&quot;E</td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>Mt Majura</td>
<td>780</td>
<td>35°14'05&quot;S, 149°10'20&quot;E</td>
<td></td>
</tr>
<tr>
<td><strong>BRO</strong> (broad leaf)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>Mt. Franklin Rd., Namadgi National Park</td>
<td>1260</td>
<td>35°25'00&quot;S, 148°47'50&quot;E</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>Mt. Tennet, Namadgi National Park</td>
<td>1200</td>
<td>35°33'50&quot;S, 149°02'50&quot;E</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>Brindabella Rd., ACT Forest</td>
<td>940</td>
<td>35°22'30&quot;S, 148°45'50&quot;E</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1.1  A typical *G. clandestina* plant twining around a bush in its habitat (A) and *G. clandestina* plants growing in a glasshouse (B). Bars represent 15 cm.
Fig. 1.2 Distribution of *G. clandestina* in Australia (based on Pfeil *et al.* 2001) (A) and the locations of study populations within the Australian National Territory (B). The narrow leaf morphotype (NAR) populations are represented by • and the broad leaf morphotype (BRO) populations by ■. See Table 1.1 for further location details of the populations.
CHASMOGAMOUS FLOWERS

A. Chasmogamous or open flowers

B. View of OP racemes at the end of flowering, showing the position of racemes in leaf axils, the initiated fruit and bracts persisting where flowers failed to set fruit

CLEISTOGAMOUS FLOWERS

C. A mature CL flower in a leaf axil (NAR morphotype)

D. A CL flower developing into a fruit with a characteristically bent style still visible (BRO morphotype)

Fig. 1.3 Chasmogamous or open flowers (OP flowers: A, B) and a cleistogamous flower (CL flowers: C, D) and developing fruit in *Glycine clandestina*. The lines represent 1 cm.

**OP flowers**: A. OP raceme at onset of anthesis with flowers at different developmental stages; B. View of OP racemes at the end of flowering, showing the position of racemes in leaf axils, the initiated fruit and bracts persisting where flowers failed to set fruit

**CL flowers**: C. A mature CL flower in a leaf axil (NAR morphotype); D. A CL flower developing into a fruit with a characteristically bent style still visible (BRO morphotype).
Chapter 2

Cleistogamy as a more reliable and cost effective seed source than chasmogamy in *Glycine clandestina*
Introduction

As outlined in Chapter 1, theory predicts that plants should employ cleistogamous (CL) in preference to chasmogamous (OP) flowers for reproduction under conditions of resource or pollinator limitation. This prediction is based on the assumption of greater economy and reliability of seed production in CL compared to OP flowers (Darwin 1897; Lloyd 1983; Schoen and Lloyd 1984). At the same time, it is expected that the ecological and phylogenetic diversity would cause differences in the role of relative economy and reliability of cleistogamous seed production among species (Schoen and Lloyd 1984).

Comparisons of fruit initiation, fruit set and seed set between the two types of flower in some CLOP species show that the relative reliability of CL and OP seed production can vary among populations and seasons (Schemske 1978; Redbo-Torstensson and Berg 1995; Berg and Redbo-Torstensson 1998), and that CL and OP seed may be produced with equal probability (e.g. Berg and Redbo-Torstensson 1998; Culley 2002; Lu 2002). One of the reasons for the lack of differences in seed production reliability is the capacity for autonomous selfing in chasmogamous flowers. Autonomous selfing has been confirmed to provide high chasmogamy fertility in pollinator absence in some CLOP species (Seed set per fruit: 40% Collomia grandiflora, Wilken 1982; Fruit set: 90% in Lamium amplexicaule, Bernstrom 1950, 100% in Ceratocapnos heterocarpa, Ruiz de Clavijo and Jimenez 1993, 94% in Polygonum thunbergii, Momose and Inoue 1993; Fruit initiation: 60% in Viola pubescens, Culley 2002).

The post-pollination limitations on seed production can also potentially diminish the difference in reliability of seed production between the two types of flowers. Only several studies in other CLOP species found near complete fruit set in CL in contrast to very poor performance of OP flowers (McNamara and Quinn 1977; Campbell 1982; Schoen 1984; Schnee and Waller 1986), showing that cleistogamous flowers may suffer from post-fertilisation limitations on fecundity. Indeed, the comparisons between fruit initiation and fruit set available for some CLOP species show that both cleistogamous and chasmogamous fecundity can be strongly limited post-pollination. For example, in Oxalis acetosella the fruit initiation rate was 20-81% for chasmogamy and 53-89% for cleistogamy, but only 29-88% and 13-68% of the initiated fruit matured respectively.
(Berg and Redbo-Torstensson 1998). In other species OP fruit set is clearly limited by factors other than pollination. For example, in *Viola pubescens* OP fruit were initiated at 60-89% rate while the fruit set was only 5-12.5% (Culley 2002), while in *Polygonum thunbergii* the resource limitation, rather than pollination, was shown to limit OP fruit set (Momose and Inoue 1993).

The greater economy of cleistogamy is primarily claimed based on the comparison of size or cost between chasmogamous and cleistogamous flowers (Jasieniuk and Lechowicz 1987; Waller 1988; Mattila and Salonen 1995; Redbo-Torstensson and Berg 1995; Berg and Redbo-Torstensson 1998; Sun 1999). However, an increased investment in CL female function, including seed production, is predicted in theory (Schoen and Lloyd 1984). Accordingly, cleistogamy would not be adaptive under resource limitation (or provide an overall resource saving) when such an increased investment in seed exceeded the savings at the flowering stage (Campbell *et al.* 1983; Schoen 1984; Schoen and Lloyd 1984). Indeed, this may be the case in some species where the size of cleistogamous seed is typically strikingly larger than chasmogamous seed, such as amphiicarps (Cheplick 1987) and some grasses (Campbell *et al.* 1983). In other species, where both flower types are produced aerially, little evidence for such stable differences in seed size has been found (Weatherwax 1928; Cope 1966; Waller 1982; Clay 1983; Berg 2000b, but see Schnee and Waller 1986; Sun 1999), so in those species the savings at the flowering stage are likely to result in greater economy of cleistogamous seed production.

As in any mixed reproductive strategy, important evidence for the role of cleistogamy comes from patterns of relative allocation to production of the two flower types, depending on the plant condition. The proportional investment in cleistogamous and chasmogamous reproduction typically varies with environmental conditions, the phenomenon dubbed ‘environmental cleistogamy’ (Uphof 1938) or ‘environmental chasmogamy’ (Waller 1980). The evidence in support of the adaptive value of CL flowers under resource limitation has been found in many species where smaller plants (Lu 2002; Le Corf 1993; Schmitt 1987; Waller 1980; Wilken 1982; Schnee and Waller 1986; Cheplick 1994; Diaz and Macnair 1998) or plants exposed to conditions limiting growth (Bell and Quinn 1987; Le Corff 1993; Mattila and Salonen 1995) were shown to employ greater proportion of CL flowers, often to the exclusion of OP flowering. The
support for the function of cleistogamy for assurance of seed set under OP pollination failure comes from observations that the timing of production of CL and OP flowers correlates with pollinator availability (e.g. Schemske 1978; Masuda and Yahara 1994), and that CL flower production correlates with diminishing success of OP flowers within a reproductive season (e.g. Gara and Muenchov 1990; Masuda and Yahara 1994).

Interestingly, recent studies of perennials describe a unique pattern of investment in cleistogamy in response to both, plant size and chasmogamous pollinator limitation. A number of perennial species showed an increase in proportion of cleistogamy with plant size and resources (Clay 1982a; Jasieniuk and Lechowicz 1987; Mattila and Salonen 1995; Berg and Redbo-Torstensson 1998; but see Bell and Quinn 1987; Le Corf 1993). As there was no reason to expect that cleistogamous seed was more costly than chasmogamous in these species, this was contrary to the prediction and observation in annuals that, the more economical, cleistogamy should be preferentially employed by smaller plants. A new form of response to chasmogamous pollinator limitation was also observed in the perennials, *Viola hirta* and *Oxalis acetosella*. In both of the species, CL flower production was increased in response to OP fertilisation failure (Redbo-Torstensson and Berg 1995; Berg and Redbo-Torstensson 1998), which was consistent with greater reliability of cleistogamy, but has never been reported for annuals. Because perennial life history imposes unique constraints on reproductive allocation, due to the demands of repeated reproduction and resource storage (Harper 1977), the unique patterns may represent a distinct perennial CLOP strategy which will require an evolutionary explanation (Jasieniuk and Lechowicz 1987; Berg and Redbo-Torstensson 1998). On the other hand, the evidence for these patterns is mainly derived from only two genera, *Viola* and *Oxalis*, which share similar ecology of vernal understorey herbs in deciduous forests of the Northern Hemisphere.

Investigation of the patterns of cleistogamous reproduction in response to plant size and chasmogamous failure in *G. clandestina* will help to ascertain the generality of these perennial reproductive strategies in response to the two key factors favouring cleistogamy, pollinator and resource limitation. In *G. clandestina* CL flowers are smaller than OP flowers and, typically of the species with aerial cleistogamy, there is no pronounced seed dimorphism that would indicate an increased investment in female function in cleistogamous flowers (Tindale 1991). Thus in *G. clandestina*, the overall
cost of seed production should be lower for cleistogamous than chasmogamous flowers, rendering cleistogamy an advantageous reproductive mode under resource limitation. On the other hand, the OP flowers in *G. clandestina* appear capable of autonomous self-pollination and thus may not suffer from pollinator limitation to a greater extent than cleistogamous flowers. This is evident from good fruit set in OP flowers in the absence of pollinators under glasshouse cultivation in this and closely related species (pers. obs.; A.D.H. Brown pers. comm.; (Brown *et al.* 1986). Moreover, Schoen and Brown (1991) suggested that OP flowers of *G. clandestina* were also capable of autonomous selfing in the field, based on the high degree of correlated selfing they found in OP flowers in a high elevation population of broad leaf (BRO) morphotype. Importantly, the differences in the level of correlated mating between this and a lower elevation, narrow leaf morphotype (NAR) population in their study can be interpreted as evidence for differences in ability for autonomous selfing and/or pollinator activity (Schoen and Brown 1991). It is likely that *G. clandestina* has a similar mechanism of autonomous selfing to the related crop, *G. max*. It is well known that in *G. max* OP flowers pollen is released prior to anthesis in a very close proximity to the stigma (Carlson and Lersten 1987) and that this ensures autonomous fertilisation within a day of anthesis (Yiu and Yeh 1989). Therefore, in *G. clandestina* the value of cleistogamy as a reliable mode of seed production under pollinator limitation may be restricted or may vary among populations.

In this chapter I examine whether the two postulated advantages of cleistogamy over chasmogamy, the greater economy and reliability of seed production, determine the function of cleistogamy in *G. clandestina* and whether the *G. clandestina* individuals employ cleistogamy in a reproductive strategy similar to other perennials. In order to determine if cleistogamy functions as a cost saving mode of reproduction under resource limitation, the following questions were addressed: 1. How does the relative investment in cleistogamy change in relation to plant size?; 2. Is cleistogamy more cost efficient at flowering as well as fruiting compared to chasmogamy? In order to determine if cleistogamy functions as a more reliable means of seed production under pollinator limitation the following questions were addressed: 1. How does the investment in cleistogamy change in response to chasmogamous fruit set?; 2. Does morphology of pollination in CL and OP flowers support autonomous self-pollination in both flower types?; 3. Are OP flowers able to autonomously initiate fruit in the field and
is fruit initiation pollinator limited?; 4. Are CL flowers more reliable at producing seed and fruit than OP flowers?
Materials and Methods

Field experiments and observations were carried out over two reproductive seasons (1998-2001), each lasting from September to February (spring-summer). Four populations were included in the study: Aranda (AR) and Taylor (TA) represented the narrow leaf morphotype (NAR), and Brindabella 2 (B2) and Mt. Franklin (FR) represented the broad leaf morphotype (BRO) (for details see Table 1.1, Chapter 1).

Individual investment in cleistogamy and chasmogamy

Census of individual reproductive patterns

The individual reproductive output was censused fortnightly over the two reproductive seasons. A count of open OP flowers, new CL flowers, new mature fruit of both types was conducted on each individual to obtain the flowering and fruit maturation phenology. All mature fruit were collected and stored separately in paper bags. All of the newly counted CL flowers were marked with liquid paper on the adjacent stem. The sum of the counts of CL and OP fruit and CL flowers over the census period provided their total number per individual. For the total number of OP flowers, a single census was conducted at the end of the OP flowering season (late December), when all of the bracts on the persistent OP racemes were counted. All of the plants that were substantially damaged (e.g. by herbivores) or died before fruit set were excluded from the analyses.

Plant size was measured as a number of leaves after the first vegetative growth spurt at a start of each reproductive season in spring (September). Such measure of plant size was considered to be most likely to reflect resource status in this perennial because it represented vegetative regeneration after dormant period in winter and relied on the stored resources (Pugliese and Kozlowski 1990).

Sampling in relation to plant size differed between the two years based on the analysis of reproductive patterns in the first year. Initially 20 plants were selected within each of three plant size classes: < 50 leaves, 50-150 leaves and >150 leaves, in order to ensure that all plant sizes were well represented. _G. clandestina_ plants showed all of the four possible combinations of cleistogamy and chasmogamy at flowering: no flowers (n=9), only CL flowers (n=39), CL and OP flowers (n=113) and only OP flowers (n=26) (Fig. 38).
2.1). There was a switch in reproductive strategy that was apparently related to plant size. The smallest plants either did not reproduce or they produced only CL flowers. In larger plants OP flowers were produced first, in most cases followed by CL flowers. Plants grown from seeds in the glasshouse experienced a similar switch in reproductive strategy as they grew and matured – their first reproduction was through cleistogamy and their next reproductive period started with OP and was followed by CL flowering. Thus, based on glasshouse observations, the CL-only flowering appeared to be confined to an early or juvenile developmental stage. This observation is also consistent with flowering patterns described in seedlings of the related G. tomentella (Kenworthy et al. 1989). This juvenile flowering or reproductive strategy was shown to be confined to the smallest plants in the field (see the section below and the results), likely to also be the youngest. In order to separate life-stage or age and plant size or resource effects on cleistogamy, in further sampling and analysis plants without flowers and with CL flowers only were grouped into a juvenile reproductive category, while those that produced OP flowers first were considered mature plants. Only plants exhibiting the mature reproductive strategy were sampled in the second year. The plants that were damaged or died in the first year were replaced with similar-sized plants. Since the majority of plants within populations were replaced (50-90%), for the analysis it was assumed that the sampling was independent between years.

**Plant size and relative investment in cleistogamy and chasmogamy**

The apparent threshold for OP reproduction was first explored within the four populations in the first year (1999). In the presence of a threshold for OP flowering, the plants bearing only CL flowers were expected to be significantly smaller than those that produced some OP flowers, and little or no overlap in sizes of plants with these two reproductive strategies was expected. The difference in mean plant size (leaf number) between mature and juvenile plants was tested using a mixed model nested analysis of variance (ANOVA) with reproductive category (mature and juvenile) as a fixed factor, and population (AR, TA, FR, B2) and plants nested within populations as random factors. Plant size was ln-transformed. Additionally, the plant size ranges between mature and juvenile plants were compared to ascertain whether there was a threshold size at which the switch from mature and juvenile reproductive strategy occurred.
Next, the hierarchy of investment in cleistogamy and chasmogamy was explored in the adult plants using data from the four populations and both years. The dependence of relative investment into OP and CL flowers and fruits on plant size, year and population tested, using a mixed model nested analysis of covariance (ANCOVA) (Quinn and Keough 2002). The proportion of CL flowers and fruits in the total reproductive output were used as response variables. There were two between plot factors (population and year) with plants as random plots nested within population and year. Population (FR, AR, TA, B2) and year (1999 and 2000) were considered random factors. Plant size (leaf number) was included as a covariate in the analysis. If plant size were important for the relative expression of cleistogamy and chasmogamy, a significant effect of the covariate would be expected. The regression slope would be positive if proportion of cleistogamy increased with size, and negative if it diminished.

In order to determine how the total fruit and flower production responded to plant size, further analyses using the ANCOVA were conducted. The ANCOVA model as described above was used, but total number of flowers and fruits were included as response variables. When the previous analysis of plant size on proportion of cleistogamy was significant, then separate ANCOVAs were computed for each reproductive mode. Otherwise, a single analysis on total reproductive output (CL plus OP) was performed.

All analyses were conducted using JMP v. 3.0.1 (SAS Institute Inc.), with α=0.05. Where multiple tests were required, a Bonferroni procedure was applied to adjust α-levels (0.05/n, where n is the number of comparisons). Where required in the mixed model analyses, quasi-F-ratios were calculated to test for significance according to Winer et al. (1991). Denominator degrees of freedom were adjusted using Satterthwaite, where the mean square error is a composite of more than one source (Winer et al. 1991) for the calculation of p, taking into account the unbalanced sampling. Before performing the ANCOVAs a test for interactions between the covariate and the other factors in the model was applied to check for the assumption of homogeneity of within group regression slopes. First the full model was fitted. This model was reduced in a stepwise process starting with the exclusion of the three-way and then two-way interactions with the covariate (all at p>0.05). Interactions between main factors that were not significant were also excluded from the models (all at p>0.2). Transformations
of data were performed to improve homogeneity of variance as determined by examination of the plots of residuals versus predicted values and the normality of data as examined by plotting the distribution of residuals from the models. Proportions were arcsine-square root transformed and counts were square-root transformed unless otherwise indicated. Throughout the results means±SE are presented.

**OP fruit set and investment in cleistogamy**

In order to test for the dependence of CL fruit production (number of CL fruit) on the failure of OP flowers to set fruit, mixed model nested ANCOVAs were used with two covariates: the probability of OP fruit maturation (number of mature OP fruit/number of OP flowers on a plant) and plant size (number of leaves). There were two between plot factors (population and year), with plants as random plots nested within population and year. Population (FR, AR, TA, B2) and year (1999 and 2000) were considered random factors. The probability of CL fruit set (number of CL fruit/number of CL flowers) and CL flower production (number of CL flowers) were also analysed in order to determine the mechanism of regulation of CL fruit production. Before executing the three ANCOVAs, the lack of correlation between the two covariates was confirmed using nested mixed model ANCOVAs with plant size as a covariate and probability of OP fruit set as a response variable. Also, homogeneity of within group regression slopes for each covariate was checked separately.

**Relative efficiencies of CL and OP seed production**

The relative efficiencies of reproduction through the two modes are best compared using plants producing the two types of flowers so that to take account of the interplant variation in the analysis. Therefore, plants producing both types of flowers were used in *G. clandestina* for the following comparisons of the relative efficiencies of reproduction through cleistogamy and chasmogamy in the four populations.

**Flowering and seed cost**

The CL and OP flowering costs were assessed based on dry weight of flowers and their supporting structures. Both types of flower were collected in the first reproductive season from each of 10 plants not included in the reproductive census in the four populations. For the OP mode, five fully open flowers and five racemes, and for the CL mode, five CL buds were collected per plant. Flowers and racemes were oven-dried at 80°C for at least 48 hours and weighed in the lots of five to the nearest 0.01 mg. The
average dry weight per flower was calculated by dividing the total dry weight by the
number of samples in the weighed batch. The average raceme weight per OP flower was
calculated as the average per raceme weight divided by the average number of flowers
per raceme in the weighed batch. The OP flowering cost was estimated as the sum of
the average dry weight of raceme per OP flower and the average weight of OP flower.
The CL flowering cost was the dry weight per CL flower.

The cost of seeds was assessed as seed weight. Seeds were sampled from individuals
that produced at least three mature fruit of both types in the first reproductive season.
The sample included both censused and uncensused plants. All of the seeds for each
fruit type were weighed in bulk to the nearest 0.01 mg and the average weight per seed
calculated.

Reproductive success
The comparisons between CL and OP flowers were conducted separately for the fruit
set and seed set component of fecundity.

The sampling for a comparison of seed set per fruit corresponded to that for seed weight
comparisons above. For each fruit the number of mature seeds and number of ovules
which failed to develop into seeds were recorded. The ‘failed’ ovules were identified at
x10-magnification under a dissecting microscope. The probability of seed maturation
within fruit was calculated as a ratio of mature seeds to total number of ovules (total
ovules – the sum of failed ovules and mature seed) within each fruit. The mean seed
number, total ovule number and probability of seed maturation were calculated across
fruit of the two types within each plant.

For the comparison of fruit set probability between CL and OP flowers, the data from
the two seasons of the reproductive census was used. Only plants with both flower types
were included. The probability of fruit maturation was expressed as a ratio of fruit to
flowers for a given flower type.

A split-plot mixed model ANOVA (Quinn and Keough 2002) was used to compare CL
and OP reproduction in terms of the flowering cost, seed costs, probability of fruit
maturation, seed number per fruit, ovule number per fruit and probability of seed
maturation within fruit. For each character the model included population (AR, TA, FR, B2) as a between plot factor, and flower or fruit type as a within plot factor (CL and OP), with plants as random plots nested within population. The analysis of probability of fruit maturation also included year as the between plot factor. Population and year were considered random factors and flower or fruit type the fixed factor in the analysis. When there was a significant interaction between the flower or fruit type and other factors, further analysis was conducted to test for the significant difference between cleistogamy and chasmogamy within each factor level using a priori non-orthogonal contrasts.

**Pollinator limitation in OP flowers**

The following study to assess pollinator limitation on fruit set in OP flowers was conducted in the second year in three out of four populations (NAR: AR and TA; BRO: FR).

The experiments were not aimed at determining the precise extent of overall pollinator limitation but rather its presence at fruit initiation in chasmogamous flowers of this species. The assessment of potential for pollinator limitation in OP flowers was important in *G. clandestina* because previous observations indicated that OP flowers may circumvent this limitation due to delayed self-pollination (pers. obs., Schoen and Brown, 1991). As these tests did not need to remove all factors that may confound the extent of pollinator limitation (e.g. inbreeding depression or interplay with resource limitation). Thus any subsequent discussion of pollinator limitation is within the restricted meaning of requirement for pollinator presence for fruit initiation and insufficient pollinator service to maximize fruit initiation in the field.

**Morphology and pollination in CL and OP flowers**

For the morphological study of the differences in pollination mechanism between CL and OP flowers, fresh flowers at various developmental stages were collected in the field. The sample of open chasmogamous flowers included three categories: flowers exposed to pollinators in the field, flowers not visited by pollinators, and flowers subject to hand-tripping in the lab (see the description of the pollen supplemented treatment below). The flowers were dissected to remove sepals and petals (perianth) and expose the gynoecium and androecium without disturbing their relative position. The reproductive organs were examined under a dissection microscope at x5 magnification.
and under a scanning electron microscope (SEM), which was also used to produce micrographs.

For viewing under the SEM, cryo preparation techniques were used on fresh material to minimise the specimen artefacts. The specimens were prepared using standard techniques in the Oxford CT1500 Cryo Preparation System, coated with 10nm layer of gold and viewed on a cryo stage of a Cambridge Instruments S360 scanning electron microscope at c. -165°C. In order to ensure high resolution and depth of field, the following settings were used: 30µm diameter final aperture, a working distance of approximately 18 mm, electron beam current of 80 pA and an accelerating voltage of 15kV.

**Efficiency of fruit initiation in OP flowers**

To test whether OP flowers can autonomously self-fertilise, three treatments were applied to each of five plants within three populations: 1. Pollinators excluded and flowers unmanipulated (Autonomous), 2. Pollinators excluded, flowers self-pollinated by hand (Supplemented), 3. Flowers exposed to natural pollination (Open pollination). Ten inflorescences were randomly assigned to each treatment. Pollinators were excluded by enclosing part of each plant within a nylon mesh bag (Fig. 2.2). Hand self-pollination was achieved by gently squeezing the keel of the flower in order to push the pollen released within the flower over the stigma (hand-tripping). This treatment was chosen as an efficient alternative to the more conventional hand cross-pollination for pollen supplementation as the flowering phenology in *G. clandestina* causes very high risk of flower damage and low efficiency of hand cross-pollination in the study species (A.D.H. Brown pers.comm.). All of the inflorescences with open flowers were removed from the plant before the experiment commenced in order to ensure that only inflorescences where all of the flowers opened during the experiment were used. Because flowers within inflorescences open sequentially, the hand self-pollinations were performed daily on the freshly opened flowers within the bags. New initiated fruits within each inflorescence were counted daily during the pollination period and every second day afterwards. The total number of flowers in each inflorescence was also recorded. Fruit initiation rate was estimated as the proportion of flowers that initiated fruit within each inflorescence. Mean initiation rate for each treatment within a plant was then calculated for the analyses.
The differences in fruit initiation rates among all the three treatments were tested using a split-plot mixed model ANOVA with plants as random plots nested within populations, population as a random between plot factor (AR, TA, FR) and treatment (three levels: Autonomous, Supplemented and Open pollination) as a fixed factor. Mean fruit initiation rates were arcsin(sqrt)-transformed for this analysis. When treatment factor was found to be significant, two a priori contrasts between pairs of treatments were performed: 1. Autonomous vs Open pollination. No difference between the two treatments was expected if natural pollination in the field was entirely due to autonomous selfing or if pollinators achieved similar efficiency of pollination as autonomous selfing; 2. Open pollination vs Supplemented. Significantly greater fruit initiation would be expected in Supplemented treatment if pollinator limitation on fruit initiation was present in the field.
Results

Individual investment in cleistogamy and chasmogamy

Based on the first year of the census, c. 19% (9/48) of the juvenile plants failed to produce any flowers. Of those that flowered c. 35% (14/39) failed to set fruit. On average, reproductive juveniles produced 10±1.3 flowers and 3 ±0.7 fruits. In contrast to juveniles, all except one of the 216 mature individuals produced fruit. On average each plant produced 29±2 CL and 244±17 OP flowers, and 7±1 CL and 18±2 OP fruit. The proportion of CL flowers per plant was the highest in the AR population (36%), intermediate in TA (13%) and B2 (10%), and the lowest in FR population (2%). The proportion of the CL fruit was at least twice as high as the proportion of CL flowers in three of the populations: 62% in AR, 36% in TA and 30% in B2. The increase in proportion of cleistogamy per plant at fruiting indicated the greater efficacy of cleistogamous flowers at fruit production in those populations. The proportion of CL flowers and fruit remained very low at 3% in FR population.

OP flowering occurs before CL flowering

The phenology of flower and mature fruit production by mature plants is depicted in Fig.2.3. Plants in the four populations commenced flowering at different times. Taylor (TA) population was first to start flowering at the end of October, followed by AR and B2 in mid-October and FR only starting in early November. Interestingly, in FR population (at the highest elevation), where individuals showed the lowest proportion of cleistogamy, the reproductive period was also much shorter than in the other three populations (less than four vs. six months). In all of the populations OP flowers were produced first, with CL flowers starting to appear only at the end of the OP flowering period. Across the populations most of the CL flowers were produced after the peak in OP flowering and majority of OP flowers opened (B2, FR) (Fig. 2.3C,D) or after OP flowering had finished (TA, AR) (Fig. 2.3A,B). There was a 1.5- to 2.5-month separation between the peak of flowering for chasmogamy and cleistogamy. Patterns in fruit maturation reflected the patterns in the sequence of flowering. This temporal separation between chasmogamy and cleistogamy created a possibility that individual chasmogamous fruit set may affect the level/expression of cleistogamy.
Plant size determines the magnitude of reproductive output and relative investment in cleistogamy

The plants with juvenile reproductive strategy (juvenile plants) were significantly smaller than plants with mature reproductive strategy (mature plants) across all of the populations (Table 2.1). There was almost no overlap in plant size between the two groups: the size of juvenile plants varied from 3 to 35 leaves while mature plants had 30 to 510 leaves. Only 2.5% (1/120) of the mature plants had 30 to 35 leaves. Therefore, a plant size-dependent switch in reproductive strategy (from juvenile to mature reproductive strategy) occurred. Plants with less than about 35 leaves employed only cleistogamous flowers, and larger plants switched to chasmogamous flowering, or chasmogamous flowering followed by cleistogamy within the same plant.

Reproductive strategy was also found to be significantly dependent on plant size within the juvenile and mature plant categories. Juvenile plants that produced flowers were significantly larger than those without flowers (20 ± 1 vs. 13 ± 1; one-way ANOVA F1,47=5.422, p=0.024), indicating that an increase in plant size increased the probability of flowering and that plant size is a good indicator of resources available for reproduction in juveniles.

The analysis of the investment in cleistogamy in relation to plant size in mature plants (mixed model nested ANCOVAs) is presented in Table 2.2. Significant factors responsible for the variation in the proportion of cleistogamy were population and year for fruit and population for flowers (Table 2.2A,D). More importantly, the response to plant size was different at the stage of flower and fruit production. At flowering, plant size had a significant negative effect on the proportion of cleistogamy in the total reproductive output (b=-1.31±0.42, p=0.003; Table 2.2A). On the other hand, the proportion of CL fruit was not significantly correlated with plant size (b=-0.82±0.80; Table 2.2D). Further analysis revealed that the number of both flower types increased significantly with plant size (Table 2.2B,C). Therefore, the lower rate of increase in CL compared with OP flowers was responsible for the decrease in the proportion of cleistogamy with plant size (OP vs CL flowers: b=1.23±0.14 vs b=0.15±0.07; Fig. 2.4A). The CL and OP fruit production also showed a significant increase with plant size (Table 2.2E, Fig. 2.4B). These results indicated a significant difference between
cleistogamy and chasmogamy in the slopes of the relationship between plant size and the number of flowers, but no significant difference for fruit.

The significant positive relationship between plant fecundity, in terms of flower and fruit number, and plant size in mature plants confirmed that plant size is a good indicator of resources available for reproduction in *G. clandestina*. Both, the size threshold for OP reproduction in juveniles and the decrease in proportion of CL flowers with increasing size in mature plants, were consistent with the hypothesis that in *G. clandestina* cleistogamy is favoured in smaller plants where reproduction is resource limited. The pattern in *G. clandestina* contrasted with observations in other perennials where the individual investment in OP reproduction diminished with plant size.

**CL fruit production responds to plant size and OP fruit set**

Table 2.3 summarises the analysis of plant size and OP fruit set as determinants of CL reproductive output. Covariate analysis indicated that OP fruit set was not related to plant size, but only varied significantly with population and year (Table 2.3A). While the production of CL fruits responded positively to plant size (b=0.20±0.04), it was negatively affected by an increased OP fruit set (b=-0.06±0.02) (Table 2.3B). The positive effect of plant size on investment in cleistogamy was evident in both the significant effect on CL flower number and the probability of CL fruit set (Table 2.3C,D). The effect of the OP fruit set was not significant for either CL flower production or CL fruit set, but both parameters showed a negative trend (b=-0.47±0.29 and b=-0.04±0.08, respectively)(Table 2.3C,D). Therefore, it is likely that OP fruit set influenced production of CL fruit via its effect on both flower number and fruit set. The positive effect of low OP fruit set on CL fruit production was consistent with the hypothesis that cleistogamy functions as a more reliable fruit source when fruit set in OP flowers is limited. This pattern was similar to those found in other perennials.

**Cleistogamy as a more efficient source of seed**

Table 2.4 and Fig. 2.5 present the results of the analysis comparing reproductive parameters for cleistogamy and chasmogamy in *G. clandestina*, including flower cost, seed cost, number of ovules per flower, seeds per mature fruit, seed set within fruit, fruit set per flower. With the exception of seed number per fruit, the values of these parameters differed between the two flower types for at least some populations.
Resource savings on CL seed production

The analysis of the cost of flowering expressed as dry mass per flower showed a significant population by flower type interaction (F3,36=4.09, p=0.01). The non-orthogonal comparisons between the two flower types demonstrated that the difference between cleistogamy and chasmogamy was significant within each population (Table 2.4A). The mean flowering cost was 8- to 14-fold higher for chasmogamy than cleistogamy within each population (Fig. 2.5A), with an average of an 8-fold difference across populations (0.3 ± 0.05 mg vs. 2.3 ± 0.05 mg). Even when rachis costs for chasmogamy were excluded, the difference was 7-fold.

In contrast to the flowering costs, the dry mass cost of a seed produced via cleistogamy significantly exceeded that for OP seed (fruit type: F1,76=35.94, p<0.0001), with the difference consistent among the four populations (Table 2.4B, Fig 2.5B). Averaged across all populations, the amount of dry matter in seed was about 1.08-fold greater for CL than OP seed (43.8± 0.4 mg vs. 40.6 ± 0.4 mg, Fig. 2.5B).

There was a difference between cleistogamy and chasmogamy in ovule number per flower (Fig. 2.5C), with OP flowers containing slightly more ovules than CL ones on average across populations (9 and 8 respectively). Although the difference in ovule number between CL and OP fruit was present in all populations (Fig. 2.5C), in the ANOVA the population x flower type interaction was significant and the non-orthogonal contrasts showed that in two of the populations (AR and TA) CL and OP flowers did not significantly differ in ovule number (Table 2.4C). Despite the significant difference in ovule number between CL and OP flowers, the number of seeds did not differ significantly between CL and OP fruit in any of the populations (Fig. 2.5D; Table 2.4D), and each fruit contained on average 6 seeds.

As expected from the difference in flower size, the CL flowers required significantly less dry matter and were therefore likely to provide resource savings over the OP flowers in G. clandestina. However, due to the significantly greater seed mass, CL seed was more costly compared to OP seed. As the seed number was similar in the two fruit types, this caused CL fruit to be more costly than OP fruit and on the per fruit basis this cost exceeded the savings made at the flowering stage in CL reproduction in G. clandestina.
Greater reliability of seed production in CL flowers

Ovules had a greater probability of developing into mature seeds within CL fruit than OP fruit across all of the populations. On average 87% and 78% of ovules turned into seeds in CL and OP fruit respectively, making CL fruit 1.07- to 1.2-fold more reliable at producing seeds than OP fruit (Fig. 2.5E). The ANOVA showed a significant fruit type effect but population x fruit type interaction was also significant (Table 2.4E). The non-orthogonal contrasts between cleistogamy and chasmogamy within each population confirmed the significant difference in all of the populations except TA, where the difference between cleistogamy and chasmogamy was the smallest (Table 2.4E, Fig. 2.5E). This lack of significant difference for TA clearly resulted from the lack of statistical power to detect the difference of the smaller magnitude. Thus it was concluded that the probability of seed set per fruit was greater for cleistogamy than chasmogamy, but that the magnitude of the difference varied among the populations.

Similarly to seed maturation within fruits, CL flowers also had a higher probability of producing fruit than OP flowers in three out of the four populations (Fig. 2.5G). In the AR, TA and B2 populations only about 4% of OP flowers produced fruits, while 29%, 18% and 16% of CL flowers matured fruit respectively. In these three populations CL flowers were 4- to 7-fold more successful than OP flowers at producing fruit. On the other hand, in the FR population both types of flowers set fruit at a similar rate of 11% (Fig. 2.5G). The results of ANOVA showed that the effect of flower type differed significantly among populations, and the paired contrasts between cleistogamy and chasmogamy confirmed that the only non-significant difference between the flower types was within the FR population (Table 2.4G).

Overall in *G. clandestina* cleistogamous flowers were more reliable at producing seeds and fruit than chasmogamous flowers, suggesting that chasmogamous flowers lacked a reliable pollination mechanism. The Mt. Franklin population of the broad leaf morphotype (BRO) located at high elevation was an exception, where equivalent fruit set in the two types of flowers indicated that they were similarly limited at fruit production.
Inefficient pollination mechanism in OP flowers

CL flowers have morphological adaptations to ensure pollination

The developmental sequence of reproductive organs and pollination within CL and OP flowers is presented in Fig. 2.6. The two flower types had the same number of anthers whose filaments were fused to form a sheath around the ovary with one free posterior anther facing the stigma. At an advanced bud stage, the androecium and gynoecium were similar in both flower types with the style extending past the whorl of anthers but bent to bring the stigma close to the anthers (Fig. 2.6A,E). While in the CL flower the stigma remained facing the anther surface and remained in its close proximity (Fig. 2.6A), in the OP flower style straightened and extended with the stigma facing away from the anthers (Fig. 2.6A,B). The stigma in both types of flower was covered with a smooth membrane.

In the OP flowers anthers proceeded to dehisce and release pollen shortly before or at anthesis (Fig. 2.6B). In the majority of OP flowers that were not exposed to pollinators, the mass of pollen was found behind and below the stigma, supported by the numerous stylar trichomes, while the stigmatic surface remained free of pollen or with few pollen grains but with the membrane intact (Fig. 2.6B). In contrast, flowers which were collected from the field after being exposed to natural pollination and in those manipulated by hand the membrane was found to be ruptured, with numerous pollen grains germinating on the stigmatic surface (Fig. 2.6C).

In CL flowers anthers remained undehisced throughout flower development, with the style curved and stigmatic surface touching the proximal anther pair (Fig. 2.6E,F). The pair of anthers became attached to the stigma and was torn off their filaments as the pistil started to elongate post-fertilisation (Fig. 2.6F,G).

This shows that in G. clandestina the morphology of CL flowers provided a mechanism for reliable self-fertilisation. In chasmogamous flowers the early pollen release in proximity of the stigma also created the possibility of autonomous self-pollination, even prior to anthesis. In contrast to CL flowers, however, the self-pollen in OP flowers did not appear to reach the stigma or germinate when flower remained undisturbed, thus preventing the premature self-fertilisation.
OP fruit initiation is limited by pollinator presence and efficiency

The three pollination treatments applied within *G. clandestina* plants resulted in substantially different fruit initiation rates per raceme (no. initiated fruit/no. flowers) across all of the three populations (AR, TA, FR) (Fig. 2.7). Fruit initiation rates were the lowest in the Autonomous, intermediate in the Open pollination and the highest for the Supplemented treatment, with the mean of 3%, 24% and 67% of flowers initiating fruit, respectively.

A split-plot mixed ANOVA showed that the effect of pollination treatment on fruit initiation rate was significant (Table 2.5). The contrast test showed that racemes that were open to the usual pollinator activity (Open pollination) had significantly higher fruit initiation rates than those which relied on ability of OP flowers for autonomous self-pollination (Autonomous) (Table 2.5). In turn, fruit initiation rate in racemes where pollen deposition was ensured by hand tripping (Supplemented) was significantly greater than that in open pollinated ones (Open pollination) (Table 2.5). Importantly, the effects of the treatments did not significantly differ among the populations, as indicated by lack of significant population by treatment effect. The mean initiation rate did not differ among populations for either Autonomous or Open pollination treatment (One-way ANOVA: $F_{2,12}=1.4$, $p=0.28$; $F_{2,12}=2.6$, $p=0.11$, respectively), indicating that OP flowers were similarly limited by inability to autonomously pollinate and effectiveness of pollinators in each population.

Thus in *G. clandestina*, contrary to the anecdotal evidence, chasmogamous flowers were not able to initiate fruit autonomously at a substantial rate in the field and were therefore likely to be limited by pollinators. Natural pollinator activity caused a significant increase in initiation over the autonomous rates in the field. The pollinator limitation in the field was confirmed by the further significant increase in the fruit initiation by pollen supplementation. Importantly, there was no evidence for any inter-population differences in the effects on initiation rates or the limitation of the rates of initiation by ability for autonomous pollination or quality of natural pollinator service.
Discussion

The comparisons between chasmogamous and cleistogamous flowers of *G. clandestina* yielded results unexpected in this species in terms of both, the relative cost effectiveness and the reliability of seed production. Contrary to predictions based on previous observations in this species, chasmogamous flowers were not capable of substantial autonomous selfing in the field and overall tended to be significantly less reliable at producing seed than cleistogamous flowers. At the same time, the relative economy of cleistogamy was likely to be diminished in *G. clandestina*, despite biomass savings at flowering stage, because of significantly greater cleistogamous seed biomass. This was surprising in an herbaceous species with aerial cleistogamy.

In terms of the relative investment in cleistogamous and chasmogamous reproduction, *G. clandestina* plants showed variation in relation to plant condition that is generally expected in CLOP species. Specifically, cleistogamy was the favoured mode of reproduction in smaller plants and in plants with lower chasmogamous fruit set, which was consistent with the postulates of greater economy and reliability of cleistogamous seed production, respectively. However, only the pattern of increase in cleistogamy in response to chasmogamous failure was consistent with those observed in other perennials.

There was little evidence of differences in the reproductive strategy between the morphotypes and elevations. One important exception was the high elevation BRO population (FR) where mature plants combined the equivalent cleistogamous and chasmogamous fruit set with an extremely low level of cleistogamy.

**Function of cleistogamy as a more reliable source of seed**

This study revealed that cleistogamy tended to be significantly more reliable at seed production than chasmogamy in *G. clandestina* at two levels. Cleistogamous flowers were 4-7 times more likely to set fruit than chasmogamous flowers in three out of four populations. Moreover, in all of the populations seed set within cleistogamous fruits was also significantly greater, exceeding by 1.07 to 1.2-fold seed set in chasmogamous fruit. By combining these two measures for each population, an estimate is derived that shows
cleistogamy was on average 1.07-8.4 times more reliable source of seeds than chasmogamy in *G. clandestina*.

**Pollinator limitation in chasmogamous and cleistogamous flowers**

An important reason for the usual low reliability of chasmogamy is pollinator limitation. Experimental manipulation confirmed that autonomous pollen deposition in chasmogamous flowers was unlikely to prevent pollinator limitation in *G. clandestina*. This was evident in that only 3% of flowers in the pollinator exclusion treatment initiated fruit in the three study populations. Moreover, pollen supplementation produced a two-fold increase in fruit initiation rate over that resulting from natural pollination (24% vs. 67%). This demonstrated that natural levels and efficiency of pollinators in the field limited OP fruit production. Together these results suggest that pollinator limitation in chasmogamous flowers is an important factor limiting their female fertility component across populations of *G. clandestina*.

The ineffective autonomous fruit initiation in OP flowers was surprising for two reasons. Firstly, it was inconsistent with previous observations in the glasshouse and mating patterns in the field that lead to the suggestion that delayed selfing may occur in OP flowers of *G. clandestina* (pers.obs.; Schoen and Brown, 1991). Secondly, the morphological observations showed that pollen was released relatively close to the stigmatic surface in OP flowers in this species. Such pre-anthesis pollen release is well documented to result in high rates of premature autonomous self-fertilisation (>90%) in OP flowers of the relative *G. max* (Summerfield and Roberts 1985; Carlson and Lersten 1987; Yiu and Yeh 1989). On the contrary the results in *G. clandestina* suggest that its OP flowers have a mechanism preventing autonomous selfing in the field.

In some other species with papilionoid (pea-like) flowers it has been shown that the stigmatic cuticle acts to prevent autonomous selfing by preventing pollen germination (Kreitner and Sorensen 1984; Lord and Heslop-Harrison 1984). Because of the papilionoid structure of OP flower, pollinator visit results in activation of a mechanism that facilitates the movement of pollen and stigma into contact with the insect (tripping) (Arroyo 1981). Tripping also ruptures the stigmatic cuticle to place pollen in contact with the stigmatic secretion and allows it to germinate (Kreitner and Sorensen 1984; Lord and Heslop-Harrison 1984). Morphological observations in *G. clandestina* in this
study revealed the presence of the stigmatic cuticle, and the lack of germination in pollen grains on the cuticle of unvisited flowers. The high chasmogamous fruit initiation rates of around 70% after pollen supplementation via hand-tripping also suggest that a similar action may be required for effective pollination of OP flowers in *G. clandestina*.

The stigmatic cuticle-based mechanism for prevention of selfing could also explain the apparent inconsistency between this study and previous observations regarding the ability for autonomous OP fruit set in *G. clandestina*. Studies in other species show that the effectiveness of stigmatic cuticle in prevention of premature selfing can be affected by environmental conditions (Brink and Cooper 1936; Arroyo 1981; Kreitner and Sorensen 1984). For example, high turgor pressure promotes autonomous cuticle rupture and high levels of autonomous fruit set (Brink and Cooper 1936). It is likely that good water supply under glasshouse conditions caused such effect in *G. clandestina* leading to high levels of autonomous fruit set in contrast to the normal conditions in the field. Extreme environmental conditions in the field could cause similar increase in effectiveness of autonomous fruit set in *G. clandestina*. However, the overall high stability of this mechanism in the field is suggested by the consistently low autonomous OP fruit initiation across several populations at a range of elevations and the two morphotypes.

In contrast to OP flowers, CL flowers in *G. clandestina* showed morphological modifications of reproductive organs consistent with adaptation for efficient autonomous pollination. In mature cleistogamous flowers styles were bent to allow the stigmatic surface to remain in contact with the most proximal anther pair, and no anther dehiscence was observed. This is one of the modifications observed in cleistogamous flowers of other CLOP species with zygomorphic flowers and it is considered one of the most effective mechanisms of autonomous fertilisation in cleistogamous flowers (reviewed in Lord 1981). It is because ontogenetic studies in those species show that in such cleistogamous flowers self-fertilisation is assured by germination of pollen within the undehisced anther, followed by the penetration of the anther wall and stigmatic surface by pollen tubes in order to reach the ovaries within the style (Lord and Hill 1987).
Differences among populations

Results of a previous study suggested that BRO populations or populations at a higher elevation in *G. clandestina* had greater ability for autonomous selfing or were more pollinator limited in OP flowers than lower elevation and NAR populations. Neither of these suggestions was supported by this study as experimental manipulation revealed that autonomous and natural fruit initiation rates did not significantly differ among the three study populations representing a range of elevations and the two morphotypes. Thus pollinator limitation was likely to equally limit the reliability of seed production in OP flowers across the elevations and morphotypes.

It is therefore surprising that in contrast to other populations, in the FR population (BRO morphotype, highest elevation) the two types of flowers produced fruit with equal probability of around 11%. As CL flowers were not less effective at initiating fruit and OP flowers were not less pollinator limited in his population than in the others, it is likely that some post-fertilisation factor absent in other populations caused the CL fruit set to drop to OP fruit set level. While the extent of post-fertilisation limit on fruit set was not directly assessed in this study, the loss of fruit and seed of both types was frequently observed in the field. Moreover, the extent of chasmogamous fruit loss reaches at least 30%, as estimated from the difference between natural fruit initiation rate in the bagging experiment and the chasmogamous fruit set in the same populations (Mean population fruit initiation vs fruit set: 32%±1.4% vs. 6%±1.8%). The observed mechanisms of fruit loss included predation (parrots, snails) and apparent abortion. Some of the fruit abortion was likely caused by predation on seeds within fruit (mainly from bruchid beetle and fly larvae – pers.obs.). Fruit abortion also provides means of reproductive output regulation in response to resource or weather limitation in the related crop *G. max* (Summerfield and Roberts 1985). It is likely that *G. clandestina* plants similarly use abortion for such regulation. Irrespective of the mechanism, the result in FR population may indicate that post-pollination limits on fruit set can decrease the relative reliability of cleistogamous seed production arising from greater pollination certainty in some populations of *G. clandestina*.

These results suggest that pollinator limitation in chasmogamous flowers is an important factor limiting the female fecundity and that cleistogamous flowers are more
reliable mode of seed production. Therefore, it would be expected that *G. clandestina* plants employ cleistogamy when chasmogamous seed production is pollinator limited.

**Employment of cleistogamy under pollinator limitation**

**Employment of cleistogamy under limitation to OP fruit set in *G. clandestina***

The study of individual reproductive patterns found that mature *G. clandestina* plants with lower chasmogamous fruit set produced higher numbers of cleistogamous fruit. Together with the greater reliability of cleistogamous seed production this suggests that in mature plants cleistogamy is employed in a conditional strategy ensuring seed production when chasmogamous fruit set is limited, either at the pollination or the post-pollination stage. As chasmogamous fruit set appeared to be limited at both pollination and post-pollination stage in *G. clandestina* such response is unlikely to be specific to the traditionally postulated pollinator limitation of OP flowers, but rather may represent a more generic response to limitations on OP fruit set. This pattern is broadly consistent with the traditional postulate that cleistogamy functions as a reliable seed source to ensure female fecundity.

This conditional strategy is undoubtedly facilitated by the sequential production of the two types of flowers in mature plants of *G. clandestina*. Observations in the field revealed that CL and OP flowers were clearly temporally and spatially disjunct, with OP flowers always arising first within a season and within plant branches. As a result the limitation on chasmogamous fruit set in any given year would become apparent while the investment in cleistogamy was taking place. This may provide the opportunity for feedback between chasmogamous and cleistogamous fruit production within a plant. The regulation of the late CL fruit production may be achieved when the early maturing OP fruit create a strong resource sink or produce phytohormones that either inhibit CL flower production or cause CL flower and fruit abortion. Such mechanisms of regulating allocation to early vs. late reproductive structures are common in plants in general (Stephenson 1981; Lee 1988; Bawa and Buckley 1989; Susko and Lovett-Doust 1999), and have been well documented in the related *G. max* (Summerfield and Roberts 1985). In *G. clandestina* the CL flower production and the CL fruit set tended to be negatively affected by the increase in OP fruit set, suggesting that both, inhibition of flower production and abortion, were involved in regulation of CL fruit production in response to OP reproductive success.
Differences among populations

There were no significant differences among the four populations in the way that mature *G. clandestina* plants employed cleistogamy in response to OP fruit limitation, nor did it differ between the two years of the study. This indicated that the conditional employment of cleistogamy under limitation to OP fruit set forms a stable part of the mature plant reproductive strategy in this species across the BRO and NAR morphotypes and a range of elevations.

However, in contrast to the other three populations, mature plants in Mt. Franklin population of a BRO morphotype at the highest elevation produced 5-15 times lower proportion of CL flowers and 10-20 times lower proportion of CL fruit. At such low levels, cleistogamy was unlikely to function for assurance of seed set under limitation in OP flowers in mature plants. At the same time plants from this population were observed to produce copious CL flowers after OP reproductive phase in the glasshouse. This suggests that environmental conditions in the field at the high elevation may restrict the production of CL flowers in mature plants in *G. clandestina*. In fact this is likely a result of the effect that higher elevation and the associated lower temperature has on reproductive phenology. This study showed that the plants in FR population started producing flowers 1.5 to 3 months later than the plants in lower elevation populations so that their reproductive period was shorter. The difference in elevation of 600-700 m at this latitude between FR and other populations corresponds to a difference in temperature of about 8 degrees at the same time of the year (Fig. 2.8). The temperature in combination with photoperiod determines the commencement of flowering in the annual *G. max* cultivars (Summerfield and Roberts 1985; Carlson and Lersten 1987), and in the related perennial species *G. tabacina* (Kenworthy *et al.* 1989). Thus the observed delay of the commencement of flowering at high elevation in *G. clandestina* is likely to be caused by the lower temperatures in spring. Because of the short reproductive period mature plants in this population have a limited opportunity to employ cleistogamy after OP flowering has finished, causing the levels of cleistogamy to be low. It follows that in *G. clandestina* populations located at high elevation the value of cleistogamy to provide for seed production under OP reproductive failure in mature plants is limited.
Comparison with other perennials

The regulation of CL fruit production in response to overall limits on OP fruit set in *G. clandestina* was similar to that observed in the two other perennial CLOP herbs, *Oxalis acetosella* and *Viola hirta* (Redbo-Torstensson and Berg 1995; Berg and Redbo-Torstensson 1998). In both of those species plants also increased investment in cleistogamy (bud production) in response to OP seed production failure (fertilisation failure). As the perennial *Glycine* has evolved in an environment distinct from that of *Oxalis* and *Viola*, this similarity is unlikely to be an outcome of strong habitat-specific ecological parallels. Instead the finding in *G. clandestina* supports the hypothesis that perennials employ cleistogamy in response to OP fruit set limitation in a unique conditional strategy.

Cleistogamy is thought to provide the fail-safe seed production in the face of unpredictable failure of chasmogamy in many CLOP species (Darwin 1897; Richards 1986; Masuda and Yahara 1992). However, in many CLOP species the character of temporal separation between cleistogamy and chasmogamy precludes the feedback between OP fruit set and CL flower production, and the regulation of CL output. The production of OP flowers before CL flowers within a reproductive season, common to *Viola, Oxalis* and *Glycine clandestina*, forms an essential pre-condition for such feedback and regulation of cleistogamy.

Some annuals also produce OP flowers prior to CL ones (Darwin 1897; Lord 1981; Schoen and Lloyd 1984). Why has the conditional CL strategy not been observed in those species? There may be two reasons for such absence. No studies on annuals report specific tests of this hypothesis and thus it may have avoided detection. It is more likely however that the well-known differences in resource allocation strategies based on life history differences make such regulation of CL output more advantageous in perennials. This is because resource allocation strategies in perennials are shaped by a trade off between the demands of current reproduction and future growth and reproduction in order to maximise the lifetime reproductive success (Lloyd 1980; Willson and Burley 1983; Stearns 1992). The downregulation of CL seed production in plants with greater OP fruit set would allow allocation of resources to storage or growth, ensuring future survival and reproductive output. In contrast, annuals are predicted to allocate resources to maximise reproductive output within one reproductive season (Harper 1977; Willson
1983), which would make diminishing the CL seed production in response to the OP fruit set a disadvantage.

**Function of cleistogamy as a more economical source of seed**

*Cost-efficiency of seed production in chasmogamous and cleistogamous flowers*

Comparison between cleistogamy and chasmogamy in *G. clandestina* revealed that on average across populations a cleistogamous flower required 9 times less biomass but a cleistogamous seed required 1.07 times more biomass compared to the same chasmogamous structures (see Table 2.6).

While the economy of cleistogamous flowers in *G. clandestina* was typical of cleistogamous species (Darwin 1897; Lord 1981; Schoen and Lloyd 1984), the greater biomass of cleistogamous compared to chasmogamous seeds and its consistency was a surprise. A difference in seed biomass of similar magnitude was observed in several other studies of species with aerial cleistogamy but in contrast to *G. clandestina* it was found to vary among populations or seasons (Weatherwax 1928; Cope 1966; Waller 1982; Clay 1983; Berg 2000b). In *G. clandestina* however the greater cleistogamous seed mass appears to constitute a stable element of the cleistogamous reproductive strategy.

Previously there was only evidence that such stable difference is an important part of reproductive strategy in species with pronounced seed dimorphism where weight differences reached a magnitude of 1.5-30-fold (Clay 1983; Schoen 1984; Cheplick 1987; Trapp and Hendrix 1988). In some of those species it has been shown that the seed weight differences are associated with the difference in the number of seeds developing at the leaf axil – the smaller chasmogamous seeds occur in greater numbers per axil than the larger cleistogamous seeds (*Triplasis purpurea*, Cheplick 1996; *Amphibromus scabrivalvis* and *Danthonia spicata*, Schnee and Waller 1986; Cheplick 1987; Cheplick and Clay 1989). The photosynthate supply from the adjacent leaf is known to limit the local seed size in plants (e.g. Maun and Cavers 1971; Vaughton and Ramsey 1997) and the trade off between seed number and seed size under limited resources is well recognised (Harper *et al.* 1970; Baskin and Baskin 1998). Thus the differences in packaging of seeds in flower axils between chasmogamy and cleistogamy in those species provides a mechanism for control of maternal biomass allocation to the
two seed types. A similar mechanism is likely at work in *G. clandestina* where both inflorescence types are subtended by a single leaf but chasmogamous inflorescences initiate up to 9 and mature up to 5 fruit compared to the maximum of one fruit in cleistogamous inflorescences. This is in contrast to other aerial CLOP plants where seed weight appears instead to primarily respond to the differences in resource limitation according to the position of fruit within a plant (e.g. *Impatiens capensis*, Waller 1982; *Mimulus nastusus*, Diaz and Macnair 1998).

In terms of reproductive biomass cleistogamous reproduction offered substantial savings over chasmogamous reproduction per flower. At the same time, the greater cleistogamous seed biomass combined with no significant difference in seed number between the two fruit types meant that a cleistogamous fruit was significantly more costly than a chasmogamous fruit. This contrast between the flowering and fruiting stage in terms of the relative biomass requirements for cleistogamy and chasmogamy suggested that, overall, cleistogamy may offer little or no saving in terms of biomass investment required for reproduction. As in *G. clandestina* there was no reason to expect such discrepancy in relative costs prior to this study, the overall biomass investment was not explicitly estimated. However, an approximation of the overall biomass saving per fruit can be derived within each population using the estimated flower biomass, seed biomass, seed number per fruit, and the proportion of failed flowers of cleistogamous and chasmogamous type (Table 2.6). According to this estimate despite the increased cost of a fruit, overall cleistogamy on average requires 1.07 fold less of reproductive biomass to produce a single fruit than does chasmogamy in *G. clandestina*. This estimate of the overall savings is within the range of estimates available for five other CLOP species, where cleistogamy required 1.07-3.2 times less biomass or energy to produce a mature fruit or seed (see Table 2.6).

The estimate of lower overall biomass required for production of cleistogamous seeds in *G. clandestina* is consistent with the hypothesis of the greater economy of cleistogamous as opposed to chasmogamous reproduction on the assumption that resources saved on flowering are reallocated to fruiting and that biomass is indicative of the limiting resources. Therefore *G. clandestina* plants should invest in cleistogamy in preference to chasmogamy under conditions where resources are limiting.
Employment of cleistogamy under resource limitation

Size dependent employment of chasmogamy in G. clandestina

This study revealed that in addition to regulation of cleistogamy in response to OP fruit set, G. clandestina plants increased the priority of investment in chasmogamy relative to cleistogamy in response to the increase in plant size. This was manifest in three patterns. Firstly, reproduction in the smallest plants was restricted to cleistogamy until they reached a size threshold of about 35 leaves. This juvenile-like reproductive strategy was then replaced in larger or mature plants by a strategy combining cleistogamy and chasmogamy. Secondly, the mature plants gave priority to OP reproduction by producing OP flowers before CL ones during the reproductive season. Finally, the proportion of CL flowers in those mature plants was found to gradually decreased with increase in their size. Overall, these patterns are consistent with the hypothesis that G. clandestina plants employ cleistogamy in a conditional strategy to ensure more economical seed production when plants are more resource limited, providing that plant size reflects resources available for reproduction.

Vegetative plant size is commonly considered to reflect directly the stored resources available for reproduction (Weiner 1988; Burd 1999). Size can also indirectly affect reproductive allocation because plants may use vegetative growth as an indicator of environmental quality (de Jong 1994; Stocklin and Favre 1994; Weiner 1988; Lloyd and Bawa 1984). Accordingly, fecundity commonly increases with plant size (e.g. Samson and Werk 1986; Weiner 1988; Schmid et al. 1995; Vaughton and Ramsey 1997; Susko and Lovett-Doust 2000). In perennials the relationship between above ground plant size, available resources and fecundity is complicated because they may allocate resources to underground storage and because of trade offs between years in growth and reproduction (Harper 1977; Samson and Werk 1986). In G. clandestina several findings confirmed that increase in plant size correlated with increase of fecundity. Most importantly, the increase in plant size measured in leaf number correlated with increased flower and fruit production of both types in mature plants. Additionally, the commencement of flower production in juvenile, CL-only plants was associated with greater leaf number. Fruit set in CL flowers was also significantly greater in larger plants. Because CL flowers are unlikely to be pollen limited, this increase in CL fruit set is likely due to the decrease in resource limitation. Thus the change in the relative
allocation from cleistogamy in favour OP reproduction with increased plant size in *G. clandestina* was likely a response to an increase of resources available for reproduction.

Another, traditionally considered, reason for a change in relative investment in favour of chasmogamy with plant size may be an adaptation to increase pollinator effectiveness and greater fertilisation success in OP flowers in larger plants (Darwin, 1877). Plant size can affect pollinator service in two ways. Firstly, the pollinator service may depend directly on plant size where plants of very small stature are invisible to pollinators among the undergrowth (direct size effect - Klinkhamer *et al.* 1997). Threshold effects associated with an abrupt switch in reproductive strategy are expected under the direct effects of plant size (Charnov 1982; Klinkhamer *et al.* 1997). Thus the direct effect of plant size on pollinators may contribute to the lack of chasmogamy in very small plants in *G. clandestina*. More importantly, the effectiveness of pollinator service in OP flowers can be indirectly related to plant size and resource status because of its dependence on the size of floral display, with smaller displays being less attractive to pollinators (budget effect - Klinkhamer *et al.* 1997). Larger displays are often more attractive to pollinators in plants (e.g. Bierzychudek 1981; Momose and Inoue 1993; Burd 1994; Murren and Ellison 1996; Wilcock and Neiland 2002) and in *G. clandestina*, the packaging of OP flowers into racemes and synchronised OP flowering in contrast to CL produced singly and over a extended time indicate that its pollinators respond positively to size of displays. Contrary to the expectations based on the increase of pollinator efficiency with plant size, the OP fruit set or OP female reproductive success in mature plants remained independent of plant size in *G. clandestina* plants. Thus it is unlikely that the increase in investment in chasmogamy with plant size is an adaptation to the increased effectiveness in OP fruit production due to indirect effects of plant size on pollinator efficiency in *G. clandestina*.

Interestingly, while flowering patterns in mature plants showed an increase in proportion of chasmogamy with plant size, this trend was absent at the fruiting stage. As a result the proportion of chasmogamous fruit was the same independent of plant size. This suggested that while the relative adaptive value of OP reproduction increased with plant size in *G. clandestina*, this increase did not depend on production or properties of chasmogamous seeds (or OP female function).
Overall, these results indicate that the plant size dependent investment in OP flower production is a response to change in resource availability and the direct effects of plant size on pollinator effectiveness. At the same time this reproductive strategy is unlikely to be based on the female function in OP flowers, either due to increased efficiency in fruit set with plant size or factors related to seed quality.

**Comparison with other CLOP**

The observations in *G. clandestina* conflict with those made in most other perennials, in which the proportion of OP flowers was noted to be lower in larger plants (*Danthonia spicata*, Clay 1982a; *Oxalis montana* Jasieniuk and Lechowicz 1987, *Viola mirabilis* Mattila and Salonen 1995, *Oxalis acetosella* Berg and Redbo-Torstensson 1998). On the contrary, the pattern of investment in the two types of flower in *G. clandestina* plants followed those typical of annuals (Lu 2002; Le Corff 1993; Schmitt, 1987; Waller, 1980; Wilken 1982; Schnee and Waller 1986; Diaz and Macnair 1998) and amphicarpic species (see review by Cheplick 1994).

Size-dependent chasmogamy in *G. clandestina* shares some features with patterns found in annuals, including the size threshold for commencement of OP reproduction followed by further increase in OP investment with plant size (Lu 2002; Le Corff 1993; Schmitt *et al.* 1987; Waller 1980; Wilken 1982; Schnee and Waller 1986; Diaz and Macnair 1998). As in *G. clandestina* the increase in proportion of OP flowers with plant size arises from the difference in the rate of increase in CL and OP flowers in some of the species (*Collomia grandiflora* Wilken 1982; *Mimulus nastusus* Diaz and Macnair 1998; *Amphicarpea bracteata* Schnee and Waller 1986). In other species, including *Impatiens* (Waller 1980; Lu 2000) and amphicarps (Cheplick 1994) it is a consequence of an increase in chasmogamy with plant size with no change in CL investment. On the other hand, there is an important difference between *G. clandestina* and the other species. No previous study has revealed a discrepancy between the relative investment in chasmogamy at flowering and fruiting stage. While some of the studies did not test this separately (Schnee and Waller 1986; Diaz and Macnair 1998, Wilken 1982; Diaz and Macnair 1998; Waller 1980; Lu 2000), others revealed similar trends at flowering and fruiting (Cheplick 1994; Schmitt *et al.* 1987). As a result, *G. clandestina* is the first known case where the relative investment in chasmogamy is may not be exclusively a
response to selection based on the female function of OP flowers. As such this aspect of CLOP strategy warrants further study in this and other species.

An ‘annual-like’ investment pattern has also been found in two other perennials (*Dichanthelium clandestinum*, Bell and Quinn 1987; *Calathea micans*, Le Corf 1993). Together with the findings of this study, this suggests that there is no uniform ‘perennial-like’ investment strategy in relation to plant size in CLOP plants, in disagreement with suggestions by some authors (e.g. Jasieniuk and Lechowicz 1987; Berg and Redbo-Torstensson 1998). Instead it appears that perennial life history is associated with both the increase and decrease of cleistogamy with plant size, and the type of response is species-specific.

**Conclusions**

In agreement with Schoen and Lloyd’s (1984) model for the evolution of cleistogamy, this study clearly shows that seed production in CL flowers is both more reliable and economical than OP flowers in *Glycine clandestina*. Moreover, *G. clandestina* plants appear to employ a conditional strategy where cleistogamy ensures seed production when production of OP seed is likely to be limited by resources, pollinators or both. This is manifest in a complex pattern of relative investment in cleistogamy and chasmogamy. The smallest *G. clandestina* plants reproduce only through cleistogamy, which is likely to be a response to resource and pollinator limitation on OP seed production at this stage. In larger plants with both types of flowers, production of cleistogamy later in the season is regulated in response to limitation in fruit set in OP flowers early in the season. This regulation is likely to provide assurance of seed production when pollinators or temporary shortage of resources restricts OP fruit set, but is of limited use at high elevations where the shorter growing season curtails CL flower production. The plants also diminish their relative investment in CL flowers as they increase size and become less resource limited.

The case of *G. clandestina* does not support the hypothesis that perennials have a unique strategy of size-dependent investment in cleistogamy, rather this appears to be a species specific response that is independent of life history. Nonetheless, it supports the hypothesis that such regulation of cleistogamy based on OP fruit set evolved in response to the perennial life history. In this light it will be of interest to further investigate the
inter-specific differences in the reproductive allocation in perennials, and the role of perennial-specific factors, like delayed costs of reproduction, on the conditional investment in cleistogamy.

The plant size dependent change in relative investment in cleistogamy in *G. clandestina* showed an intriguing pattern – the increased proportion of OP flowers with plant size was not paralleled by increased proportion of OP fruit. This suggests that the increase in proportion of OP flowers with size in this species is unlikely to be selected based solely on the properties of seeds or female function. In contrast to CL flowers, in OP flowers the reproductive output also includes seeds sired by pollen exported to other plants or male function. The potential for the role of the chasmogamous male function in selection for plant-size dependent investment in *G. clandestina* is further discussed in Chapter 5.

Despite the demonstrable benefits of reproduction via cleistogamy, *G. clandestina* retains the chasmogamous mode of reproduction, which even dominates as the plant size increases. One of the traditionally postulated benefits of chasmogamous reproduction that could explain its retention is the avoidance of inbreeding depression (ID) in progeny due to chasmogamous outcrossing. In *G. clandestina* this advantage is likely to be diminished because of the greater cleistogamous seed biomass that is likely to enhance relative performance of cleistogamous seedlings. The effects of ID and flower type-related seed biomass differences on seedling performance are investigated in Chapter 4.
Results of a mixed model ANOVA comparing plant size (ln(leaf number)) between *G. clandestina* plants employing two different reproductive strategies: Juvenile (J) – CL flowering only and Adult or Mature (A) – OP and then optional CL flowering. Reproductive strategy was considered a fixed factor, population (AR, TA, B2, FR) was a random factor and plants were a random factor nested within populations. Satterthwaite’s approximation was used to calculate the denominator degrees of freedom where the mean square is a composite of more than one source (*, quasi F-ratios, Winer *et al.*, 1991).

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population*</td>
<td>3 (158.1)</td>
<td>0.21</td>
<td>0.65</td>
<td>0.58</td>
</tr>
<tr>
<td>Plant[Population]</td>
<td>111</td>
<td>0.31</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>Reproductive strategy</td>
<td>1</td>
<td>92.5</td>
<td>246.41</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>150</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2 Results of mixed model nested analyses of covariance (ANCOVA) exploring the impact of plant size on proportion of cleistogamy and magnitude of reproductive output at flowering and fruiting stage over two years (1999 and 2000) and four populations (AR, TA, B2, FR) of *G. clandestina*. Population, year and plant nested within population and year were defined as random factors and plant size (sqrt(leaf number)) was included as a covariate. Satterthwaite’s approximation was used to calculate the denominator degrees of freedom where the mean square is a composite of more than one source (*, quasi F-ratios, Winer et al., 1991). Response variables included: A. % CL flowers - arcsin(CL flowers/total flowers); B. CL flower production - sqrt(CL flower no.); C. OP flower production - sqrt(OP flower no.); D. %CL fruit - arcsin(CL fruit no./total fruit no.); E. Fruit production - sqrt(total fruit no.).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Source</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P</th>
<th>b±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FLOWERS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A) %CL flowers</td>
<td>Plant size</td>
<td>1</td>
<td>1341.28</td>
<td>9.57</td>
<td>0.003</td>
<td>-1.31±0.42</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (164.5)</td>
<td>7146.87</td>
<td>51.48</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (176.1)</td>
<td>290.06</td>
<td>2.09</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>138.72</td>
<td>0.99</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>62</td>
<td>140.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B) CL flower production</td>
<td>Plant size</td>
<td>1</td>
<td>18.23</td>
<td>4.06</td>
<td>0.048</td>
<td>0.15±0.07</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (2.9)</td>
<td>206.45</td>
<td>7.67</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (3.03)</td>
<td>53.87</td>
<td>2.05</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population xYear*</td>
<td>3 (175.2)</td>
<td>26.9</td>
<td>6.03</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>143</td>
<td>4.45</td>
<td>0.99</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>62</td>
<td>4.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C) OP flower production</td>
<td>Plant size</td>
<td>1</td>
<td>1193.84</td>
<td>76.79</td>
<td>&lt;.0001</td>
<td>1.23±0.14</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (159.3)</td>
<td>498.79</td>
<td>24.78</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (168.7)</td>
<td>53.68</td>
<td>2.69</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>20.51</td>
<td>1.32</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>62</td>
<td>15.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FRUIT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D.) %CL fruit</td>
<td>Plant size</td>
<td>1</td>
<td>544.65</td>
<td>1.04</td>
<td>0.31</td>
<td>-0.82±0.80</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (163.3)</td>
<td>21679.9</td>
<td>39.05</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (175.9)</td>
<td>4244.49</td>
<td>7.66</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>558.13</td>
<td>1.07</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>65</td>
<td>520.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.) Fruit production</td>
<td>Plant size</td>
<td>1</td>
<td>96.26</td>
<td>39.44</td>
<td>&lt;.0001</td>
<td>0.35±0.05</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (2.9)</td>
<td>54.64</td>
<td>9.01</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (3.1)</td>
<td>13.99</td>
<td>2.34</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population xYear*</td>
<td>3 (172.84)</td>
<td>6.06</td>
<td>2.31</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>143</td>
<td>2.84</td>
<td>1.08</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>62</td>
<td>2.44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

68
Table 2.3 Results of mixed model nested analyses of covariance (ANCOVA) exploring the impact of plant size and OP fruit set on investment in cleistogamy over two years (1999 and 2000) and four populations (AR, TA, B2, FR) of *G. clandestina*. Population, year and plant nested within population and year were defined as random factors, and plant size (sqrt(leaf number)) and OP fruit set (arcsin(OP fruit no./OP flower no.)) were included as covariates. Satterthwaite's approximation was used to calculate the denominator degrees of freedom where the mean square is a composite of more than one source (*, quasi F-ratios, Winer et al., 1991). A. ANCOVA testing of correlation between the two covariates. ANCOVAs testing the effect of both covariates on: B. CL fruit production- sqrt(CL fruit no.); C. CL fruit set- arcsin(CL fruit no./CL flower no.); D. CL flower production - sqrt(CL flower no.).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Source</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>b±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) OP fruit set</td>
<td>Plant size</td>
<td>1</td>
<td>18.52</td>
<td>0.48</td>
<td>0.49</td>
<td>0.1±0.22</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (160.44)</td>
<td>853.62</td>
<td>18.41</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (170.39)</td>
<td>687.51</td>
<td>14.95</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>47.03</td>
<td>1.23</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>62</td>
<td>38.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B) CL fruit production</td>
<td>Plant size</td>
<td>1</td>
<td>30.57</td>
<td>21.78</td>
<td>&lt;.0001</td>
<td>0.20±0.04</td>
</tr>
<tr>
<td></td>
<td>OP fruit set</td>
<td>1</td>
<td>7.47</td>
<td>5.32</td>
<td>0.02</td>
<td>-0.06±0.02</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (3.281)</td>
<td>34.27</td>
<td>7.9</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (3.6289)</td>
<td>1.19</td>
<td>0.3</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PopulationxYear*</td>
<td>3 (165.72)</td>
<td>4.82</td>
<td>2.34</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>2.14</td>
<td>1.52</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>58</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C) CL fruit set</td>
<td>Plant size</td>
<td>1</td>
<td>1547.92</td>
<td>8.94</td>
<td>0.004</td>
<td>1.59±0.53</td>
</tr>
<tr>
<td></td>
<td>OP fruit set</td>
<td>1</td>
<td>441.72</td>
<td>2.55</td>
<td>0.11</td>
<td>-0.47±0.29</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (3.0074)</td>
<td>1781.51</td>
<td>0.79</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (3.2168)</td>
<td>298.99</td>
<td>0.18</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PopulationxYear*</td>
<td>3 (153.43)</td>
<td>2304.62</td>
<td>10.22</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>125</td>
<td>234.75</td>
<td>1.35</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>51</td>
<td>173.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D) CL flower production</td>
<td>Plant size</td>
<td>1</td>
<td>20.84</td>
<td>4.45</td>
<td>0.04</td>
<td>0.17±0.05</td>
</tr>
<tr>
<td></td>
<td>OP fruit set</td>
<td>1</td>
<td>3.89</td>
<td>0.83</td>
<td>0.36</td>
<td>-0.04±0.08</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3 (3.2163)</td>
<td>94.55</td>
<td>4.94</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Year*</td>
<td>1 (3.4633)</td>
<td>30.33</td>
<td>1.76</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PopulationxYear*</td>
<td>3 (177.31)</td>
<td>21.49</td>
<td>4.79</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plant[Population,Year]</td>
<td>142</td>
<td>4.45</td>
<td>0.95</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>58</td>
<td>4.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.4  The comparison of CL and OP reproductive mode in *Glycine clandestina* in respect of cost per flower (flower dry mass) (A), cost per seed (seed dry mass) (B), number of ovules per flower (C), seed number per fruit (D, E), probability of seed set within fruit (F), and probability of fruit set per flower (G) using a mixed model split-plot analysis of variance (ANOVA). Population (AR, TA, B2, FR) was included in all of the analyses with year (1999, 2000) added in the analysis G as a between plot random factors. Flower or fruit type (OP and CL) was a fixed within plot factor. Plants were random plots nested within population (A-F) or within population and season (G). Both flower/fruit types were examined on the same plant. The non-orthogonal, planned comparisons between the two flower or fruit types within populations were conducted using a synthetic denominator from the model. Satterthwaite's approximation was used to calculate the denominator degrees of freedom where the mean square is a composite of more than one source (*, quasi F-ratios, Winer et al., 1991).

<table>
<thead>
<tr>
<th>Character</th>
<th>Source</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A.) Flower cost (mg)</strong></td>
<td>Population*</td>
<td>3 (3.027)</td>
<td>0.22</td>
<td>0.9</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Plant [Population]</td>
<td>36</td>
<td>0.06</td>
<td>1.04</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>Flower type*</td>
<td>1 (3 )</td>
<td>82.03</td>
<td>340.35</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>Population x Fruit type</td>
<td>3</td>
<td>0.24</td>
<td>4.09</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>AR CL vs OP</td>
<td>1</td>
<td>19.83</td>
<td>336.51</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>TA CL vs OP</td>
<td>1</td>
<td>15.39</td>
<td>261.18</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>B2 CL vs OP</td>
<td>1</td>
<td>21.38</td>
<td>362.75</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>FR CL vs OP</td>
<td>1</td>
<td>26.15</td>
<td>443.66</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>36</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B.) Seed cost (mg)</strong></td>
<td>Population*</td>
<td>3 (73 )</td>
<td>702.09</td>
<td>14.37</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Plant [Population]</td>
<td>73</td>
<td>48.87</td>
<td>4.54</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Fruit type</td>
<td>1</td>
<td>386.89</td>
<td>35.94</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>76</td>
<td>10.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C.) Ovule number/fruit</strong></td>
<td>Population*</td>
<td>3 (4.4757)</td>
<td>1.44</td>
<td>9.74</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Plant [Population]</td>
<td>74</td>
<td>0.04</td>
<td>2.71</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Flower type*</td>
<td>1 (3.0302)</td>
<td>0.88</td>
<td>7.54</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Population x Fruit type</td>
<td>3</td>
<td>0.12</td>
<td>7.62</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>AR CL vs OP</td>
<td>1</td>
<td>0.04</td>
<td>2.54</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>TA CL vs OP</td>
<td>1</td>
<td>0.03</td>
<td>2.01</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>B2 CL vs OP</td>
<td>1</td>
<td>0.42</td>
<td>26.68</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>FR CL vs OP</td>
<td>1</td>
<td>0.58</td>
<td>36.48</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>74</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D.) Seed number/fruit</strong></td>
<td>Population*</td>
<td>3 (17.623)</td>
<td>3.29</td>
<td>20.38</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Plant [Population]</td>
<td>82</td>
<td>0.13</td>
<td>4.76</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Fruit type*</td>
<td>1 (3.0621)</td>
<td>0.03</td>
<td>0.55</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Population x Fruit type</td>
<td>3</td>
<td>0.06</td>
<td>2.34</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>82</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character</td>
<td>Source</td>
<td>DF</td>
<td>MS</td>
<td>F</td>
<td>Prob&gt;F</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----------------------------</td>
<td>----</td>
<td>-----</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>E.) Seed number/fruit</strong></td>
<td>Population</td>
<td>3</td>
<td>1.75</td>
<td>12.92†</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>(Non-predated fruit)</td>
<td>Plant [Population]</td>
<td>74</td>
<td>0.11</td>
<td>5.47</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Fruit type</td>
<td>1</td>
<td>0.02</td>
<td>0.39†</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Population x Fruit type</td>
<td>3</td>
<td>0.05</td>
<td>2.35</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>74</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>F.) Prob(Seed set/fruit)</strong></td>
<td>Population</td>
<td>3</td>
<td>95.37</td>
<td>0.38†</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Plant [Population]</td>
<td>74</td>
<td>155.07</td>
<td>2.6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Fruit type</td>
<td>1</td>
<td>2464.03</td>
<td>16.38†</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Population x Fruit type</td>
<td>3</td>
<td>153.94</td>
<td>2.58</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>AR CL vs OP</td>
<td>1</td>
<td>697.83</td>
<td>11.7</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>TA CL vs OP</td>
<td>1</td>
<td>216.86</td>
<td>3.64</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>B2 CL vs OP</td>
<td>1</td>
<td>339.35</td>
<td>5.69</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>FR CL vs OP</td>
<td>1</td>
<td>1363.32</td>
<td>22.86</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>74</td>
<td>59.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>G.) Prob(Fruit set)</strong></td>
<td>Year*</td>
<td>1</td>
<td>1001.21</td>
<td>1.51</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Population*</td>
<td>3</td>
<td>505.28</td>
<td>0.27</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>Year x Population*</td>
<td>3</td>
<td>729.85</td>
<td>5.51</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Plant [Population, Year]</td>
<td>127</td>
<td>129.46</td>
<td>0.83</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Flower type*</td>
<td>1</td>
<td>9462.88</td>
<td>6.98</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Population x Flower type</td>
<td>3</td>
<td>1469.74</td>
<td>9.46</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>AR CL vs OP</td>
<td>1</td>
<td>11735.71</td>
<td>75.57</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>TA CL vs OP</td>
<td>1</td>
<td>3612.14</td>
<td>23.26</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>B2 CL vs OP</td>
<td>1</td>
<td>4473.54</td>
<td>28.81</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>FR CL vs OP</td>
<td>1</td>
<td>70.03</td>
<td>0.45</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>239</td>
<td>155.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.5 Results from the mixed model split-plot ANOVA comparing fruit initiation per inflorescence (arcsin(mean no. initiated fruit/no. flowers per raceme)) among three treatments (Autonomous, Open pollinated and Supplemented) testing for the efficiency of pollination in OP flowers of G. clandestina. The treatment was a fixed within plot factor, and population was a random between plot factor, with plants as random plots nested within populations. The three treatments were applied on the same plant. The non-orthogonal, planned comparisons between pairs of treatments were conducted using a synthetic denominator from the model. Satterthwaite’s approximation was used to calculate the denominator degrees of freedom where the mean square is a composite of more than one source (*, quasi F-ratios, Winer et al., 1991).

<table>
<thead>
<tr>
<th>Source</th>
<th>MS</th>
<th>DF</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>318.34</td>
<td>2 (12)</td>
<td>4.48</td>
<td>0.04</td>
</tr>
<tr>
<td>Plant[Population]</td>
<td>71.04</td>
<td>12</td>
<td>1.38</td>
<td>0.23</td>
</tr>
<tr>
<td>Treatment</td>
<td>7764.17</td>
<td>2</td>
<td>151.13</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Autonomous vs Open pollinated</td>
<td>2846.20</td>
<td>1</td>
<td>55.40</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Supplemented vs Open pollinated</td>
<td>5020.09</td>
<td>1</td>
<td>97.72</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>51.37</td>
<td>28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.6  Estimates of the relative requirements of reproductive biomass (mg dry mass) for production of a chasmogamous (OP) and a cleistogamous (CL) fruit in (A) *Glycine clandestina* (this study) and (B) six other CLOP species. For each estimate a ratio of chasmogamous and cleistogamous biomass requirement is given (OP/CL). Total biomass requirement/fruit for other studies are reported as given by the authors. Components of the reproductive biomass calculated based on the reported data. Components of reproductive biomass: 1. Average flower biomass; 2. Average flower biomass x (1-Average fruit set); 3. Average seed no./fruit x Average seed biomass

<table>
<thead>
<tr>
<th>Species</th>
<th>Flower type</th>
<th>Components of reproductive biomass</th>
<th>Total biomass requirement/fruit</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1. Flower</td>
<td>2. Excess flowers</td>
<td>3. Fruit</td>
</tr>
<tr>
<td>(A) <em>Glycine clandestina</em> Population</td>
<td>AR</td>
<td>OP</td>
<td>2.34</td>
<td>46.02</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.35</td>
<td>0.85</td>
<td>355.31</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>6.72</td>
<td>54.07</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>OP</td>
<td>2.01</td>
<td>43.59</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.25</td>
<td>1.17</td>
<td>378.37</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>7.96</td>
<td>37.35</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>OP</td>
<td>2.26</td>
<td>55.03</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.19</td>
<td>1.00</td>
<td>240.73</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>11.77</td>
<td>55.06</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>FR</td>
<td>OP</td>
<td>2.49</td>
<td>18.18</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.20</td>
<td>1.94</td>
<td>198.69</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>12.21</td>
<td>9.36</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>OP</td>
<td>2.28</td>
<td>40.7</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.25</td>
<td>1.24</td>
<td>293.28</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>9.19</td>
<td>32.83</td>
<td>0.93</td>
</tr>
<tr>
<td>(B) <em>Microalena polynoda</em></td>
<td>OP</td>
<td>1.62</td>
<td>5.42</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1.26</td>
<td>0.03</td>
<td>2.81</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>1.29</td>
<td>210.91</td>
<td>0.4</td>
</tr>
<tr>
<td><em>Dichanthelium clandestinum</em></td>
<td>OP</td>
<td>1.88</td>
<td>4.42</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1.76</td>
<td>2.15</td>
<td>1.07-1.16</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>1.07-1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Calathea micans</em></td>
<td>OP</td>
<td>12</td>
<td>45.14</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>3</td>
<td>2.36</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>4</td>
<td>19.15</td>
<td>1.6</td>
</tr>
<tr>
<td><em>Impatiens pallida</em></td>
<td>OP</td>
<td>17.2</td>
<td>480-426</td>
<td>20-100</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>1</td>
<td>197-150</td>
<td>19.5-58.5</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>17.2</td>
<td>2.43-2.83</td>
<td>1.03-1.71</td>
</tr>
<tr>
<td><em>Impatiens canescens</em></td>
<td>OP</td>
<td>11.5</td>
<td>70-163</td>
<td>13.3-66.5</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.4</td>
<td>46-59</td>
<td>11.5-34.5</td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>28.75</td>
<td>1.51-2.77</td>
<td>1.16-1.93</td>
</tr>
<tr>
<td><em>Impatiens canescens</em></td>
<td>OP</td>
<td>9.84</td>
<td>18.45</td>
<td>Waller 1979</td>
</tr>
<tr>
<td></td>
<td>CL</td>
<td>0.25</td>
<td>12.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OP/CL</td>
<td>39.36</td>
<td>1.48</td>
<td>73</td>
</tr>
</tbody>
</table>
Fig. 2.1 The mean size of plants falling into four classes of individual flowering strategy observed in four populations of *G. clandestina*: none – no flowers produced, CL – only CL flowers produced, CLOP – both CL and OP flowers produced, OP – only OP flowers produced. Thin bars represent SE.

Fig. 2.2 Set up of the pollinator exclusion experiment with a nylon mesh bag enclosing part of the plant in order to exclude pollinators and supported with a stake to prevent damage to the plant. A branch with opening OP racemes is visible in the foreground. Each plant was fenced off to prevent damage from herbivores.
Fig. 2.3 Phenology of CL and OP flower and mature fruit production in the four *G. clandestina* populations (A – AR, B – TA, C – B2, D – FR) as determined by fortnightly census of the total number of flowers and mature fruit produced by the censused plants per population over c. 21 weeks in spring and summer (mid-August 1999 to mid-April 2000). Black – chasmogamy, grey - cleistogamy, broken line – flowers, solid line – fruit.
Fig. 2.4 Relationship between plant size (sqrt(leaf no.)) and flower (sqrt(flower no.)) (A) and fruit (sqrt(fruit no.)) (B) production via cleistogamy (CL, gray square) and chasmogamy (OP, black cross) per plant across four *G. clandestina* populations. The difference in slopes between cleistogamy and chasmogamy was tested using ANCOVA on proportion of cleistogamy (see Table 2.2). Regression lines are given with relevant equations as determined by nested ANCOVA (see Table 2.2).
Fig. 2.5 Differences between OP (white bars) and CL (grey bars) flowers and fruit on G. clandestina plants with both flower and fruit types in the four populations (AR, TA, FR, B2) in terms of: (A) cost per flower in dry mass (mg); (B) cost per seed in dry mass (mg); (C) ovules per flower (sqrt(no. ovules)); (D) seeds per fruit (sqrt(no. seeds)); (E) probability of seed set per mature fruit (arcsin(no. seeds/no. ovules)). Presented are the least square means and corresponding SE (represented by the bars) from the mixed model split-plot ANOVAs testing for the significance of the differences (see Table 2.4).
KEY TO LABELLING: The bars represent 250 \( \mu m \); a – anther; ad – dehisced anther; c – calyx; f – filament; hw – lowermost whorl of stigmatic papillae; s – stigmatic surface, covered with stigmatic cuticle or membrane which serves as containment for the stigmatic exudate; sm – staminal tube; sy – style; o – ovary; p – pollen grain; pg – germinated pollen grains.

Fig. 2.6  SEM view of developmental sequence of gynoecium and androecium in OP (A-D) and CL (E-G) flowers in *G. clandestina* illustrating the divergent pollination mechanisms in the two types of flowers. The two types of flowers are shown at three progressive developmental stages: *Advanced bud* – OP bud shortly before anthesis (A) and CL bud before the emergence of the style between the sepals (E); *Pollination* – OP flower after anthesis and before pollination (the few pollen grains on the stigma deposited during preparation process) (B) and after a couple of days of exposure to pollinators with pollen germinating on stigmatic surface (C); CL flower at the stage of the bent style emerging between the sepal lobes with two anthers attached to the stigma via the germinated pollen tubes and fractured filaments (F); *Fruit growth* – early stages of fruit expansion after fertilisation in the two flowers. Corolla still attached in OP flower but the style extending and straightening (D), and in CL flower bent style fully emerged from the calyx, the two anthers still attached to the stigma (G).
Fig. 2.7 Least square means of fruit initiation rate per inflorescence (arcsin(\sqrt{\text{initiated fruit/total flowers per inflo}})) in three treatments applied to *G. clandestina* plants: Autonomous - pollinators excluded and flowers unmanipulated; Open pollination - flowers exposed to natural pollination; Supplemented - pollinators excluded, flowers self-pollinated by hand. The vertical bar represents SE from the model. The ANOVA and paired contrast analysis for differences among treatments is presented in Table 2.5.

Fig. 2.8 Long-term mean day temperature over *G. clandestina* reproductive period at upper and lower limit of the range of elevations of the study populations, near the study location near Canberra, Australia. Start of flowering is marked for populations located at the elevations corresponding to the lower elevation limit (AR and TA from NAR morphotype, and B2 from BRO morphotype) and the high elevation limit (FR from BRO morphotype). The temperature data was sourced from The Commonwealth Bureau of Meteorology, Australia (2003). Station 071010 KIANDRA CHALET (1300 m) and Station 070014 CANBERRA AIRPORT (600 m).
Chapter 3

Limited inter-population variation in chasmogamous inbreeding and small-scale genetic structure in *G. clandestina* – implications for the advantage to chasmogamy from avoidance of inbreeding depression
Introduction

As outlined in Chapter 1, the relative advantage to chasmogamy over cleistogamy derived from the avoidance of inbreeding depression (ID) can be affected by two factors: the mating patterns in OP flowers and the population genetic structure. Selfing or biparental inbreeding (outcrossing to close relatives) in OP flowers will diminish the difference in inbreeding between OP and obligatorily selfed CL progeny and thus the relative superiority of OP seed (Darwin 1897; Schoen and Lloyd 1984; Masuda et al. 2001). A degree of small-scale spatial genetic structure within populations can also affect inbreeding of progeny from outcrossed matings, either leading to biparental inbreeding and diminished expression of ID between cleistogamy and chasmogamy or outbreeding and enhanced expression of ID (Uyenoyama 1986; Waller 1993).

Theoretically, variation in either of these factors among populations could change the relative advantage of cleistogamy and chasmogamy.

In early considerations of CLOP selfing was assumed to be minimal in OP flowers because of the expectation that the function of chasmogamy is to provide outcrossed progeny in CLOP species (Darwin 1897; Uphof 1938; Schemske 1978; Waller 1980; Antlfinger 1986). However, now there is abundant evidence that selfing rates can be high in OP flowers, with estimates ranging between 0.3 to 0.9 (Wilken 1982; Mitchell-Olds and Waller 1985; Brown et al. 1986; Waller and Knight 1989a; Schoen and Brown 1991; Stewart 1994; Lu 2000; Culley 2002), and only three reported instances of selfing below 0.12 (Mitchell-Olds and Waller 1985; Lu 2000; Culley 2002). In addition to selfing, biparental inbreeding can further contribute to inbreeding in OP flowers as reported in studies of Impatiens (I. capensis, Waller and Knight 1989a; Lu 2000; I. pallida, Stewart 1994).

Within CLOP species the level of inbreeding in OP flowers may vary broadly potentially affecting the relative advantage to chasmogamy over cleistogamy derived from the avoidance of inbreeding depression. Studies in I. capensis show that selfing in chasmogamous flowers can vary substantially among populations with the recorded range spanning from 0.03 to 0.70 (Mitchell-Olds and Waller 1985; Waller and Knight 1989a). Substantial variation in OP selfing rate within populations between years has
also been recorded in *I. pallida* and *Viola pubescens* (*V. pubescens*: 0.07-0.60, Culley 2002; and *I. pallida*: 0.48-0.66, Stewart 1994).

Apart from inbreeding in OP flowers, the strong genetic differentiation and small-scale spatial genetic structure of CLOP populations is also likely to affect the relative advantage of OP reproduction. Strong genetic differentiation at multiple scales is likely to be a feature of CLOP systems because of their habitual self-fertilisation. Levels of cleistogamy in many species indicate that populations self-fertilise at least at an intermediate rate (Schemske 1978; Clay 1982a; Clay 1982b; Wilken 1982; Schoen 1984; Bell and Quinn 1987; Jasieniuk and Lechowicz 1987; Le Corff 1993; Mattila and Salonen 1995; Redbo-Torstensson and Berg 1995; Cheplick 1996; Berg and Redbo-Torstensson 1998; Diaz and Macnair 1998). Direct estimates of total selfing rate are rare but they show that combined contributions of mating in OP and CL flowers can result in predominant selfing, even when OP flowers are outcrossing (Wilken 1982; Lu 2000). Predominant selfing is also apparent in indirect estimates based on inbreeding levels detected in mature plants in several CLOP species (Mitchell-Olds and Waller 1985; Waller and Knight 1989b; Schoen and Brown 1991; Cole and Biesboer 1992; Stewart 1994; Sun 1999).

Knight and Waller (1987) recognised that variation in inbreeding among populations could profoundly affect the relative advantage of outcrossing and OP reproduction and lead to differences in relative investment in cleistogamy and chasmogamy among populations. Great variation in inbreeding levels was revealed in studies of *Viola pubescens* and *Impatiens capensis* (Brown et al. 1986; Knight and Waller 1987; Culley and Wolfe 2001; Culley 2002). Such variation is likely more widespread since the two sources of inbreeding in CLOP, cleistogamous selfing (e.g. Waller 1980; Clay 1982a; Clay 1982b; Campbell et al. 1983; Bell and Quinn 1987; Jasieniuk and Lechowicz 1987; Le Corff 1993; Mattila and Salonen 1995; Berg and Redbo-Torstensson 1998; Diaz and Macnair 1998) and OP selfing (see above) often show strong inter-population differences.

Knowledge of the effects of habitual selfing and its variation on genetic structure in CLOP populations is limited, mainly due to the lack of variable population genetic markers (e.g. Waller and Knight 1989b; Cole and Biesboer 1992; Sun 1999). In four
CLOP species where inter-population genetic differentiation was successfully studied using allozymes, the high inbreeding levels were shown to be associated with high genetic differentiation between populations typical of predominant (Waller and Knight 1989b; Cole and Biesboer 1992; Sun 1999) or intermediate selfers (Culley and Wolfe 2001). Only one study has examined genetic structure within CLOP populations in two *Impatiens* species, using morphological characters assumed to represent neutral loci (Schoen and Latta 1989). Schoen and Latta (1989) reported significant small-scale genetic structure within all of the studied populations, which was consistent with the habitual selfing and restricted dispersal in the species, but they were not able to test for differences in the degree or strength of this substructuring among the populations.

Investigation of mating and genetic patterns in populations of *Glycine clandestina* will contribute to our understanding of the potential for the interpopulation variation in OP mating and small-scale spatial genetic structure to affect the relative advantage of chasmogamy derived from avoidance of inbreeding. Previous estimates of mating in OP flowers in *G. clandestina* and closely related *G. argyrea* showed that OP flowers self-fertilise at substantial rates and that the habitual selfing is reflected in inbreeding in mature individuals (Schoen and Brown 1991). The high selfing in OP flowers when combined with obligatory CL selfing in this species are likely to restrict gene flow between and within populations. Strong differentiation in leaf morphology previously noted among *G. clandestina* populations (Pfeil et al. 2001) suggest that genetic differentiation is strong at the interpopulation scale, and that local selection may enhance it at relatively short geographic distances. It is likely that strong small-scale spatial genetic structure also exists within *G. clandestina* populations as the effects of high inbreeding in this species are unlikely to be counteracted by gene flow through ballistically dispersed seed.

More importantly, in a previous study Schoen and Brown (1991) found that OP selfing rate and inbreeding level was almost twice as high in high elevation population of broad leaf (BRO) morphotype than in a low elevation population of narrow leaf (NAR) morphotype of *G. clandestina* included in their study. This clearly suggests that, similarly to other CLOP species, populations of *G. clandestina* may differ in aspects that affect the relative advantage to chasmogamy derived from avoidance of inbreeding. Moreover, the difference observed by Schoen and Brown (1991) may represent ecotypic
differences in reproductive strategy between BRO and NAR morphotype of *G. clandestina*. Under this assumption the BRO morphotype populations would be expected to show higher level of OP selfing, greater level of inbreeding, enhanced small-scale spatial genetic structure with greater potential for biparental inbreeding than NAR populations.

The objectives in this chapter were two-fold: to examine the patterns of selfing and associated genetic structure of *Glycine clandestina* populations, and to determine if there are differences in OP selfing, total selfing and small-scale spatial genetic structure between *G. clandestina* populations of BRO and NAR morphotypes that could potentially affect the relative advantage to chasmogamous reproduction derived from inbreeding avoidance. In order to achieve these objectives, patterns of population genetic structure at neutral loci were analysed among and within seven *G. clandestina* populations and this was combined with mating analysis in four representative populations. In contrast to studies in other CLOP species, which were often limited by variability of the traditional allozyme markers, the availability of informative microsatellites (SSR) allowed for the required resolution of this population genetic structure in *G. clandestina* at multiple scales. The 11 microsatellite loci were transferred from the related *Glycine max* (Hempel and Peakall, 2003). The following specific questions were addressed:

1. What are the patterns of genetic differentiation at three scales: between morphotypes, among populations and within populations?

2. What is the level of inbreeding in mature plants and does it correspond to direct estimates of total selfing combining CL- and OP-selfing?

3. What is the level of inbreeding in OP flowers, including selfing and biparental inbreeding?

4. Is there evidence for inter-morphotype and inter-population differences in inbreeding, OP mating patterns, and small-scale spatial genetic structure?
Materials and Methods

Molecular analyses

Sampling and DNA extraction

Mature plants that were previously sampled for the microsatellite polymorphism analysis (Hempel and Peakall 2003, see Appendix 1) were used here to investigate patterns of differentiation and inbreeding across the seven *G. clandestina* populations (Table 1.1). These individuals represent a random sub-sample of all of the individuals recorded at intervals of a minimum of 3 m along two to three parallel transects placed at least 5 m apart, spanning most of the length of each population. Additional sampling was conducted in four of the populations for the analysis of the mating system in OP flowers. Fruits arising from OP flowers were collected from 15 of the genotyped plants, and 5-6 seeds were selected from separate fruits for genotyping. Maternal plants were chosen from a limited area to ensure they came from a randomly mating deme. DNA was extracted from whole seeds using modified CTAB method as per Krauss (1999). The Mendelian patterns of segregation in CL self-fertilised seeds confirmed that the method produced progeny genotypes (data not shown).

Microsatellite markers

Patterns of genetic variation in mature plants were assessed using genotypes generated in polymorphism screening at 11 polymorphic microsatellite loci developed for the purpose of this study (see Hempel and Peakall, 2003 in the Appendix 1). An additional locus developed by Peakall *et al.* (1998) was also included. For the summary of locus properties see Table 3.1. Five of the loci, which were variable across all seven populations (Table 3.1), were used to assess inter- and intra-population differentiation based on the total of 148 individuals (14 to 42 per population). For calculations of population and individual inbreeding coefficients I also included all of the individuals for which a genotype at one or more loci was missing and additional loci that were informative only in some populations. This data set represented 251 individuals. Additional genotyping was conducted for seeds used in OP mating system estimates. It was sufficient to only use three or four loci per seed within each of the four populations because of their high information content and the predominantly selfing mating system. The estimate of mating system parameters from genotypes of progeny arrays requires polymorphic loci (Brown 1989). In this study, loci were selected in order to maximize...
information content and accuracy of OP mating system estimate, i.e. the loci with good
information content were selected in each population. The high degree of differentiation
among populations of this species meant that different loci fit this criterion within each
population.

**PCR conditions and product visualisation**

Additional samples required for mating system analysis were genotyped following
conditions used in the marker development process for the polymorphism screening
(Hempel and Peakall 2003 in the Appendix). PCR products from two to four loci were
combined at the precipitation stage and loaded into the same lane when their sizes were
known not to overlap.

**Genetic structure at three spatial scales**

**Inter-morphotype and inter-population differentiation**

Genetic differentiation among the seven *G. clandestina* populations was explored using
a qualitative and a quantitative approach. First, Principal Coordinate Analysis (PCA)
was applied to the multilocus inter-individual genetic distance in order to visualise the
patterns of inter-population and inter-morphotype differentiation. Then Analysis of
Molecular Variance (AMOVA) was used to quantify the amount of variation distributed
among populations and morphotypes and to determine the significance of the observed
level of differentiation. The AMOVA was conducted with and without the grouping of
populations into morphotypes (Peakall *et al.* 1995; Maguire *et al.* 2002).

Inter-individual genetic distance based on an allelic distance (AD) measure formed an
input into AMOVA to produce *F*-statistics measuring the extent of genetic
differentiation among the populations (Weir and Cockerham 1984; Michalakis and
Excoffier 1996). The value of this differentiation measure is bounded by [0,1], with
values significantly greater than zero indicating significant genetic differentiation. The
significance of *F*>0 was determined based on 1000 permutations, with null distribution
generated by shuffling membership of individual genotypes among the populations. The
AD distance measure assumes an infinite allele mutation model (IAM) at the marker
loci, and represents the distance between alleles with distance equal 0 if they are the
same and 1 if they differ.
Alternative genetic distance measures exist for microsatellites, which assume stepwise mutation models (SMM) or their modifications (Raybould et al. 1998; Gustafsson 2000). It is still debated which model best fits the mutation mechanism at SSR loci (Estoup and Cornuet 1999), and structural differences among loci used in this study suggest that their mutation mechanism may differ (Hempel and Peakall 2003). The analyses based on IAM and SMM mutation models were compared in order to ensure that assumptions about mutation mechanism did not influence the conclusions. For SMM model the microsatellite distance measure (MD) according to (Slatkin 1995) was used as the input for AMOVA. In the MD the inter-individual distance is calculated as the sum of the squared size difference between the two alleles in comparisons to generate R-statistics, analogous to F-statistics.

The conclusions of significant inter-population and inter-morphotype differentiation or its variation with geographical distance (see below) based on R-statistics were not different to those based on F-statistic. However, the R-statistics showed greater variation among loci and greater magnitude of differences between the two morphotypes. SMM based-measures have been noted to occasionally produce higher estimates of interpopulation differentiation than IAM-based measures (Viard et al. 1996; Goodman 1998). The higher SMM-based values can be interpreted as more accurate when populations are more distantly related so that the divergence is due to a greater extent to mutation process rather than genetic drift (Rousset 1996). On the other hand, SMM-based differentiation statistics can be less reliable because of their large variance caused by the combination of the reliance on inter-allelic size differences and the heterogenous mutation processes across microsatellite loci (Estoup et al. 2002), and because they are vulnerable to biases by patterns at a few among the genotyped loci (Landry et al. 2002). Simulations show that AD-based measures may be preferable for use with small samples (<10) and small number of loci (<20) (Gaggotti et al. 1999). Because R-statistics showed substantial heterogeneity across loci, fewer than 20 loci were used, and because the geographical scale in our study corresponds to a relatively recent interpopulation divergence, the more conventional F-statistics was chosen as more suitable indicator of population differentiation for this study.

Sequencing of alleles at the locus Sat131 revealed an inter-morphotype size homoplasy with distinct array structure at the locus for NAR and BRO (see Hempel and Peakall
2003 in Appendix 1). If not taken into account in coding of genetic distance this could lead to underestimate of the magnitude of differentiation between the morphotypes. Additional analysis was conducted, where all of the alleles at BRO were coded as different to those at NAR. Because taking homoplasy into account did not change conclusions about significance and had no impact on the magnitude of overall, multilocus differentiation estimates, only the results based on allele size differences are presented.

The population differentiation analyses were run for the multilocus data using a set of individuals with full five-locus genotypes (148 individuals) as well as for single-locus data with a greater number of individuals per locus (total individuals 231-251 per locus) in order to assess consistency of the trends across loci.

In order to test among population genetic differentiation was proportional to geographic distance, Mantel test of matrix correspondence (Mantel 1967; Smouse et al. 1986; Peakall et al. 1995) was performed on the respective ln(linearised pairwise total genetic distance) and ln(linear geographic distance). Linearisation of differentiation statistics ($F_{ST}$) was carried out following (Rousset 1997) based on a stepping stone model:

$$F_{ST, linearized} = F_{ST} / (1 - F_{ST})$$

Mantel test was performed on interpopulation differentiation measures for all pairwise population comparisons as well as within each morphotype. The significance of Mantel test was based on 1000 permutations, with genotypes shuffled among populations. A significant linear relationship between genetic and geographic distance would indicate that differentiation followed the stepping stone model or isolation by distance (Kimura and Weiss 1964), where gene flow between populations is limited but regular. A lack of linear relationship would suggest the island model of differentiation (Wright 1969), where the gene flow is sporadic and coupled with frequent population extinction and colonisation.
Genetic structure within populations

**Distribution of genotypes within populations**

The five-locus genotypes of the same 148 individuals that were used in interpopulation differentiation analysis were also used in the analysis of genetic differentiation within populations using the method of spatial autocorrelation analysis developed by Smouse and Peakall (1999), and modified by Peakall et al. (2003). For this analysis the inter-individual genetic distance was coded using genotypic distance measure (GD) which integrates information across alleles (Peakall et al. 1995; Smouse and Peakall 1999). For example, for a single-locus, with i-th, j-th, k-th and l-th different alleles, a set of squared genetic distances is defined as $d^2(ij, ij) = 0$, $d^2(ii, ij) = 1$, $d^2(ij, ik) = 1$, $d^2(ik, kl) = 2$, $d^2(ij, jk) = 3$, and $d^2(ii, jj) = 4$. Unlike classical spatial autocorrelation analysis, this procedure combines information across alleles and loci, and if required, across populations, to strengthen the spatial signal by reducing stochastic noise (Peakall et al. 2003).

From the input of pairwise squared GD and linear geographic distance matrices the analysis generates autocorrelation coefficient, $r$. It is a proper correlation coefficient, bounded by $[-1,+1]$, and is closely related to Moran’s-I. The autocorrelation coefficient is expected to be positive between individuals separated by shorter distances when the distribution of genotypes in space is non-random with genetically related individuals aggregated into subpopulations or neighbourhoods (i.e. there is small-scale spatial genetic structure).

To visualise the patterns of spatial genetic autocorrelation, the traditional autocorrelograms were employed, plotting the autocorrelation coefficient $r$ as a function of distance. Permutation based 95% confidence intervals were constructed around the null hypothesis of $r = 0$ (no structure) for all distance classes. Additionally, Peakall et al. (2003) was followed to generate a bootstrap distribution of $r$ within each distance class to provide 95% confidence intervals for the estimates. Analysis was conducted with distance class size of 10 m, and maximum distance of 100 m.

Under non-random genotype distribution due to restricted gene flow the autocorrelogram is expected to show positive $r$ in the shortest distance classes followed by fluctuation of $r$ between negative and positive values. In order to test for the positive
non-random structure, a permutation-based one-tailed test for \( r > 0 \) was applied at the first distance class. When positive spatial genetic structure was found, the first x-intercept provided an estimate of the spatial scale of the non-random (positive) genetic structure (i.e. the patch, neighbourhood or subpopulation size)(Heywood 1991).

In addition to population-by-population analysis, a new procedure developed in Peakall et al. (2003) was applied to calculate \( r \) across multiple populations (\( rc \)), in order to increase the power of the analysis and allow for assessment of overall patterns for NAR and BRO populations. This is a better estimator than a simple arithmetic average across populations, since each pairwise comparison in each population contributes equally to the estimate. In order to test for differences in within-population differentiation between NAR and BRO, the overlap of 95% bootstrap confidence intervals around their respective \( rc \)-values within the first distance class was checked.

**Distribution of plants within populations**

A non-random spatial distribution of plants within populations can enhance within population differentiation (Doligez et al. 1998) or may indicate processes that do so (e.g. limits on seed dispersal or spatial heterogeneity of microhabitats, pollinator behaviour modification)(Linhart and Grant 1996). Therefore, the relationship between spatial and genetic structure was also investigated across populations. To determine whether distribution of *G. clandestina* individuals within populations was random, a spatial structure was analysed as described above, but with modified inputs as suggested by Peakall et al. (2003). The linear pairwise geographic distance matrix was calculated as the Euclidean distance between x- and y-coordinates for all of the transect positions at the three-meter intervals irrespective of plant presence. In order to determine if there was any relationship between the spatial and genetic autocorrelation results, linear regression analysis was performed between the \( r \) in the first distance class and the x-intercepts for the seven populations from the respective spatial and genetic data sets.

**Self-fertilisation rates within populations**

The rate of selfing within *G. clandestina* populations was investigated using two approaches. First, inbreeding in mature plants was estimated within all seven populations from their microsatellite genotypes. Second, this estimate was validated in four of the populations using direct estimates of selfing, which are a more reliable but also more laborious.
Self-fertilisation rates – indirect estimate from inbreeding in mature plants

Inbreeding in mature plants was estimated using Wright’s inbreeding coefficient (Wright 1965):

\[ F = 1 - \frac{H_o}{H_e}, \]

where \( H_o \) = observed heterozygosity, \( H_e \) = heterozygosity expected from the allele frequencies. \( F \) is bound by \([-1, 1]\) and values of \( F > 0 \) can be interpreted as evidence of excess homozygosity over that expected from allele frequencies under Hardy-Weinberg equilibrium (Wright 1965). Under the assumption that self-fertilisation is the only cause of excess homozygosity, the magnitude of \( F \) provides a reliable measure of the long-term self-fertilisation rate (Byers and Waller 1999).

Inbreeding coefficient was estimated at each locus where the number of genotypes was \( \geq 20 \) with five to 10 loci per population. The population inbreeding coefficient \( (F_{POP}) \) was estimated as an average across loci. Within each population the significance of \( F > 0 \) was tested at each locus and for the average across loci \( (F_{POP}) \). Bootstrap tests following (Van Dongen and Backeljau 1995) were applied to single-locus \( F \) since they are superior to exact tests for small samples (Van Dongen 1994; Van Dongen and Backeljau 1995), while for \( F_{POP} \) one-tailed t-tests were used.

The differences among populations in inbreeding patterns were tested by comparing the average per locus inbreeding coefficient \( (F_{POP}) \) using ANOVA, with populations (seven levels) nested within morphotypes (two levels: NAR and BRO).

Does inbreeding level in mature plants reflect self-fertilisation rates in populations?

In partially selfing populations an estimate of mating patterns from the inbreeding coefficient can be biased upwards if there is significant within population differentiation or downwards if inbreeding depression reduces the relative frequency of homozygous individuals. In order to establish the potential for such bias in \( G. \ clandestina \), selfing rates were estimated directly within four populations chosen to represent the range of adult plant inbreeding levels (as estimated by \( F \)) within each morphotype (AR and TA for NAR, FR and B2 for BRO).
Total population selfing rate

The estimate of total population selfing rate ($s_{TOT}$) necessitated the assessment of the proportion of reproductive output derived from self-fertilisation via obligatory CL selfing ($s_{CL}$) and self-fertilisation in OP flowers ($s_{OP}$) at a population level. Data on reproductive output for the 1999 reproductive season was used, which was collected for the study that is presented in Chapter 2.

As previously noted in Chapter 2, a sharp plant size threshold for reproductive strategy in *G. clandestina* was found with small juvenile plants only producing CL flowers and larger, mature plants producing either both CL and OP or only OP flowers. A survey of 20 by 20 m quadrats within each of the four populations at the end of the 1999 revealed that the juveniles constituted c. 60% of the populations. Given the mean reproductive output per mature and juvenile plant and their relative frequencies within the quadrats it was estimated that the juveniles contributed less than 10% of the total fruit set. In mature plants the proportion of CL and OP fruits did not vary with size, nor was there any significant difference in seed number between CL and OP fruits. Therefore, the number of fruits was treated as representative of the number of progeny produced through each reproductive mode within a population. Based on these considerations, estimates of the proportion of progeny resulting from CL selfing ($s_{CL}$), OP selfing ($s_{OP}$) were made in order to derive total selfing in a population ($s_{TOT}$). Variances were estimated using binomial error distribution.

For the total selfing ($s_{TOT}$) estimate, CL selfing ($s_{CL}$) was estimated as proportion of CL fruit, and the proportion of progeny derived from OP selfing ($s_{OP}$) was derived from a product of proportion of OP fruit and population selfing rate in OP flowers, $s_m$. The variances of these parameters were estimated following Mendenhall *et al.* (Mendenhall *et al.* 1981):

\[
\text{Var}(s_{CL}) = \left( \frac{\text{CL fruit no.}}{\text{Total fruit no.}} \right)^2 \left( \frac{\text{OP fruit no.}}{\text{Total fruit no.}} \right)\frac{\text{Total fruit no.}}{\text{Total fruit no.}}
\]

\[
\text{Var}(s_{OP}) = \left( \frac{\text{OP fruit no.}}{\text{Total fruit no.}} \right)^2 \text{Var}(s_m)^2 + \left( \frac{\text{CL fruit no.}}{\text{Total fruit no.}} \right)^2 \left( \frac{\text{OP fruit no.}}{\text{Total fruit no.}} \right) \left( \frac{\text{CL fruit no.}}{\text{Total fruit no.}} \right) + \text{Var}(s_m) \left( \frac{\text{OP fruit no.}}{\text{Total fruit no.}} \right) \left( \frac{\text{CL fruit no.}}{\text{Total fruit no.}} \right) \left( \frac{\text{Total fruit no.}}{\text{Total fruit no.}} \right)
\]
Var (s_{TOT}) = Var(s_{CL}) + Var(s_{OP})

For the estimate of \( s_m \) and its variance see the description of the estimate of selfing in OP flowers in the section below. Comparison of 95% confidence intervals based on these variances \((1.96 \times \sqrt{\text{Var}})\), were used for tests of significant difference among populations in \( s_{TOT}, s_{CL}, \) and \( s_{OP} \).

**Comparison between expected and observed inbreeding**

In order to determine whether the observed inbreeding in mature plants reflected the amount of total selfing within populations I derived the expected inbreeding coefficient from the observed total selfing rates \( (s_{TOT}) \) within four populations using:

\[
F_{\text{EXP}} = \frac{s_{TOT}}{2-s_{TOT}} \quad \text{(Jain 1979)},
\]

where \( s_{TOT} \) is the observed total selfing rate. This formula assumes that selfing is the primary factor contributing to departure of inbreeding coefficient from zero (i.e. no selection, subdivision of gene frequency, or outcrossing between relatives) and that population has achieved an inbreeding equilibrium, after selfing over several generations at a similar rate (Ritland 1983). The inbreeding coefficients expected from the total selfing \( (F_{\text{EXP}}) \) was compared with those observed in mature plants within populations \( (F_{\text{POP}}) \) using 95% confidence intervals around the estimates of the latter. No difference was expected between the two inbreeding coefficients if \( F_{\text{POP}} \) accurately reflected the total selfing rate within populations.

The \( F_{\text{POP}} \) can reflect both, selfing rates and the genetic substructure within a population. In order to test whether one or both of these factors influenced \( F_{\text{POP}} \) estimates, non-parametric test for correlation between \( F_{\text{POP}} \) and \( F_{\text{EXP}} \) (measure of selfing rates) as well as \( F_{\text{POP}} \) and the first distance class \( r \) (strength of population sub-structure) were performed. A significant correlation was expected between \( F_{\text{POP}} \) and \( F_{\text{EXP}} \) if the inbreeding coefficient \( F_{\text{POP}} \) is an accurate reflection of selfing rates. Similarly, a significant correlation was predicted between \( F_{\text{POP}} \) and \( r \) if population substructure is a confounding factor.
Selfing rate in OP flowers

Mating patterns in OP flowers were assessed based on microsatellite genotypes in maternal plants and progeny arrays using multilocus method described by Ritland (2002). The errors around the estimates were generated from 500 random bootstrap samples, with family as the unit of resampling. Chasmogamous multilocus outcrossing rate ($t_m$) was estimated using EM maximum likelihood algorithm, with pollen and ovule gene frequencies estimated separately. Because the estimate is based on the assumption of random outcrossing (with probability $t_m$) and selfing (with probability $s_m = 1-t_m$), the OP selfing rate was estimated as $s_m = 1-t_m$ (Ritland 1984).

The output from the analysis of mating system in OP flowers also allowed for estimates of biparental inbreeding, which was calculated as the difference between the multilocus ($t_m$) and minimum variance single locus ($t_s$) outcrossing rates. The difference is expected to be positive when outcrosses occur between relatives. The magnitude of the difference gives a lower bound for the level of biparental inbreeding (Ritland 1990b). A one-tailed t-test was used on the bootstrap values to establish if ($t_m-t_s$) > 0. The proportion of outcrossed matings that resulted in biparental inbreeding was estimated as $m_b = \frac{t_m-t_s}{t_m}$ (Ritland 2002), which is equivalent to $m_b$ of Waller and Knight (1989a). In order to test whether the magnitude of biparental inbreeding was correlated with the magnitude of within-population differentiation the positive correlations between the autocorrelation coefficient ($r$) in the first distance class and $t_m-t_s$ was tested using Spearman’s Rho.

In order to test for significance of the differences in multilocus selfing rate ($s_m$) and the biparental inbreeding level ($t_m-t_s$) between morphotypes and among populations, a nested ANOVA was performed on the bootstrap values for the estimates. The analysis was conducted with the populations nested within morphotypes. Bootstrap values were used in the parametric tests on mating system parameters after ensuring that the bootstrap output for each parameter estimate followed a normal distribution (K. Ritland, pers. comm.).

Statistical analyses

Population genetic analyses were performed using a customised version of the software package, GenAlEx (Peakall and Smouse 2001), available from http://www.anu.edu.au/BoZo/GenAlEx. When bootstrapping was used to estimate
variance of a parameter value in the population genetic analyses, confidence intervals were constructed based on the percentile method (p.25-26 Quinn and Keogh, 2002). Mating system analysis in OP flowers was conducted using the MLTR software package (Ritland 2001; http://genetics.forestry.ubc.ca/ritland/programs.html). JMP v. 3.0.1 (SAS Institute Inc.) was used for other statistical tests. Where appropriate Tukey HSD was used for multiple pairwise comparisons. For all analyses, $\alpha=0.05$. 
Results

Genetic structure at three spatial scales

**Strong differentiation between morphotypes and among populations**

Figure 3.1 shows a plot of the first two principal components of PCA on a pairwise multilocus inter-individual genetic distance. The first two coordinates accounted for 31% of inter-individual genetic variation with a pattern of individual genotypes forming two distinct clusters corresponding to NAR and BRO, within which individuals from different populations were only partially intermingled. This pattern indicated that the two morphotypes were genetically distinct and that genetic differentiation among populations within morphotypes was also substantial.

Pairwise inter-population differentiation coefficients presented in Table 3.3 showed significant differentiation in all pairwise population comparisons. Overall genetic differentiation among the seven populations was significant and substantial as shown by AMOVA without grouping populations into morphotypes ($F_{ST}=0.29$, $p<0.001$). The results of the hierarchical Analysis of Molecular Variance (AMOVA) are presented in Table 3.2. The hierarchical multi-locus AMOVA confirmed that genetic differentiation between the two morphotypes was significant ($F_{RT}=0.21$, $p=0.001$), with 25% of the total interpopulation genetic variation found between NAR and BRO. Genetic differentiation among populations within morphotypes was also significant ($F_{SR}=0.19$, $p=0.001$) and accounted for 16% of the total variation.

Mantel tests revealed positive and significant correlation between interpopulation genetic and geographic distance matrices ($r_{xy}=0.542$, $r^2=0.294$, $y=0.280x-0.890$, $p=0.04$). The correlation was also positive within each morphotype but was not significant for NAR (NAR: $r_{xy}=0.751$, $r^2=0.564$, $y=0.598x-2.579$, $p=0.17$; BRO: $r_{xy}=0.617$, $r^2=0.38$, $y=0.319x-3.351$, $p=0.04$). These results indicate a relationship between genetic and geographic distance consistent with the differentiation due to isolation by distance in a stepping-stone model.

**Strong spatial genetic structure within populations**

The outcome of combined spatial genetic autocorrelation analysis across populations at NAR and BRO is illustrated in Figure 3.2 with the correlograms showing the genetic
correlation $r_c$ as a function of distance between genotypes. For both, BRO and NAR, the correlation value is high, positive and significant in the first distance class with an x-intercept of 25 and 31 m respectively, followed by correlogram oscillations around $r_c = 0$. This pattern is consistent with significant positive small-scale spatial genetic structure for both NAR and BRO populations. Since confidence intervals around $r_c$ values overlap and the x-intercepts are similar, and consistent with a pattern of local spatial genetic structure at a similar scale of about 30 m (Fig. 3.2, Table 3.4), it can be concluded that overall patterns of local positive spatial genetic structure did not differ between the two morphotypes.

On the other hand, when populations were considered separately $r$ was positive and significant in the first distance class for five of the seven of populations, but was close to zero and not significant in one population from each morphotype (FR from BRO and TA from NAR) (Table 3.4). Where $r$ was significant, its value varied from 0.062 at B2 (BRO) to 0.471 at AR (NAR). Therefore, despite lack of overall difference between NAR and BRO, there was evidence of differences in the presence and magnitude of significant genetic substructure among G. clandestina populations.

Spatial autocorrelation analyses of the distribution of plants within populations were significant in the first distance class for both NAR and BRO indicating a patchy distribution of plants (Table 3.4). The regression analysis showed that there was a positive and strong relationship between the $r$-values and x-intercepts from the analysis of genotype and plant distribution across the seven populations, but that the result was borderline insignificant ($r^2 = 0.43, F_{1,6} = 3.811, p = 0.099$ and $r^2 = 0.44, F_{1,6} = 3.911, p = 0.095$, respectively). This suggested a trend for genetic substructuring to be correlated with the aggregation of plants, and the lack of significance was likely a result of lack of statistical power.

**Inbreeding and selfing in populations**

*Substantial inbreeding level in mature plants*

Single locus inbreeding coefficients (Table 3.5) and population inbreeding coefficients (Table 3.6) ($F_{POP}$) were positive and high in all seven populations. The bootstrap confidence intervals around single-locus $F$ did not overlap zero with three exceptions (Satt478 at B2, TE, MA, and Satt228 at TA; Table 3.5). One-tailed t-tests showed that
\( F_{\text{POP}} \), combining information across loci, was significantly greater than zero in all seven populations. Therefore all populations of \( G. \) clandestina in this study experienced a significant excess of homozygosity over that expected under Hardy-Weinberg equilibrium, indicating high rates of self-fertilisation.

Mean \( F_{\text{POP}} \)-estimates were similar for each morphotype (NAR: \( F_{\text{POP}} = 0.544 \); BRO: \( F_{\text{POP}} = 0.558 \)), but varied among populations from 0.448 at TA to 0.732 at AR. Nested ANOVA showed no significant difference in mean \( F_{\text{POP}} \) between the morphotypes (\( F_{\text{POP}}: F_{1, 38} = 0.114, p=0.7 \)), but that they varied significantly among populations nested within the morphotypes (\( F_{\text{POP}}: F_{6, 38} = 4.816, p=0.002 \)). Thus no significant difference was detected in mean inbreeding coefficient of mature plants between NAR and BRO morphotype.

Further tests on population means using Tukey HSD showed that differences among populations were limited. For \( F_{\text{POP}} \), the mean was significantly greater for AR than for B1, MA and TA (Tukey \( q=3.111, p=0.05 \)). This was a reflection of the limited variation in population inbreeding coefficients across six of the populations (\( F_{\text{POP}} 0.45-0.60 \)) with only the AR population experiencing markedly higher inbreeding levels (\( F_{\text{POP}} 0.75 \)) (Table 3.6). Therefore, there was little variation in inbreeding patterns among populations. These findings show that variation in long-term self-fertilisation rates is limited among \( G. \) clandestina populations and it does not follow the genetic and morphotypic differentiation between NAR and BRO.

**Inbreeding in mature plants reflects selfing rates within populations**

In the four populations where direct estimates of mating patterns were made, self-fertilised progeny were produced through both cleistogamy and chasmogamy (Fig. 3.3).

The contribution to selfed progeny through OP flowers was a result of substantial levels of self-fertilisation in this reproductive mode, as indicated by significant values of multilocus selfing rates (\( s_{m} \)) (Table 3.7). Selfing rates in OP flowers (\( s_{m} \)) for BRO populations were slightly but significantly higher than those for NAR populations (NAR: \( 0.52 \pm 0.02 \) vs. BRO: \( 0.62 \pm 0.01 \); \( F_{1, 400} = 39.31, p<0.0001 \)). The population effect was also significant (\( F_{2, 400} = 277.29, p<0.0001 \)) and Tukey’s HSD test indicated that differences among all of the populations were significant (Table 3.7). As the population with the highest and the lowest selfing in OP flowers both came from the NAR
morphotype \( s_m = 0.37 \pm 0.01 \) at TA and \( 0.67 \pm 0.01 \) at AR), it is unlikely that the significant difference between NAR and BRO reflects consistent differences in OP selfing between their populations.

When CL and OP selfing were combined in the estimate of the total population selfing rate \( s_{POP} \) in the four populations, there were no significant differences among them, despite the contributions of CL selfing \( s_{CL} \) and OP selfing \( s_{OP} \) differing significantly in some inter-population comparisons (Fig. 3.3).

Limited variation in the direct estimates of self-fertilisation rates among populations was consistent with trends in the mature plant inbreeding levels. The population inbreeding coefficients expected from the total selfing rates \( F_{EX} \) and inbreeding coefficients observed in mature plants \( F_{POP} \) were similar and not significantly different among the four populations (Table 3.6). Additionally, the correlation between \( F_{POP} \) and the total selfing rate in those populations was positive and significant (Spearman’s Rho = 1.0000, \( p = 0.0000 \)). On the other hand the correlation between the strength of intra-population structure \( r \) and \( F_{POP} \) was not significant (Spearman’s Rho = 0.464, \( p = 0.294 \)).

The lack of significant differences between \( F_{POP} \) and \( F_{EX} \), and the significant correlation between them, but not between \( F_{POP} \) and \( r \), provide evidence that \( F_{POP} \) accurately reflects total selfing rates in the \( G. \) clandestina populations, despite the significant genetic substructure detected within them. It also indicates that inbreeding depression does not detectably reduce the relative frequency of homozygotes arising from high levels of selfing within the populations. Therefore any differences in inbreeding coefficients among populations can be interpreted as differences in the level of inbreeding due to selfing. The inbreeding coefficient may be used to estimate long term selfing rates from the equilibrium formula, \( s = 2F_{POP}/(1+F_{POP}) \) (Jain 1979). For NAR these are \( AR = 0.85 \), \( MA = 0.62 \), \( TA = 0.62 \) and for BRO, \( B1 = 0.65 \), \( B2 = 0.75 \), \( TE = 0.72 \), \( FR = 0.73 \).

**Biparental inbreeding in outcrossed OP progeny**

Biparental inbreeding among the outcrossed OP progeny was significant in all four populations where mating patterns were estimated (Table 3.7), and it accounted for at least 6 to 14% of all OP matings \( (t_m-t_s) \) which corresponded to 12 to 33% of outcrossed
OP matings ($m_b$). The level of biparental inbreeding differed significantly between morphotypes ($t_{m-t_s}$ NAR vs BRO: $0.11\pm0.003$ vs $0.07\pm0.002$; $F_{1,400} = 11.71$, $p=0.0007$) and among populations ($F_{2,38}=7.47$, $p=0.0007$)(Table 3.7). However, pairwise inter-population comparisons using Tukey’s HSD test showed that only in the AR population the value of $t_{m-t_s}$ was significantly higher than in the other three populations (Table 3.7). Thus the biparental inbreeding in OP flowers was not consistently higher in the NAR compared to BRO populations, and the inter-population variation was limited.
Discussion

In accordance with expectations, genetic variation at microsatellite loci in *Glycine clandestina* showed patterns typical of habitual selfer: high inbreeding in mature individuals, significant genetic differentiation among populations and significant fine scale genetic structure within seven study populations. The study of four populations confirmed that the inbreeding level in populations was consistent with that expected from the observed level of OP and CL selfing. On the other hand, contrary to hypothesis based on the previous study of *G. clandestina* no consistent differences were found between populations of the two studied morphotypes, BRO and NAR in any of the measured parameters: the level of inbreeding in mature plants, the strength of genetic structure, total selfing, the level of OP selfing, and the potential for OP biparental inbreeding, despite some differences among populations.

**Inbreeding and genetic differentiation among populations**

*Selfing shapes genetic differentiation among populations in G. clandestina*

The seven *G. clandestina* populations from BRO and NAR morphotypes included in the study were significantly genetically differentiated at the 5 microsatellite loci ($F_{ST} = 0.29$). The genetic differentiation was significant both between the morphotypes (25% of total variation, $F_{HT}=0.21$) and among populations within them (15% of total variation, $F_{SR}=0.19$). It was also shown that despite the morphological discontinuity apparent between BRO and NAR populations, the degree of genetic differentiation was proportional to geographical distance among all of the populations. This was consistent with the stepping stone model of genetic differentiation via isolation by distance and genetic drift (Kimura and Weiss 1964) or genetic hitchhiking caused by a gradual cline in a local selection gradient (Linhart and Grant 1996).

Both processes, the local genetic drift and the genetic hitchhiking, are enhanced by habitual selfing because of the reduction in gene dispersal via pollen, effective population size and recombination rates (Narrain 1966; Wright 1978; Pollack 1987; Charlesworth *et al.* 1997; Nordborg and Donnelly 1997). In *G. clandestina* the habitual selfing at the rate of 62 to 85 % was apparent from the mating patterns and the inbreeding levels in the populations observed in this study. In fact the value of $F_{ST} = 0.29$ for overall inter-population differentiation in *G. clandestina* lies between those
observed for predominant outcrossers and predominant selfers at microsatellite loci (see Table 3.8), suggesting differentiation level typical of a habitually selfing plant with a mixed mating system. Thus selfing appears to be a major factor shaping inter-population genetic structure in this species.

*Comparison with other CLOP species*

In contrast to *G. clandestina*, which was shown here to be an intermediate selfer (0.10<s<0.90), the majority of previous studies of population genetics in CLOP examined predominant selfers (s>0.90), as was evident from inbreeding level in their populations (Knight and Waller 1987; Cole and Biesboer 1992; Sun 1999) or lack of genetic variation at allozyme loci (Schoen 1984; Lesica et al. 1988; Cole and Biesboer 1992). One exception is *Viola pubescens* where low level of total selfing could be surmised from predominant outcrossing in OP flowers and insignificant inbreeding in mature individuals (Culley and Wolfe 2001).

Those studies using allozyme markers in CLOP species, including *V. pubescens*, reported significant and substantial inter-population differentiation (Table 3.8). When microsatellite and allozyme markers are compared directly, they tend to reveal equivalent patterns, but differ in magnitude of differentiation (e.g. Bonnin et al. 2001; Freville et al. 2001). Therefore, it is not possible to directly compare the findings from this study of *G. clandestina* with earlier allozyme studies of other CLOP species, but the comparison of the patterns of differentiation can be made.

Across the previous allozyme-based studies, the extent of interpopulation differentiation in a species correlated with the levels of selfing experienced by its populations (Table 3.8). The extreme selfers, *Lespedeza capitata* and *Scutellaria indica*, exhibit the highest levels of interpopulation differentiation (Cole and Biesboer 1992; Sun 1999), which decreases with the selfing rate in *Impatiens capensis* (Knight and Waller 1987), to the lowest level in the predominantly outcrossing *Viola pubescens* (Culley and Wolfe 2001). Thus as in *G. clandestina*, the degree of selfing appears to be an important factor shaping the distribution of genetic variation among populations of CLOP species generally.

The study of *I. capensis* (Waller and Knight 1989a) appears to be the only other study to date that has examined relationship between population genetic and geographic distance
in a CLOP species. In contrast to the stepping stone model of differentiation in *G. clandestina*, *I. capensis* showed evidence of an island model of population differentiation suggesting that in addition to selfing, frequent extinction and colonisation had a major effect on distribution of genetic variation in this species (Waller and Knight 1989a). These contrasting genetic patterns parallel the life history differences between the two species: *Impatiens* is a colonising annual whose populations are often ephemeral (Waller and Knight 1989a), while *G. clandestina* is perennial with long-lived and stable populations.

**Limited differences between morphotypes and among populations of *G. clandestina***

**Uniform selfing and inbreeding**

An average inbreeding coefficient in mature individuals detected at microsatellite loci within seven populations amounted to $F_{pop}=0.55$ for both morphotypes, and variation among populations was limited to the tendency for greater inbreeding in one of the NAR populations. As the inbreeding coefficient was confirmed to reflect total selfing rates, this meant that in populations of both morphotypes the long term selfing accounted for on average around 71% of matings, and which ranged from 62 to 85% among populations.

Such uniform inbreeding and total selfing across populations was even more surprising, considering that selfing in OP flowers and especially the contribution of the obligatory cleistogamous selfing varied significantly among a subset of four populations. The reproductive patterns and OP mating estimates in four representative populations showed that the contribution of cleistogamy to the matings within a population ranged between 2 and 60%, and confirmed that much of the inbreeding was due to selfing in OP flowers alone, which ranged from 37 to 67%. However, there was again no evidence that the populations within the morphotypes shared distinct mating patterns. Even though OP selfing was significantly higher in BRO than NAR morphotype ($s_m$ BRO vs NAR: 0.62 vs. 0.52), that difference of 10% was small compared to the difference of 30% between the population with the highest and the lowest OP selfing rate which both were found within one, NAR morphotype ($s_m$ AR vs. TA: 0.67 vs. 0.37).

**Comparison with the previous study in *G. clandestina***

Most strikingly, this study revealed one of the NAR morphotype populations at the lowest elevation of 630 m to have the highest OP selfing rates ($s_m=0.67$) and overall
inbreeding \((F_{pop}=0.73)\) with values closest to those found in the BRO population located at a high elevation of 1500 m in Schoen and Brown’s study (1991) \((s_r=0.89, F_{pop}=0.73)\). Irrespective of their elevation, the BRO populations, together with the remaining NAR populations in this study were inbred and were selfing in OP flowers at a much lower rate, closer to the NAR population from the previous study (respectively \(F_{pop}=0.48-0.60\) vs 0.38 and \(s=0.37-0.62\) vs 0.53). From this, a general pattern emerges where in most \(G.\ clandestina\) populations OP selfing and overall inbreeding remains within relatively narrow intermediate range, with only some populations reaching substantially higher levels, with no relation either to morphotype or elevation. This is at odds with the suggestion by Schoen and Brown (1991) that increase selfing in OP flowers and inbreeding in \(G.\ clandestina\) populations may be associated with decrease in quality of pollinator service typical for higher elevations. Nevertheless, local differences in pollinator composition and behaviour may still be responsible for the higher OP selfing and the higher inbreeding observed in some of the populations of this species.

**Spatial genetic structure within populations and biparental inbreeding**

The comparison between BRO and NAR morphotype populations also revealed no significant differences in spatial autocorrelation patterns of multilocus microsatellite genotypes within \(G.\ clandestina\) populations. The two morphotypes did not differ in the significance and magnitude of \(re\) (the multipopulation spatial autocorrelation coefficient) in the first distance class, the x-intercept, and the patterns of oscillation between positive and negative values of \(re\) with increased distance. The results suggest that overall, \(G.\ clandestina\) populations have significant patchy genetic structure where individuals positioned within 30 m from each other are genetically more similar than those separated by greater physical distances.

Although on average the two morphotypes did not differ in their small scale genetic structure, substantial variation in its strength among populations within the morphotypes was apparent in the four-fold differences in the first distance class \(r\)-values among populations (BRO: \(r=0.084-0.47\) and NAR: \(r=0.062-0.226\)). Additionally, in two populations, TA from NAR and FR from BRO, the spatial autocorrelation failed to detect significant genetic structure. Thus in \(G.\ clandestina\) small-scale spatial genetic structure and its variation may be an important factor affecting selection of reproductive system within populations, irrespective of their morphotype.
One important consequence of spatial aggregation of genetically related individuals in conjunction with limited pollen dispersal typical in plants is matings between relatives or biparental inbreeding, which reduces the advantage of outcrossing from inbreeding avoidance (Uyenoyama 1986; Waller 1993). The analysis of OP mating detected significant biparental inbreeding in all four populations that were studied as representative of the two morphotypes. Across these populations the minimum estimate of the proportion of OP outcross matings between relatives was uniform at around $m_6 = 16\%$, with one exception of the AR population from NAR where it reached $m_6 = 33\%$. Thus as expected from small scale genetic structure analysis the quality of OP outcross matings was apparently affected by spatial aggregation of genotypes within those *G. clandestina* populations, independent of the morphotype. However, in an apparent contradiction biparental inbreeding was detected in the two populations, FR and TA, where autocorrelation analysis failed to detect significant small-scale spatial genetic structure. One reason for the failure of autocorrelation to detect structure is a discrepancy between the scale of sampling and analysis and the actual scale of patch size (Heywood 1991). As 10 m was the minimum size of distance class for which the spatial autocorrelation analysis was possible in this study, the significant structure evident from biparental inbreeding is likely present at the scales $<10$ m in those two populations.

The small-scale genetic structure of the nature found in *G. clandestina* has been predicted to occur in plants solely due to restriction of gene flow via pollen and seed movement typical of plants, i.e. isolation by distance (Turner *et al.* 1982; Sokal and Wartenberg 1983; Sokal and Jacquez 1991; Epperson 1995). Reviews of empirical studies show that small-scale genetic structure is characteristic of populations in a wide range of plants and that, similarly to the genetic structure at the inter-population level, selfing is a major factor enhancing such genetic structure within populations (Epperson 1989; Ennos 2001; Vekemans and Hardy 2004). The parallels in the overall inbreeding and the within-population genetic structure between BRO and NAR morphotypes detected in *G. clandestina* are in agreement with those observations.

However, variation in small-scale spatial genetic structure among populations was not significantly correlated with the variation in inbreeding across populations. This was not
a surprise as inbreeding was relatively similar across populations while the \( r \) varied widely. In contrast to inbreeding, a strong trend for positive correlation between the strength and the scale of genetic structure and the spatial distribution of individuals indicated that it is important in contributing to the observed variation in population structure. Spatial patchiness of individual distribution can affect the formation of genetic structure directly due enhancing the effects of isolation by distance (Doligez et al. 1998). Additionally, if the spatial distribution of individuals corresponds to patchiness of suitable habitat within G. clandestina populations and if selection differs between the patches, hitchhiking across the genotypes may also affect the scale and degree of differentiation at neutral loci, the effect thought to be particularly strong in selfers (Epperson 1989; Linhart and Grant 1996). Finally, the relatively small spatial scale of genetic structure, such as may the case for FR and TA populations in this study, is expected in populations recovering from a recent disturbance of the equilibrium (Slatkin 1993; Hardy and Vekemans 1999). In G. clandestina fire can cause decimation within populations and stimulate germination of dormant seed to produce such effect. Thus while selfing, the main factor likely to affect within population substructuring, may be relatively uniform across studied populations of G. clandestina, there is evidence that other factors such as patchiness of individual distribution or recent disturbance event cause some interpopulation variation.

Comparison with other CLOP species

Very little is known about within population structure in other CLOP species, and this is the first study that examined it using molecular markers. Explicit studies of within population genetic structure in CLOP are also confined to Impatiens species, where the presence of strong genetic structure at scales from 1 to >30 m has been well established by exploration of morphological characters and local adaptation (Shoemaker and Waller unpubl. cited in Schemske 1984; Schoen and Latta 1989; Waller and Knight 1989b; Argyres and Schmitt 1992; Paoletti and Holsinger 1999; Donohue et al. 2000; Heschel et al. 2002). Structure in Impatiens populations has been shown to arise both from local selection (Argyres and Schmitt 1991; Paoletti and Holsinger 1999; Donohue et al. 2000; Heschel et al. 2002) and genetic drift (Schemske 1984; Knight and Waller 1987; Schoen and Latta 1989). Similarly to this study of G. clandestina, the population substructuring in Impatiens is associated with biparental inbreeding in OP flowers (Waller and Knight 1989a; Stewart 1994; Lu 2000). Estimates of the proportion of outcrossed matings to relatives (\( m_b \)) in Impatiens range from 0.08 to 0.52 (Waller and
Knight 1989a; Lu 2000), which is comparable to the range of 0.12 to 0.33 found across populations in *G. clandestina*. On the other hand, a study of OP mating patterns in another CLOP, *Viola pubescens*, has revealed no significant biparental inbreeding (Culley 2002). This indicates that in this predominant outcrosser and in contrast to intermediate to high inbreeders *Impatiens* and *G. clandestina*, kinship genetic structure does not have a significant effect on the quality of OP matings and the chasmogamy advantage due to inbreeding avoidance.

**Implications for the evolutionary dynamics of CLOP in G. clandestina**

The patterns of inbreeding and population genetic structure in *G. clandestina* have several important implications for the traditionally recognised advantage to OP reproduction derived from the avoidance of inbreeding and inbreeding depression.

Substantial selfing combined with biparental inbreeding in OP flowers of *G. clandestina* diminishes differences in inbreeding between open pollinated chasmogamous progeny and obligately self-fertilised cleistogamous progeny. Thus any differences in fitness that may favour chasmogamy over cleistogamy due to inbreeding depression are likely to be substantially reduced in *G. clandestina*. The extent of this reduction was probably underestimated here because the estimate of biparental inbreeding used in this study, although popular, in theory only gives the minimum of the actual biparental inbreeding rates (Brown 1989; Leclerc-Potvin and Ritland 1994; Ritland 2002). Griffin and Eckert (2003) have recently shown that this estimate may be a magnitude smaller than more direct estimates in the study of *Aquilegia canadensis* (3 vs 30%). If a similar difference applies in *G. clandestina* biparental inbreeding could be as high as 80%. This, in combination with average selfing rates of 57%, will produce highly inbred OP progeny.

The importance of inbreeding in OP flowers of CLOP species for maintaining the evolutionary balance between the two modes of reproduction is now widely appreciated. Furthermore, detailed studies of OP inbreeding in *Impatiens* have demonstrated that inter-plant variation is an important characteristic of both biparental inbreeding and selfing (Stewart 1994; Lu 2000; Lu 2002). In fact, modelling incorporating selfing in OP flowers as a dynamic variable in form of geitonogamous selfing that increases with plant size shows that it may be responsible for the overall stability of the mixture of cleistogamy and chasmogamy (Masuda et al. 2001). Such effect of floral display on selfing rate in OP flowers has been demonstrated in *Impatiens* (Stewart 1994). Since in
G. clandestina selfing in OP flowers is almost solely a result of geitonogamy (Chapter 2), it is likely that the relative advantage of producing seeds via chasmogamy may also decrease with plant size. As plant size also profoundly affects the allocation to the two reproductive modes (Chapter 2), this further emphasises that the future investigations of size-related variation in advantage to chasmogamy may be of interest in G. clandestina.

Any advantage of chasmogamy derived from the avoidance of inbreeding hinges on the substantial level of depression in fitness of selfed relative to outcrossed progeny within populations. Much of this inbreeding depression is thought to be due to expression of deleterious recessive mutations in homozygous inbred progeny (dominance-based inbreeding depression), which under habitual inbreeding are exposed to selection and removed (Lande and Schemske 1985; Charlesworth and Charlesworth 1987b). The congruence between expected and observed inbreeding coefficient in G. clandestina in this study suggests the inbred progeny does not experience enhanced mortality expected under inbreeding depression (Ritland 1990a). Given the habitual selfing in this species, this suggests that at least purging of deleterious alleles affecting vitality has occurred within these G. clandestina populations, and that the overall level of inbreeding depression is likely to be reduced. Since they experience relatively uniform levels of inbreeding, the reduction in inbreeding depression should be similar across populations.

The existence of the small-scale genetic structure within populations of G. clandestina could also affect the expression of inbreeding depression. In addition to dominance inbreeding depression, some of the inbreeding depression is thought to have basis in the superior fitness of heterozygotes over homozygotes or heterosis (overdominance inbreeding depression). Inbreeding depression due to overdominance cannot be purged and, importantly, may be enhanced under genetic differentiation. Because isolation by distance causes random fixation of different alleles within patches, the outcrossing events between inbred patches cause the progeny to be superior due to heterosis (Fenster 1991b; Byers 1998). Thus, the advantage to chasmogamy derived from avoidance of inbreeding may be enhanced by inter-patch outcrossing in the structured populations of G. clandestina. Variation in the strength and scale of the kinship structure may affect the importance of this effect across populations.
From the inter-morphotypic comparisons it is clear that these effects do not differ between the two morphotypes of *G. clandestina* included in this study in line with lack of differences in reproductive strategy described in Chapter 2. The variation among populations is also likely to be limited, mainly due to uniform level of overall inbreeding and total selfing, which should in turn limit variation in purging and kinship structure. Only narrow interpopulation variation in kinship structure as well as biparental inbreeding and selfing in OP flowers as well as kinship structure not related to total inbreeding should affect the chasmogamy advantage due to inbreeding avoidance. For example, it is unlikely that the dramatically low proportion of cleistogamy observed in FR population from BRO morphotype would be an outcome of increased advantage to chasmogamy due to variation in any of the discussed factors. As the geographic and habitat range included in the study did not exhaust the range of this species, naturally it is likely that a greater range these factors vary more broadly in *G. clandestina*, with effects on the balance between cleistogamy and chasmogamy.

**Conclusions**

This study has confirmed intermediate to high levels of inbreeding in populations of *G. clandestina*, resulting from the combination of chasmogamy and cleistogamy. The observed degree of differentiation among populations is comparable with other selfing species with the patterns consistent with isolation by distance or gradual selection gradient. Inbreeding in *G. clandestina* is at the lower end of the range than most of the CLOP species where population genetic structure was studied to date.

No support was found for the hypothesis based on the previous study by Schoen and Brown (1991) that populations of broad leaf (BRO) morphotype may experience higher level of OP selfing, greater level of inbreeding, enhanced small-scale spatial genetic structure with greater potential for biparental inbreeding than populations of narrow leaf (NAR) morphotype. In combination the results of this and the previous study indicate that some populations of *G. clandestina* may experience higher selfing and inbreeding than others, with no consistent pattern in relation to either morphotype or elevation. Further studies are necessary to reveal the ecological correlates of such variation. Surprisingly, the potential for striking differences in CL selfing found in Chapter 2 was not reflected in variation in the inbreeding level, which dictates caution in interpretation.
of wide variation in cleistogamy among populations typical for CLOP species as evidence for differences in overall inbreeding.

In agreement with the lack of intermorphotypic differences in inbreeding there was no evidence for difference in the small-scale genetic structure. Combined evidence from autocorrelation and biparental inbreeding showed that all populations had significant kinship structure but that its scale and strength varied among populations. Inbreeding contributed little to these inter-population differences, and rather they were related to the spatial distribution of individuals, possibly reflecting heterogenous selection environment. Together with studies in Impatiens these finding confirm that small scale genetic structure constitutes an important element of the genetic background that likely influences the balance between chasmogamy and cleistogamy in CLOP species with intermediate to high inbreeding.

It appears that in G. clandestina the advantage to chasmogamy from the avoidance of inbreeding depression is substantially diminished due to the high rates of selfing and biparental inbreeding in chasmogamous flowers. Purging of dominance inbreeding depression is likely under the substantial habitual inbreeding of this species. At the same time the existence of population genetic structure may counteract these effects by increased expression of overdominance based expression in inter-patch outcrosses. While this study showed that these effects on chasmogamy advantage do not differ between broad and narrow leaf morphotypes of G. clandestina, there is restricted variation among the studied populations. Further investigations of the dynamics of relative advantage to chasmogamy derived from outcrossing would benefit from the studies of the contradictory effects of small-scale population genetic structure via biparental inbreeding vs. outbreeding and variation of OP selfing and biparental inbreeding with individual plant size. However, before the complexities of these interactions can be explored in G. clandestina it is important to establish the nature and magnitude of a key factor in this interplay - inbreeding depression. I explore this question in the next chapter (Chapter 4).
Table 3.1  SSR loci used in the study of *G. clandestina*, transferred from *G. max*. Unless otherwise indicated, the SSR unit structure was inferred from SSR sequence in *G. max* in conjunction with allele size variation in *G. clandestina*. Np - number of populations where marker was polymorphic; N - number of genotyped individuals; Na - total number of alleles; He - expected heterozygosity. For detailed description of loci see Hempel and Peakall (2003).

<table>
<thead>
<tr>
<th>Locus</th>
<th>SSR unit</th>
<th>Np</th>
<th>N</th>
<th>Na</th>
<th>He</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt478</td>
<td>tri-nuc</td>
<td>7</td>
<td>236</td>
<td>7</td>
<td>0.75</td>
</tr>
<tr>
<td>Sat040</td>
<td>di-nuc</td>
<td>7</td>
<td>231</td>
<td>27</td>
<td>0.86</td>
</tr>
<tr>
<td>AG48</td>
<td>(AG)n</td>
<td>7</td>
<td>232</td>
<td>11</td>
<td>0.80</td>
</tr>
<tr>
<td>Satt462</td>
<td>tri-nuc</td>
<td>7</td>
<td>247</td>
<td>18</td>
<td>0.83</td>
</tr>
<tr>
<td>Sat131</td>
<td>(AT)n(T)2A</td>
<td>7</td>
<td>251</td>
<td>15</td>
<td>0.87</td>
</tr>
<tr>
<td>Sat064</td>
<td>(AT)n</td>
<td>2</td>
<td>51</td>
<td>2</td>
<td>0.46</td>
</tr>
<tr>
<td>Satt163</td>
<td>tri-nuc</td>
<td>2</td>
<td>51</td>
<td>16</td>
<td>0.88</td>
</tr>
<tr>
<td>Satt373</td>
<td>(ATT)3</td>
<td>2</td>
<td>47</td>
<td>2</td>
<td>0.40</td>
</tr>
<tr>
<td>Satt199</td>
<td>(AT)n</td>
<td>2</td>
<td>45</td>
<td>9</td>
<td>0.78</td>
</tr>
<tr>
<td>Satt155</td>
<td>tri-nuc</td>
<td>2</td>
<td>48</td>
<td>2</td>
<td>0.46</td>
</tr>
<tr>
<td>Satt350</td>
<td>tri-nuc</td>
<td>2</td>
<td>48</td>
<td>7</td>
<td>0.61</td>
</tr>
<tr>
<td>Satt288</td>
<td>tri-nuc</td>
<td>2</td>
<td>44</td>
<td>7</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Mean (SE) 10.25 (2.19)

1 – Locus sequenced in *G. clandestina* by Hempel and Peakall (2003)
2 – Locus sequenced in *G. clandestina* by Peakall et al. (1998)
3 – Size homoplasies present between BRO and NAR morphotype, A – repeat structure in NAR, B – repeat structure in BRO.

Table 3.2  Results of hierarchical Analysis of Molecular Variance (AMOVA) describing genetic differentiation among seven populations of *Glycine clandestina* representing the two morphotypes – NAR and BRO, based on microsatellite genotypes at five loci from 148 individuals. All differentiation measure values were significantly greater than zero and were significant at <0.05 level based on 1000 permutations. The values of F-statistics are estimates of the differentiation between morphotypes ($F_{RT}$), the differentiation among populations within morphotypes ($F_{SR}$), and the differentiation among populations and morphotypes ($F_{ST}$).

<table>
<thead>
<tr>
<th></th>
<th>$F_{RT}$</th>
<th>$F_{SR}$</th>
<th>$F_{ST}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multilocus</td>
<td>0.21</td>
<td>0.19</td>
<td>0.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus</th>
<th>$F_{RT}$</th>
<th>$F_{SR}$</th>
<th>$F_{ST}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt478</td>
<td>0.33</td>
<td>0.23</td>
<td>0.48</td>
</tr>
<tr>
<td>Sat040</td>
<td>0.28</td>
<td>0.05</td>
<td>0.31</td>
</tr>
<tr>
<td>AG48</td>
<td>0.33</td>
<td>0.37</td>
<td>0.58</td>
</tr>
<tr>
<td>Satt462</td>
<td>0.24</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>Sat131</td>
<td>0.14</td>
<td>0.21</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Table 3.3  
Pairwise inter-population genetic differentiation ($F_{ST}$) at five microsatellite loci in *G. clandestina*. All values are significantly greater than zero at 0.05 level based on 1000 permutations.

<table>
<thead>
<tr>
<th></th>
<th>B1</th>
<th>B2</th>
<th>FR</th>
<th>TE</th>
<th>AR</th>
<th>MA</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>0.117</td>
<td>0.077</td>
<td>0.155</td>
<td>0.248</td>
<td>0.216</td>
<td>0.399</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>0.178</td>
<td>0.132</td>
<td>0.346</td>
<td>0.331</td>
<td>0.540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>0.247</td>
<td>0.353</td>
<td>0.351</td>
<td>0.542</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>0.379</td>
<td>0.360</td>
<td>0.562</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td></td>
<td>0.123</td>
<td>0.308</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td></td>
<td>0.258</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.4  
Results of spatial autocorrelation analysis of individual multilocus genotypes as well as plant distribution using distance class size of 10 m. The first distance class analysis is shown for each population and populations combined within each morphotype (All rc) where : n - number of pairwise comparisons; r - spatial autocorrelation coefficient; rc - spatial autocorrelation coefficient in analysis combined across populations; x-intercept – the distance of the first intercept of autocorrelogram with x-axis, which reflects the patch size. All significant autocorrelation values are in bold. Significance of autocorrelation and 95% confidence intervals (CI) around null hypothesis of $r=0$ were determined based on 1000 permutations.

<table>
<thead>
<tr>
<th></th>
<th>BRO populations</th>
<th></th>
<th>NAR populations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (rc)</td>
<td>B2</td>
<td>TE</td>
<td>FR</td>
</tr>
<tr>
<td>Autocorrelation of genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>91</td>
<td>22</td>
<td>12</td>
<td>49</td>
</tr>
<tr>
<td>r</td>
<td>0.108</td>
<td>0.062</td>
<td>0.226</td>
<td>0.009</td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td>0.053</td>
<td>0.059</td>
<td>0.124</td>
<td>0.043</td>
</tr>
<tr>
<td>Lower 95% CI</td>
<td>-0.050</td>
<td>-0.127</td>
<td>-0.212</td>
<td>-0.119</td>
</tr>
<tr>
<td>Prob $r &gt;$ permuted $r$</td>
<td>0.001</td>
<td>0.029</td>
<td>0.001</td>
<td>0.122</td>
</tr>
<tr>
<td>X - intercept</td>
<td>31</td>
<td>20</td>
<td>36</td>
<td>22</td>
</tr>
</tbody>
</table>

Autocorrelation of plant distribution

|             |             |             |             |             |             |             |             |             |             |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| n           | 1052        | 296         | 184         | 191         | 381         | 668         | 248         | 229         | 191         |
| r           | 0.076       | 0.132       | 0.027       | 0.052       | 0.068       | **0.176**   | **0.220**   | 0.024       | **0.306**   |
| Upper 95% CI| 0.063       | 0.130       | 0.147       | 0.119       | 0.119       | 0.075       | 0.121       | 0.129       | 0.149       |
| Lower 95% CI| -0.059      | -0.127      | -0.148      | -0.099      | -0.080      | -0.069      | -0.113      | -0.116      | -0.134      |
| Prob $r >$ permuted $r$ | 0.030 | 0.030 | 0.360 | 0.180 | 0.090 | **0.001** | **0.001** | 0.336 | **0.001** |
| X-intercept | 47          | 49          | 63           | 15          | 49          | 19          | 62          | 13          | 18          |
Table 3.5  Single-locus inbreeding coefficients for seven populations of *G. clandestina* based on mature plant genotypes at microsatellite loci. CI - 95% confidence intervals based on 1000 bootstraps; *F* - inbreeding coefficient; NAR - narrow leaf morphotype; BRO - broad leaf morphotype.

<table>
<thead>
<tr>
<th>Population</th>
<th>Satt478</th>
<th>Sat040</th>
<th>Ag48</th>
<th>Satt462</th>
<th>Sat131</th>
<th>Satt163</th>
<th>Satt373</th>
<th>Satt155</th>
<th>Satt350</th>
<th>Satt228</th>
<th>Sat064</th>
<th>Satt199</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAR AR F</td>
<td>0.569</td>
<td>0.773</td>
<td>0.742</td>
<td>0.763</td>
<td>0.817</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(0.295-0.795)</td>
<td>(0.516-0.957)</td>
<td>(0.541-0.910)</td>
<td>(0.578-0.934)</td>
<td>(0.656-0.961)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA F</td>
<td>0.266</td>
<td>0.598</td>
<td>0.532</td>
<td>0.469</td>
<td>0.404</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(-0.028-0.573)</td>
<td>(0.323-0.847)</td>
<td>(0.289-0.742)</td>
<td>(0.249-0.696)</td>
<td>(0.173-0.611)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA F</td>
<td>0.489</td>
<td>0.576</td>
<td>0.377</td>
<td>0.591</td>
<td>0.334</td>
<td>0.336</td>
<td>0.474</td>
<td>0.4</td>
<td>0.558</td>
<td>0.361</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(0.287-0.673)</td>
<td>(0.341-0.801)</td>
<td>(-0.035-0.727)</td>
<td>(0.304-0.823)</td>
<td>(0.112-0.553)</td>
<td>(0.108-0.570)</td>
<td>(0.061-0.828)</td>
<td>(0.017-0.754)</td>
<td>(0.178-0.868)</td>
<td>(-0.026-1.071)</td>
<td></td>
</tr>
<tr>
<td>BRO B1 F</td>
<td>0.325</td>
<td>0.522</td>
<td>0.556</td>
<td>0.281</td>
<td>0.727</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(0.035-0.602)</td>
<td>(0.285-0.750)</td>
<td>(0.231-0.858)</td>
<td>(0.068-0.508)</td>
<td>(0.474-0.945)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2 F</td>
<td>0.298</td>
<td>0.655</td>
<td>0.681</td>
<td>0.663</td>
<td>0.734</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(-0.138-0.744)</td>
<td>(0.447-0.852)</td>
<td>(0.312-0.941)</td>
<td>(0.502-0.803)</td>
<td>(0.571-0.871)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE F</td>
<td>0.556</td>
<td>0.525</td>
<td>0.684</td>
<td>0.488</td>
<td>0.572</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(-0.063-1.034)</td>
<td>(0.293-0.754)</td>
<td>(0.283-1.011)</td>
<td>(0.207-0.754)</td>
<td>(0.325-0.807)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR F</td>
<td>0.538</td>
<td>0.676</td>
<td>0.64</td>
<td>0.604</td>
<td>0.658</td>
<td>0.545</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.505</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>(0.262-0.757)</td>
<td>(0.543-0.805)</td>
<td>(0.430-0.827)</td>
<td>(0.355-0.842)</td>
<td>(0.471-0.838)</td>
<td>(0.249-0.804)</td>
<td>(0.096-0.889)</td>
<td>(0.255-0.672)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.6  Mean inbreeding coefficients observed in mature plants ($F_{POP}$ – mean population inbreeding coefficient) and those expected from mating patterns within populations in G. clandestina. Means that do not significantly differ among populations using Tukey HSD test at alpha=0.05 are marked with the same letter. All observed inbreeding coefficients were significantly greater than zero using t-test at alpha 0.05. n – number of loci; CI – 95% confidence intervals; $F$ – inbreeding coefficient.

<table>
<thead>
<tr>
<th>Population</th>
<th></th>
<th>$F_{observed}$</th>
<th>$F_{expected from}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_{POP}$</td>
<td>CL-selfing</td>
<td>CL and OP selfing</td>
</tr>
<tr>
<td>NAR AR</td>
<td>$0.732^a$</td>
<td>0.43</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>(0.649 - 0.816)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA F</td>
<td>$0.453^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.342 - 0.565)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TA F</td>
<td>$0.449^b$</td>
<td>0.3</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>(0.387-0.512)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>$0.544$</td>
<td>0.365</td>
<td>0.625</td>
</tr>
<tr>
<td>BRO B1</td>
<td>$0.482^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.322-0.641)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2 F</td>
<td>$0.606^ab$</td>
<td>0.11</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>(0.453-0.759)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TE F</td>
<td>$0.564^ab$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.500-0.629)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR F</td>
<td>$0.579^ab$</td>
<td>0.01</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(0.526-0.632)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>$0.558$</td>
<td>0.06</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Table 3.7  Maximum likelihood estimates of mating system parameters for open flowers (OP) based on microsatellite genotypes in 15 families (maternal parent and 5-6 seeds) within each of four G. clandestina populations. $s_m$ is multilocus selfing rate ($s_m=1-t_m$); $t_m$ is multilocus outcrossing rate; $t_s$ is a minimum variance single locus outcrossing rate; $t_m-t_s$ is a minimum estimate of biparental inbreeding rate; $m_b$ is a minimum proportion of biparentally inbred progeny among outcrosses from $(t_m-t_s)/t_s$ (Ritland, 2001). Values that significantly differ among populations at $\alpha=0.05$ according to Tukey HSD test on 500 bootstrap values are marked with different letters.

<table>
<thead>
<tr>
<th>Morphotype</th>
<th>Population</th>
<th>Loci</th>
<th>$s_m$ (SE)*</th>
<th>$t_m - t_s$ (SE)*</th>
<th>$m_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAR AR</td>
<td>SAT373, SATT478, AG148, SATT350</td>
<td>0.667 (0.008) a</td>
<td>0.141 (0.002) a</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td>SAT373, SATT478, AG148, SATT163</td>
<td>0.373 (0.010) b</td>
<td>0.075 (0.004) b</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>BRO B2</td>
<td>SAT131, SAT064, SAT462</td>
<td>0.624 (0.007) c</td>
<td>0.092 (0.002) b</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>SAT064, SAT199, SAT131</td>
<td>0.607 (0.006) d</td>
<td>0.051 (0.003) b</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>

*All values significantly greater than zero according to one tailed t-test on 500 bootstraps, $\alpha=0.05$. 

116
Table 3.8  A review of majority of published inter-population differentiation studies based on microsatellites in plants as well as allozymes in CLOP plants for comparison with the results of this study in *G. clandestina*. Selfing estimates were either calculated from the Wrights $F_{ST}$ using equilibrium formula (marked with *) or given by the authors of the study (unmarked, out – judged to be predominantly outcrossing).

<table>
<thead>
<tr>
<th>Species</th>
<th>Differentiation among populations ($F_{ST}$)</th>
<th>Selfing rate</th>
<th>No. populations</th>
<th>No. loci</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsatellites</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUTCROSSERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Fraxinus excelsior</em></td>
<td>0.09</td>
<td>0.03</td>
<td>10</td>
<td>6</td>
<td>(Heuertz <em>et al.</em> 2001)</td>
</tr>
<tr>
<td><em>Melaleuca alternifolia</em></td>
<td>0.07</td>
<td>0.14*</td>
<td>20</td>
<td>5</td>
<td>(Rossetto <em>et al.</em> 1999)</td>
</tr>
<tr>
<td><em>Centaurea corymbosa</em></td>
<td>0.23</td>
<td>out/rare</td>
<td>6</td>
<td>6</td>
<td>(Freville <em>et al.</em> 2001)</td>
</tr>
<tr>
<td><em>Gymnadenia conopsea</em></td>
<td>0.06</td>
<td>out/rare</td>
<td>10</td>
<td>3</td>
<td>(Gustafsson 2000)</td>
</tr>
<tr>
<td><strong>MIXED MATING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Avicennia marina</em></td>
<td>0.41</td>
<td>0.29*</td>
<td>13 (worldwide)</td>
<td>3</td>
<td>(Maguire <em>et al.</em> 2002)</td>
</tr>
<tr>
<td><em>Mimulus guttatus</em></td>
<td>0.32</td>
<td>0.51</td>
<td>3</td>
<td>6</td>
<td>(Awadalla and Ritland 1997)</td>
</tr>
<tr>
<td><em>Grevillea macleayana</em></td>
<td>0.22</td>
<td>mixed</td>
<td>8</td>
<td>6</td>
<td>(Englund <em>et al.</em> 2002)</td>
</tr>
<tr>
<td><em>Glycine clandestina</em></td>
<td>0.29</td>
<td>0.71</td>
<td>7</td>
<td>5</td>
<td>This study</td>
</tr>
<tr>
<td><strong>SELFERS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Bulinus truncatus</em>#</td>
<td>0.54</td>
<td>&gt; 0.82*</td>
<td>22</td>
<td>4</td>
<td>(Viard <em>et al.</em> 1997)</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>0.64</td>
<td>&gt;0.90</td>
<td>11</td>
<td>3</td>
<td>(Bergelson <em>et al.</em> 1998)</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>0.64</td>
<td>&gt;0.90</td>
<td>2</td>
<td>5</td>
<td>(Kuittinen <em>et al.</em> 1997)</td>
</tr>
<tr>
<td><em>Medicago truncata</em></td>
<td>0.37</td>
<td>&gt;0.90</td>
<td>3</td>
<td>5</td>
<td>(Bonnin <em>et al.</em> 2001)</td>
</tr>
<tr>
<td><em>Mimulus lacinatus</em></td>
<td>0.6</td>
<td>0.83*</td>
<td>3</td>
<td>6</td>
<td>(Awadalla and Ritland 1997)</td>
</tr>
<tr>
<td><em>Elymus fibrosus</em></td>
<td>0.54</td>
<td>&gt;0.80</td>
<td>10</td>
<td>6</td>
<td>(Sun <em>et al.</em> 1998)</td>
</tr>
<tr>
<td><strong>Allozymes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CLOP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Impatiens capensis</em></td>
<td>0.46</td>
<td>0.72*</td>
<td>11</td>
<td>N/A</td>
<td>(Waller and Knight 1989b)</td>
</tr>
<tr>
<td><em>Lespedeza capitata</em></td>
<td>0.51</td>
<td>0.98</td>
<td>12</td>
<td>N/A</td>
<td>(Cole and Biesboer 1992)</td>
</tr>
<tr>
<td><em>Scutellaria indica</em></td>
<td>0.92</td>
<td>1.00*</td>
<td>20</td>
<td>N/A</td>
<td>(Sun 1999)</td>
</tr>
<tr>
<td><em>Viola pubescens</em></td>
<td>0.29</td>
<td>0*</td>
<td>6</td>
<td>N/A</td>
<td>(Culley and Wolfe 2001)</td>
</tr>
<tr>
<td>#Snail.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 3.1 Principal coordinate analysis (PCA) plots of first two coordinates for pairwise inter-individual genetic distance. Percent of variation explained by each coordinate is given. Individuals from BRO populations (B1, B2, FR, TE) are visualised using solid symbols and delineated with solid lines. Individuals from NAR populations (AR, MA, TA) are visualised using line symbols and delineated with dashed lines.

Fig. 3.2 Correlograms reflecting patterns in spatial autocorrelation of multilocus microsatellite genotypes in combined analysis across (A) BRO and (B) NAR populations, with distance class size of 10 m. The solid line represents the correlogram with broken lines corresponding to upper and lower limits of 95% confidence interval around null hypothesis derived from 1000 permutations. Error bars around values for each distance class are 95% confidence intervals based on 1000 bootstraps; n - number of individuals in a distance class.
Fig. 3.3 Proportion of population reproductive output (proportion of fruit) contributed by self-fertilisation in CL mode, OP mode and total selfing (combination of selfing through cleistogamy and OP). The error bars represent 95% confidence intervals based on binomial error distribution. Means that differ significantly between populations based on comparison of 95% confidence intervals are marked by different letters. Numbers in brackets give the total number of fruit.
Chapter 4

Non-genetic effects on selfed seed from cleistogamous flowers cause better seedling performance than outcrossing in chasmogamous flowers of Glycine clandestina
Introduction

The depression in fitness of inbred relative to outcrossed progeny (inbreeding depression) is considered a central force favouring outcrossing in plant mating systems (Lande and Schemske 1985; Campbell 1986; Uyenoyama 1986; Charlesworth and Charlesworth 1987b; Holsinger 1988; Holsinger 1991; Damgaard et al. 1992; Jarne and Charlesworth 1993; Barrett 1995; Ronfort and Couvet 1995; Sakai 1995; Barrett and Harder 1996). Inbreeding depression is also thought to favour chasmogamous reproduction in systems that mix obligatory cleistogamous selfing with at least partial chasmogamous outcrossing (Darwin 1897; Schoen and Lloyd 1984; Masuda et al. 2001). However, the dynamics of ID can be complex and there is much uncertainty about its importance in habitual selfers, like CLOP, mainly because it is predicted to be subject to purging in selfing populations (Schemske and Lande 1985; Charlesworth et al. 1990; Husband and Schemske 1996; Byers and Waller 1999; Crnokrak and Barrett 2002).

Because ID is important to understanding of the balance between cleistogamy and chasmogamy, many studies have attempted to determine its the magnitude in CLOP species (Cope 1966; Wilken 1982; Clay 1983; Schoen 1984; Waller 1984; Clay and Antonovics 1985; Mitchell-Olds and Waller 1985; Antlfinger 1986; Trapp and Hendrix 1988; Gara and Muenchov 1990; Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990; McCall et al. 1994; Berg and Redbo-Torstensson 1999; Berg 2000b; Culley 2000). While there is evidence for the expected superiority of OP to CL progeny in some CLOP species (Waller 1984; Mitchell-Olds and Waller 1985; Gara and Muenchov 1990; Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990), in others little or no fitness difference has been observed (Wilken 1982; Antlfinger 1986; Trapp and Hendrix 1988; Berg and Redbo-Torstensson 1999). Some of this ambiguity in the evidence as to the importance of ID in CLOP species may be due to methodological bias.

Most these studies have compared CL and OP progeny to detect ID (but see Bryan and Muenchow 1991; Culley 2000). This is convenient because floral manipulations are not required. However, such comparisons rely on two assumptions: (1) OP progeny are outcrossed, and (2) reproductive mode has no effect on progeny quality. Both of these assumptions are likely to be violated in CLOP. The assumption that OP pollination
produces only outcrossed progeny is likely untrue because OP flowers are self-fertile in CLOP (Lord 1981) and because the estimates of mating show that large proportion of progeny in OP flowers results from inbreeding (Chapter 3, Wilken 1982; Mitchell-Olds and Waller 1985; Waller and Knight 1989a; Stewart 1994; Lu 2000; Culley 2002). A failure to account for inbreeding in OP flowers will underestimate the magnitude of ID. Furthermore, even when outcrossed OP progeny are ensured by hand manipulation (e.g. Wilken 1982; Schmitt and Gamble 1990), comparisons with cleistogamy may be confounded by the effects of reproductive mode. This is possible because seeds produced via the two reproductive modes often develop at different positions within a plant or at distinct times during the season. Consequently, they can experience differences in resource provisioning or seed maturation environment (pre-conditioning), leading to non-genetic differences in seed characteristic (Baskin and Baskin 1998, p.214). One of the seed characteristics that often differ systematically between cleistogamy and chasmogamy is seed mass (Campbell et al. 1983; Schoen and Lloyd 1984; Cheplick 1994).

It is well recognised that seed mass, the earliest juvenile component of fitness, can be affected by both inbreeding depression and non-genetic differences in resource allocation (Carr and Dundash 1995; Norman et al. 1995; Ramsey and Vaughton 1996; Affre and Thompson 1999; Sassaki and Felippe 1999; Wennstrom et al. 2002). Larger seed mass usually correlates with greater seedling size and juvenile performance (Harper 1977; Willson 1983). While the effects of inbreeding depression on seed size and juvenile performance would be expected to disadvantage self-fertilised CL compared with outcrossed OP progeny, the reproductive mode effects may act in favour of either of the progeny types. Thus, comparisons of CL and OP progeny to assess ID confound the effects of mating type and reproductive mode, especially when early juvenile performance is considered.

Similarly to many other CLOP species, *G. clandestina* typically produces three types of progeny: outcrossed OP and self-fertilised CL and OP, because of the intermediate selfing in OP flowers (Schoen and Brown, 1991, and Chapter 3) and coexistence of cleistogamy with chasmogamy (see Chapter 2). Additionally, *G. clandestina* is characterised by a consistent difference in seed mass in favour of CL seeds (Chapter 2). Therefore both ID and reproductive mode are likely to affect early progeny performance.
in *G. clandestina*. The difference in seed mass between cleistogamy and chasmogamy suggests that non-genetic effects of reproductive mode should enhance performance of seedlings derived from selfing in CL flowers compared to selfing in OP flowers. On the other hand because of the habitual selfing in *G. clandestina* (Chapter 3) and potential purging of ID, little difference in seedling performance is expected in favour of outcrossing over selfing.

In this chapter the effects of reproductive mode and the type of mating on early progeny performance were independently assessed in a population of *G. clandestina*. The comparison of CL self with OP self progeny was made to determine the reproductive mode effect and OP self with OP outcross progeny were compared to show the effect of the mating type. A similar design has been used previously in two other studies (Bryan and Muenchow 1991; Culley 2000) but this is the first CLOP species with seed size dimorphism to apply this approach. The two mating types (self and outcross) among OP progeny were identified *a posteriori* in naturally pollinated seed arrays using microsatellite markers. The two specific questions addressed by the study were:

1. What are the effects of the reproductive mode and the mating type (ID) on early performance indicators in progeny, including seed size, early growth, date of emergence and seedling size?

2. Is the effect of seed size, early growth and date of emergence on seedling size modified by reproductive mode and mating type (ID)?
Materials and Methods

Sampling and design

Seeds were obtained from the Mt Taylor population (TA) representing the narrow leaf morphotype (NAR) (see Table 1.1, Chapter 1). The TA population is one of the four populations where mating and reproductive patterns were estimated in addition to the fine-scale population genetic structure as described in Chapter 2 and 3. The population was selected because seed parents produced sufficient number of CL and OP seed simultaneously to facilitate a family-structured experimental design and large numbers of seed parents were already genotyped which assisted in use of microsatellites for a posteriori identification of selfed and outcrossed OP progeny.

The use of naturally pollinated OP progeny was an important feature of the experimental design. In many Legume species position of the ovule within open-pollinated fruits interacts with mating type and affects seed mass (Horovitz et al. 1976; Nakamura 1988; Bawa and Buckley 1989; Fenster 1991a; Ibarra-Perez et al. 1996). Ovule-position effects on seed mass arise due to morphological constraints on the development of distal ovules within the legumes (Watson and Casper 1984). Within a style, self-pollen may be often outcompeted by cross pollen in the race for ‘superior’, usually central, ovule positions. As a result smaller seeds in ‘inferior’ distal positions tend to be self-fertilised and those in superior positions are outcrossed. In G. clandestina, fruits clearly show ‘inferior position’ effects on seed size, and distal positions yield smaller seeds than those from central positions. Because natural pollination results in the mixture of selfing and outcrossing within an OP fruit (Schoen and Brown 1991), the use of natural progeny arrays reveal a more realistic picture of the consequences of selfing for juvenile performance if the performance is affected by seed weight. This is in contrast to hand pollination, which may alter naturally occurring fitness differences between outcrossed and selfed progeny arising from seed mass differences within a legume.

Naturally pollinated mature fruits of both reproductive modes were collected in summer 1999 (November and December) from each of seven seed parents within the population. Each fruit was identified as OP or CL based on its position on the plant; OP flowers are
produced in racemes while CL flowers are sessile in leaf axils above OP flowers within a shoot (Chapter 2).

For comparison of progeny types within each family at least five seeds of each of the three progeny types were needed. Given the germination rate and the outcrossing rate in OP flowers (c. 60%), this required 7 CL and 23 OP seeds per family (total N=210 seeds). The seeds were selected randomly from as many separate fruit as possible.

Seedling size at seven weeks was used to assess juvenile performance. Additionally, the assessment was made of seed weight, time to emergence, and early growth rate because they were expected to affect seedling size.

**Glasshouse experiment**

Following collection, fruits were stored in paper bags at room temperature for 9 to 12 months. In spring 2000 (October) seeds were weighed individually to the nearest 0.0001 g. To break physical dormancy, the seeds were heated to 80°C for 4 minutes, and then by soaked in distilled water overnight before planting. Using this treatment 97% of seeds germinated. The treatment was necessary because seed dormancy causes low (c. 10%) germination rates in untreated seeds in *G. clandestina* (Auld and O'Connell 1991).

Each seed was planted at c. 1 cm depth in a separate pot (763 cm³), in 1:4 sand:soil+peat+compost, with no added fertiliser. Pots were randomly arranged on glasshouse benches and grown at a 18°C night and 30°C day regime under natural light, and watered daily until full expansion of the third trifoliate leaf (about 7 weeks). A census of all pots was conducted daily. The following parameters were recorded for each seedling: days to cotyledon emergence from soil (days to emergence), days from emergence to first simple leaf expansion (early growth), and the width and length (to the nearest 0.5 mm) of the mid-leaflet in the first trifoliate leaf when leaves were fully expanded. Seedling size was estimated as a product of the mid-leaflet’s width and length. At the end of the census leaves were collected from OP seedlings for DNA extractions.

**Determination of mating type in OP progeny**

Microsatellite markers were used to determine whether OP progeny were selfed or outcrossed by comparing the genotypes of the parent and the progeny within each
A seedling was classified as outcrossed if it possessed an allele not present in the parent at one or more loci. The knowledge of parental genotypes and microsatellite locus polymorphism within the population was utilised to maximise the probability of outcross detection with minimum genotyping effort. Parent-locus combinations were selected to ensure that seed parents were homozygotes for relatively rare alleles. The probability of outcross detection was based on the frequency of non-parental alleles in the population, multiplied across loci, which, given two loci, was at least 0.85, with a mean of 0.93. To further economise on genotyping effort, the screening of seedlings was conducted in two steps, with the second locus only included if the seedling was not a heterozygote at the first locus.

DNA was extracted from fresh leaves, PCR conducted and its products resolved following the protocol outlined in Chapter 2 and Hempel and Peakall (2003) (see Appendix 1). Each seedling was genotyped at one to two of the four microsatellite loci (Satt164, Satt478, Satt350, Satt131) shown to be polymorphic in the TA population (see Chapter 2 and Hempel and Peakall 2003).

**Statistical analysis**

*Does reproductive mode and mating type affect early progeny performance?*

The significance of the differences in performance among the three progeny types was tested for the four measured characters (seed weight, days to emergence, early growth rate and seedling size) using a mixed model analysis of variance (ANOVA) with progeny type as a fixed factor (three levels: CL self, OP self and OP outcross), and family and family by progeny type interaction as random factors. For characters in which progeny type was significant, nonorthogonal planned contrasts were used between OP-self and OP-outcross and CL-self and OP-self to test for mating type and reproductive mode effects, respectively. The tests were conducted using synthetic denominators estimated from the ANOVA.

These and other analyses were conducted in JMP v. 3.0.1 (SAS Institute Inc.). Examination of residuals indicated that data transformation was not required for any of the models. Means (±SE) are presented in the text.
How does seed mass and seedling development relate to seedling size?

Because CL and OP seeds were known to differ in mass, it was tested whether there was a relationship between seed mass and seedling size and whether it differed among the progeny types. A mixed model analysis of covariance (ANCOVA) was used with seed mass as a covariate, progeny type (three levels: CL self, OP self and OP outcross) as a fixed factor, and family and family by progeny type interaction as random factors. The two remaining characters, days to emergence and early growth, were also included in the analysis as they describe the rate of seedling development and thus may not be independent of seedling size. It was expected that if seedling size differences were only due to differences in seed mass, other covariates, progeny type and progeny by seed mass interaction would not be significant in the model.

When any progeny type by covariate interaction was significant, further tests were made for significance of differences in slope between the two pairs of progeny types by sequentially removing one of the progeny types from the model and checking if covariate by progeny type interaction was still significant.

Depression in seedling performance due to mating type and reproductive mode

Measures of depression in performance ($D$) due to mating type ($D_M$, equivalent to $\delta$, Charlesworth and Charlesworth 1987b) and reproductive mode ($D_R$) within each family were calculated using a standard inbreeding depression formula:

$$D = 1 - \frac{w_1}{w_2} (\text{Charlesworth and Charlesworth 1987b})$$

where $w_1$ is performance of OP self, and $w_2$ is performance of OP outcross (for $D_M$) or CL self (for $D_R$). Values of $D > 0$ indicate depression of performance of one type of progeny relative to the other, either due to inbreeding depression ($D_M$) or due to the advantage of CL compared with the OP reproductive mode ($D_R$). Family $D_M$ and $D_R$ values were averaged to give a population estimate as suggested by Johnston and Schoen (1994). An alternative measure of relative performance (RP) was also calculated (Agren and Schemske 1993). This measure reduces downward bias caused by any families with negative $D$, which can have a large effect on the estimate of population $D$ (Johnston and Schoen 1994). RP values gave comparable results to $D$ and therefore they are not shown here. $D$ was taken to be significant for a character if the ANOVA showed significant differences in the pairwise progeny contrasts.
Seedling size was used as the principal performance proxy in calculations of relative fitness. Any of the other three measured characters were also included when they differed in pairwise comparisons between progeny types, but were also shown to be independent from seedling size by the ANCOVA.
Results

The 210 seeds produced a total of 195 seedlings by the time of the termination of the experiment at seven weeks, with 45 of CL and 150 of OP origin. Genotyping identified 70 of the OP seedlings as outcrosses. This resulted in the average of 6 CL self, 11 OP self and 9 OP outcross progeny within each family.

Reproductive mode affects seed mass and seedling size

The average performance of the three progeny types with respect to the four measured parameters is shown in Fig. 4.1. A clear hierarchy in seedling size was apparent with seedlings derived from CL selfing being the largest, OP outcresses being intermediate, and OP selves being the smallest (Fig. 4.1D). This hierarchy closely reflected the pattern shown by seed mass (Fig. 4.1A), indicating that these two parameters may be correlated. The significant effect of progeny type in the mixed ANOVA, for both seedling size and seed mass (Table 4.1) confirmed that there were significant differences in performance among the progeny types. The other two parameters, days to emergence and early growth rate did not follow this pattern (Fig. 4.1B,C) and the ANOVAs showed no significant differences among the progeny types (Table 4.1). Subsequent paired contrast for mating type (OP self vs. OP outcross) on seed mass and seedling size showed no significant effects (Table 4.1), indicating that there was no evidence for significant ID. On the other hand, reproductive mode effect (OP self vs. CL self) was significant (Table 4.1).

Even though this study did not aim to explicitly analyse the family (seed parent) effects, it should be noted that either family or both family and family by progeny effects explained significant amount of the variation in the data for seed mass, and seedling size and days to emergence (Table 4.1). This indicated that not only the mean performance may vary among the families but also that progeny types from different families have a potential to show contrasting patterns of seedling performance.

Reproductive mode affects the relationship between seed mass and seedling size

The analysis of covariance (ANCOVA) to test for effects of seed mass, days to emergence and early growth on seedling size is presented in Table 4.2. Only seed mass showed a significant effect on seedling size. The significant seed mass by progeny
interaction indicated that progeny types differed in the way seed mass influenced seedling size.

When progeny types were sequentially excluded from the model, OP self and OP outcross progeny did not significantly differ in their relationship between seed mass and seedling size ($F_{1,123}=1.26, p=0.26$), but there was significant difference in slopes between both of these progeny types and CL selfed group (OP self vs. CL self: $F_{1,98}=7.76, p=0.006$; OP outcross vs. CL self: $F_{1,91}=10.49, p=0.002$). When the relationship between seed mass and seedling size was examined for each progeny type separately, family included as a random effect, the models showed a significant positive relationship for CL self and OP self, not for OP outcrosses (Table 4.2). These results confirm that mating type did not affect the relationship between seed mass and seedling size. More importantly, they indicate that some of the superiority of CL self compared to OP self seedlings was due to the greater rate of increase of CL self seedling size in response to increase in the seed mass.

Only seedling size was included to estimate relative performance depression due to reproductive mode ($D_h$) and mating type ($D_m$) effects. Even though days to emergence and early growth showed independence from seedling size in the ANCOVA, and thus may independently affect juvenile fitness, they did not significantly differ with progeny type. Performance depression was found for reproductive mode ($D_R = 0.21±0.07$), but not for mating type ($D_M$ to $0.06±0.05$) (see ANCOVA above).
Discussion

As expected from the differences between cleistogamy and chasmogamy in seed size in *G. clandestina* (Chapter 2), selfing in cleistogamous flowers produced superior seedlings compared to selfing in chasmogamous flowers. This difference could be attributed to both the greater seed mass and a greater response of seedling size to seed mass in CL self progeny. In contrast, the type of mating in OP flowers had no significant effect on seedling performance, which is consistent with lack of inbreeding depression in early seedling performance in *G. clandestina*.

**Seedling performance enhanced by greater cleistogamous seed mass**

In Chapter 2 it was shown that *G. clandestina*’s reproductive strategy was characterised by consistently higher mass of cleistogamous as opposed to chasmogamous seeds. The comparison of selfed seeds between the two reproductive modes in this study confirmed a non-genetic character of this effect. Moreover it showed that the seed mass differences were paralleled by a 21% depression in chasmogamous compared to cleistogamous seedling size. Interestingly, the increase in seedling size in response to seed mass was also greater for cleistogamous seedlings. Such difference was unlikely due to the difference in rate of seedling development since neither the days to emergence nor early seedling growth differed significantly. Therefore the difference in performance of seedlings from CL and OP selfing in *G. clandestina* appears to be a direct reflection of the seed size differences as well as the differences in the response of seedling size to seed mass.

In plants differences in seed mass often result in differences in seedling size and this effect is due to the greater content of carbohydrates and nutrients, or larger sized embryos in seeds with greater biomass (Lee and Fenner 1989; Milberg and Lamont 1997). There is also evidence that nutrient concentration can increase disproportionately with seed mass, in turn causing a disproportionate increase in seedling size (Vaughton and Ramsey 2001). Such difference in allocation of resources may account for direct and indirect effects of seed biomass on seedling size observed in *G. clandestina*. As discussed in Chapter 2, the differences between the two reproductive modes in local competition for resources during seed maturation as well as differences in timing of seed production could cause those allocation differences in *G. clandestina*.
Even though seedling size is not a direct measure of early/juvenile fitness, the non-genetic differences in seedling size in *G. clandestina* are likely to affect early/juvenile fitness through effects on establishment success. Studies of the effects of differential maternal allocation show that as a rule the larger the seed and seedling, the greater its chance of survival (Sacchi and Price 1992; Bonfil 1998; Eriksson 1999; Shaukat *et al.* 1999; Seiwa 2000; Navarro and Guitian 2003). *G. clandestina* seedlings are likely to be most limited by nutrients and water. Nutrient poor soils are prevalent in Australia and a characteristic of the study sites. *G. clandestina* seedlings are also likely to be under much water stress because mass germination is triggered by good rainfalls in early summer (pers. obs.), followed by very high temperatures and lack of further rainfall. Studies of other Australian plants have found that it is the increased nutrient supply in larger seeds that ensures seedling survival in nutrient poor environment (e.g. Stock *et al.* 1990; Milberg and Lamont 1997; Vaughton and Ramsey 2001; Lamont and Groom 2002). A larger root system in larger seedlings is also expected to enhance survival under water stress (Jobidon *et al.* 1998; Donovan *et al.* 1993; O'Connor 1996; but see Chacon and Bustamante 2001). Therefore, overall, cleistogamous selfing in *G. clandestina* should result in seedlings with enhanced survival over those derived from selfing in OP flowers.

**Outcrossing does not improve seedling performance**

In contrast to the reproductive mode, the type of mating did not appear to affect early seedling performance in *G. clandestina*. In comparisons between naturally selfed and outcrossed chasmogamous seeds no significant differences were found in any of the measured parameters: days to emergence, early growth rate, seed mass and seedling size, or the relationship between seed mass and seedling size. Nonetheless, a trend was apparent for 5% depression in performance of selfed progeny in the same parameters as those that were affected by the reproductive mode - seed weight and seedling size. Thus while the results show that overall inbreeding depression in seedling performance was not significant, they also indicate that outcrossing may have small positive effect on seed and seedling size.

It is possible that under field conditions the reduction in fitness of juveniles due to ID in *G. clandestina* is greater than indicated by this glasshouse experiment. This is because inbreeding depression in direct measures of fitness, like survivorship, tends to exceed levels estimated using morphological characters, like seedling size (DeRose and Roff
Additionally, the expression of inbreeding depression is usually enhanced under harsh experimental regimes (e.g. Schmitt and Ehrhardt 1990; Dundash et al. 1997; Ramsey and Vaughton 1998). The benign glasshouse conditions in this study were reflected in the high germination rates of 97% and the lack of substantial mortality over the 7 weeks of duration. By comparison the usual field environment experienced by *G. clandestina* is harsh and seedling mortality is an important selective force, which is evident in 85% death rate within the first year from germination (pers. obsv.).

On the other hand the lack of substantial inbreeding depression in juvenile performance is consistent with purging of recessive deleterious mutations due to habitual selfing, as would be expected under long-term selfing rates of 60% observed in the study population (Chapter 2). An increase in inbreeding has been confirmed to produce reductions in inbreeding depression (Byers and Waller 1999; Crnokrak and Barrett 2002) the effect reflected in 43% reduction of lifetime ID in selfing compared to the outcrossing plants (Husband and Schemske 1996). Such reduction in the lifetime ID in selfers occurs primarily due to reduced ID at juvenile life stage and in mortality, both thought to be due highly recessive lethal or very deleterious mutations that are easily purged even with little inbreeding (Lande and Schemske 1985; Charlesworth and Charlesworth 1990; Willis 1999). Indeed the size superiority of outcrossed seedlings of 5% observed in this study is close to values for early performance indicators reported in prevalent selfers (Husband and Schemske 1996).

Further evidence that highly deleterious or lethal loads have been purged in the study population comes from the examination of seed-set within fruit of CL and OP type. Because they were produced under harsh field conditions, the presence of early inbreeding depression would be expected to manifest itself in increased mortality of embryos in fully CL fruit compared to that in OP pollinated fruit where only 37% of seed resulted from selfing in the study population (Chapter 3). At the same time in Chapter 2 the seed set (no. mature seed/total no. embryos) within CL fruit was shown to be significantly greater than in OP fruit in this and other three populations. While some of this difference could be attributed to lower efficiency of pollination in OP flowers, analysis of additional data for the study population shows that initiated chasmogamous seed also tended to be more likely to abort (OP vs. CL: Proportion of fruit with abortion 0.64±0.14 vs. 0.40±0.06; Proportion of aborted embryos 0.15±0.04 vs. 0.10±0.02). This
is exactly the opposite of the effect on seed mortality in the two types that early ID should cause. Together with the low level of inbreeding depression in juvenile performance observed here, this suggest that genetic load responsible for early juvenile depression may be low in G. clandestina, although its effect on seedling performance should be enhanced in the field.

**Evolutionary implications for CLOP in G. clandestina**

The comparison of reproductive mode and mating type effects on juvenile size in G. clandestina demonstrates that in terms of size cleistogamous seedlings outperform chasmogamous ones, whether the latter originated from outcrossing or selfing. The superiority of CL selfed seedlings could be as high as 20% compared to OP selfs, or at least 15% compared to OP outcrosses, assuming 5% inbreeding depression. These patterns are unlikely to be confined to the study population, since the investigations in Chapter 2 and 3 have found that superiority of CL seed mass over chasmogamy was typical and that factors affecting expression of inbreeding depression were relatively constant across all of the studied populations. This suggests that advantage to chasmogamy reproduction in G. clandestina generally does not arise from avoidance of ID due to outcrossing in terms of early progeny performance.

The outcome of the interaction between the effects of the reproductive mode and the mating type on juvenile fitness is likely more complex that the seedling size hierarchy observed in the glasshouse suggests. For the seedling size to affect juvenile fitness the differences in size should be reflected in seedling establishment rates and enhanced survival of CL progeny, which is likely under harsh conditions in the field (Venable and Brown 1988). However, the similarity expected and observed inbreeding levels in all six G. clandestina populations (Chapter 2) is contrary to the expectation of excess homozygosity should the survival of inbred CL seedlings be enhanced. This suggests either that seedling size does not influence juvenile survival of G. clandestina in the field, or that ID depression is enhanced leading to no net difference in mortality. Further studies are required to tease these effects apart.

The effects of ID and reproductive mode on early progeny performance are unlikely to be reflected in fitness components at later life stages. Firstly, the effects of enhanced CL seedling performance are likely to wane at the later life stages as commonly occurs with non-genetic seedling size differences of similar magnitude in plants (reviewed in
Stanton 1984; Baskin and Baskin 1998; also see Weis et al. 1987; Ouborg and Vantreuren 1995; Cervantes et al. 1998; Castro 1999; for an exception see Tremayne and Richards 2000). Persistence or even enhancement of seedling size effects are only expected under size-asymmetric competition for light in dense stands of uniform age (Berntson and Wayne 2000), which is typical of many annuals (Schwinning and Fox 1995; Schwinning and Weiner 1998). In G. clandestina such competition is unlikely because of the overlapping generations, the relatively low seedling density (pers. obs.), and because of symmetrical nature of competition for water and nutrients which are the resources that likely to limit G. clandestina seedlings (Berntson and Wayne 2000).

Secondly, the effects of ID should become more apparent at later life stages, as in other habitual selfers where ID in growth and fecundity remains largely unaltered and of magnitude comparable to that in outcrossers (Husband and Schemske 1996). This is because late ID purging is retarded due to its basis in mildly deleterious and polygenic mutations (Lande and Schemske 1985; Charlesworth and Charlesworth 1987b; Charlesworth et al. 1990) or overdominance (Charlesworth and Charlesworth 1987b; Byers and Waller 1999). An average of 23% depression in fitness of selfed as compared with outcrossed progeny is still retained in selfers (Husband and Schemske 1996), and it can be as high as 50% (Holsford and Ellstrand 1990). Given the expected decrease in the effects of seed mass and seedling size favouring CL progeny, and the increasing expression of ID in favour of OP outcrosses, the relative performance of the latter should be enhanced at later life stages.

Finally, in all of the studies that have reliably assessed ID in CLOP, a significant and sometimes substantial ID was detected at later life stages, even where it was insubstantial or absent early in plant life (Wilken 1982; Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990; Bryan and Muenchow 1991; Culley 2000; Cope 1966). In general this relative increase in importance of ID over plant life corresponded to waning of the non-genetic effects of reproductive mode (Cope 1966; Schmitt and Ehrhardt 1990; Bryan 1993 cited in Culley 2000) although they could remain strong under certain conditions (Cope 1966; Schmitt and Ehrhardt 1990) or increase in some species (Culley 2000).
Similar patterns may be expected in *G. clandestina*. Overall, as the late acting ID can be quite substantial, it is likely that despite the early superiority of selfed CL progeny outcrossing provides some advantage to OP reproduction in *G. clandestina*.

**Comparison with other CLOP species**

**Effects of reproductive mode on progeny fitness**

*Glycine clandestina* shares the beneficial effect of seed mass on CL juvenile performance with other CLOP species where the CL seeds are heavier than OP seeds, but where a striking seed mass dimorphism is accompanied by amphicarpy or other differences in dispersal. In those species the seed mass advantage to cleistogamy of 1.5-40-fold resulted in a pronounced superiority of CL progeny not only at the juvenile stage (seedling size and survival) but also at the adult stage (size and fecundity) (Clay 1983; Schoen 1984 and references therein; Trapp and Hendrix 1988; Campbell *et al.* 1983; Cheplick and Clay 1989; Cheplick 1996). These studies mostly underestimate the effect of reproductive mode because they compare CL and OP seed without accounting for outcrossing in OP flowers (but see Trapp and Hendrix 1988). By comparison, the 1.07 fold difference in seed mass between cleistogamy and chasmogamy in *G. clandestina* (Chapter 2) is small and its effects, although shown here to be important for seedling performance, are unlikely to match the magnitude or the duration of the reproductive mode effects in those species.

An example of a reproductive mode effect on a similar scale to that *G. clandestina* has been described in *Lespedeza cuneata* (Cope 1966). In this case, however, the non-genetic effects on seed mass and seedling size were in favour of chasmogamy rather than cleistogamy. The study by Cope (1966) showed that the reproductive mode effect in *L. cuneata* was responsible for 13% superiority in OP seedling size at germination and an average OP seedling size superiority of 7.8% over first 4 months of growth. At the same time only less than 5% of overall OP seedling size superiority could be accounted for by mating type effect. Similarly to *G. clandestina*, early performance was significantly correlated with seed mass. Cope’s findings (1966) contradicted earlier assumptions that the 20% seed mass superiority and 20 to 40% greater overall performance in biomass and fecundity typical of *L. cuneata* was entirely due to ID (Donnelly 1955).
Together, the findings in *G. clandestina* and *L. cuneata* demonstrate that modest differences in seed mass between cleistogamy and chasmogamy can arise because of non-genetic effects and that they have consequences for seedling performance. This has important implications for studies of ID in other species with superiority of OP seed mass of comparable magnitude (3-25%) where it may be interpreted incorrectly as evidence for ID in comparisons of CL and OP progeny (*Impatiens capensis*, Waller 1982; Schmitt and Gamble 1990; Mitchell-Olds and Waller 1985; *Viola hirta*, Berg and Redbo-Torstensson 1999; *Oxalis acetosella*, Berg 2000b). Among these species, *Impatiens capensis* has become an important model system for study of the role of ID in CLOP in particular (Waller 1984; Mitchell-Olds and Waller 1985; Antlfinger 1986; Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990) and dynamics of ID in general (McCall et al. 1994; Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990). There is strong evidence in *Impatiens* that differences in seed mass can influence seedling fitness (Waller 1985; McCall et al. 1994; Schmitt and Ehrhardt 1990) and, under asymmetric competition, lifetime fitness (Schmitt and Ehrhardt 1990; Waller 1985). Yet where observed, the significant differences in seed mass between cleistogamy and chasmogamy have variously been assumed to be non-genetic (e.g. Schmitt and Ehrhardt 1990) or due to ID (e.g. McCall et al. 1994). As both assumptions potentially lead to important bias in ID estimates (especially in juvenile fitness), the real basis of this seed size dimorphism should be established before estimates of the magnitude of ID are made based on comparison between cleistogamy and chasmogamy.

A similar experimental design to this study, involving the comparison of CL self, OP self and OP outcross progeny to separate reproductive mode and mating type effects, was also employed in the study of *Viola canadensis* (Culley 2000) and *Triodanis perfoliata* (Bryan and Muenchow 1991). In contrast to *G. clandestina*, reproductive mode did not influence seed mass in either of these species, but germination behaviour was affected in favour of cleistogamy (Bryan and Muenchow 1991; Culley 2000). In *V. canadensis* cleistogamy tended to enhance germination success (Culley 2000) and in *T. perfoliata* cleistogamy significantly enhanced the speed of germination (Bryan and Muenchow 1991). Both of these effects can be interpreted to be due to dormancy differences between the two reproductive modes, resulting from the differences in the environment during seed development (preconditioning) between them (Baskin and Baskin 1998). As seed dormancy in *G. clandestina* is very high (Auld and O'Connell
1991; Morrison et al. 1992; pers.obs.), it had to be broken so as not to impede the experiment in this study. This may explain the lack of reproductive mode-related differences in germination behaviour in G. clandestina here. However, when preliminary comparisons were made they showed that CL germination rate tended to be greater than OP (pers.obs.), which given high seed viability of both types, suggests dormancy differences between them. Such difference in dormancy may combine with the seed mass effects to further enhance early progeny performance in G. clandestina.

**Effects of inbreeding depression on progeny fitness**

As in G. clandestina, the comparison of OP self and OP outcross progeny provided little evidence for mating type effect or ID in early progeny performance either in V. canadensis or in T. perfoliata (Bryan and Muenchow 1991; Culley 2000). Only ID in speed of germination in the latter species was significant (Bryan and Muenchow 1991). One other study in Collomia grandiflora (Wilken 1982) compared outcrossed and selfed progeny in OP flowers but did not test for early ID. More substantial early ID was detected in two studies of I. capensis based on comparison of cleistogamy and chasmogamy adjusted for outcrossing in OP flowers (ID calibrated by OP selfing rate in Schmitt and Gamble 1990; full OP outcrossing ensured in Schmitt and Ehrhardt 1990). In both studies ID was unrelated to seed mass, and it was evident in depression of the order of 9-31% in each: germination success (Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990), seedling size (Schmitt and Ehrhardt 1990), and seedling survival (Schmitt and Gamble 1990). The expression of the early as well as the late ID in those studies was substantially enhanced by stressful experimental conditions (Schmitt and Ehrhardt 1990; Schmitt and Gamble 1990). Particularly interesting was the pattern of increase in the magnitude of ID with increased distance from parental site, resulting from the decrease in fitness of CL progeny (Schmitt and Gamble 1990). It was suggested that this was a consequence of the environmental heterogeneity within populations contributing to local genetic differentiation and the better adaptation of inbred progeny to the parental site (Schmitt and Gamble 1990). In G. clandestina populations are significantly genetically differentiated at small scales (Chapter 3) and thus it may be particularly pertinent to continue investigations of expression of ID in this species in relation to distance from parental site.
Conclusions

This appears to be the first study to examine the effects of reproductive mode and mating type on seedling performance in a CLOP species with seed mass dimorphism in favour of cleistogamy. Reproductive mode effects on seed mass have been shown to be more important in the determination of juvenile performance than mating type, because of preferential provisioning of CL seeds. As the result of preferential provisioning, the seedlings derived from selfing in CL flowers outperformed those from selfing in OP flowers by 20% in terms of size. Moreover, outcrossing in OP flowers tended to produce at most a 5% advantage in seedling size over OP selfing, which is consistent with purging of ID. Thus on average in G. clandestina CL seedlings outperformed OP seedlings, even if the latter were outcrossed. This suggested that advantage to OP reproduction in G. clandestina generally does not arise from avoidance of ID due to outcrossing in terms of early progeny performance, assuming that the seedling size differences are reflected in seedling establishment rates in the field.

More generally, the results in G. clandestina emphasise that relatively small difference in seed mass result in significant differences in seedling performance, and such effects should be taken into account in other CLOP studies, especially where the reproductive mode differences could be misinterpreted as inbreeding depression.

The differences in seedling size that followed from the seed mass differences are likely to be reflected in the fitness of juveniles in G. clandestina. However, studies of other habitual selfers and CLOP species suggest that in later life stages the non-genetic effects of seed mass in favour of cleistogamy should wane and the effects of ID in favour of outcrossing in OP flowers should increase. As a consequence, I predict that the advantage to chasmogamy from avoidance of inbreeding will become apparent over the progeny’s lifetime in G. clandestina. Further tests of this prediction in G. clandestina should take into account the potential for differences in dormancy between reproductive modes and environmentally induced variation of expression of ID in general, and its increase with distance from the parent in particular.

Given that ID appears to have been at least partially purged from G. clandestina populations and that reproductive mode effects work to enhance relative quality of CL progeny, the overall advantage to chasmogamy derived from avoidance of inbreeding
depression would be considerably reduced. Therefore, other fitness gains from chasmogamy are expected to play an important role in balancing out the advantages of CL reproduction in *G. clandestina*. Fitness gains from chasmogamous male function are one possibility that will be discussed further in the next chapter.
Table 4.1  Results from mixed model ANOVA examining the effects of progeny type (CL self, OP self and OP outcross) and maternal family on seed mass, days to emergence, early growth and seedling size in G. clandestina. Data are shown in Fig. 1. Progeny type and family were considered fixed and random factors, respectively.

<table>
<thead>
<tr>
<th>Character</th>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed mass</td>
<td>Progeny type</td>
<td>2</td>
<td>1659.4</td>
<td>829.68</td>
<td>12.89</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>CL self vs OP self</td>
<td>1</td>
<td>1633.51</td>
<td></td>
<td>25.36</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>OP self vs OP outcross</td>
<td>1</td>
<td>153.305</td>
<td></td>
<td>2.38</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>6</td>
<td>1022.0</td>
<td>170.33</td>
<td>2.67</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Family x Progeny type</td>
<td>12</td>
<td>803.1</td>
<td>66.92</td>
<td>3.11</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>174</td>
<td>3740.4</td>
<td>21.497</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days to emergence</td>
<td>Progeny type</td>
<td>2</td>
<td>72.6</td>
<td>36.298</td>
<td>3.98</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>6</td>
<td>462.9</td>
<td>77.15</td>
<td>8.49</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>Family x Progeny type</td>
<td>12</td>
<td>111.6</td>
<td>9.296</td>
<td>1.54</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>174</td>
<td>1049.6</td>
<td>6.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early growth</td>
<td>Progeny type</td>
<td>2</td>
<td>0.586</td>
<td>0.293</td>
<td>0.209</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>6</td>
<td>2.942</td>
<td>0.490</td>
<td>0.352</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Family x Progeny type</td>
<td>12</td>
<td>17.212</td>
<td>1.434</td>
<td>1.708</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>171</td>
<td>143.564</td>
<td>0.839</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seedling size</td>
<td>Progeny type</td>
<td>2</td>
<td>10866.6</td>
<td>5433.31</td>
<td>5.71</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>CL self vs OP self</td>
<td>1</td>
<td>10527.7</td>
<td></td>
<td>11.06</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>OP self vs OP outcross</td>
<td>1</td>
<td>707.287</td>
<td></td>
<td>0.74</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>6</td>
<td>19287.3</td>
<td>3214.54</td>
<td>3.40</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Family x Progeny type</td>
<td>12</td>
<td>11865.3</td>
<td>988.773</td>
<td>3.17</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>173</td>
<td>53902.5</td>
<td>311.58</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2  Results from mixed model ANOVA examining the effects of days to emergence, early growth and seed mass in conjunction with progeny type (CL self, OP self and OP outcross) and seed parent (family) on seedling size in G. clandestina. Progeny type and family were considered fixed and random factors, respectively and days to emergence, early growth and seed weight are included as covariates.

<table>
<thead>
<tr>
<th>Character</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days to emergence</td>
<td>1</td>
<td>398.887</td>
<td>398.88</td>
<td>1.48</td>
<td>0.225</td>
</tr>
<tr>
<td>Early growth</td>
<td>1</td>
<td>566.799</td>
<td>283.399</td>
<td>1.05</td>
<td>0.350</td>
</tr>
<tr>
<td>Seed mass</td>
<td>1</td>
<td>5852.32</td>
<td>5852.32</td>
<td>21.78</td>
<td>0.000</td>
</tr>
<tr>
<td>Progeny type</td>
<td>2</td>
<td>2907.31</td>
<td>1453.66</td>
<td>5.36</td>
<td>0.006</td>
</tr>
<tr>
<td>Seed weight x Progeny type</td>
<td>2</td>
<td>2648.12</td>
<td>1324.06</td>
<td>4.92</td>
<td>0.008</td>
</tr>
<tr>
<td>Family</td>
<td>6</td>
<td>17418.9</td>
<td>2903.15</td>
<td>5.44</td>
<td>0.005</td>
</tr>
<tr>
<td>Family x Progeny type</td>
<td>12</td>
<td>6643</td>
<td>553.584</td>
<td>2.06</td>
<td>0.022</td>
</tr>
<tr>
<td>Error</td>
<td>167</td>
<td>44866.14</td>
<td>268.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 4.1 Average early performance of each of the three types of progeny produced by seven *G. clandestina* seed parents: seed size (A), days to emergence (B), early growth rate (C) and seedling size (D). Error bars show SE of the means. Analysis of differences using mixed model ANOVAs is shown in Table 1.
Chapter 5

Importance of male function in favouring chasmogamy in *Glycine clandestina* – discussion and new hypotheses

‘It is a favourite popular delusion that the scientific inquirer is under a sort of moral obligation to abstain from going beyond [the] generalisation of observed facts... But anyone which is practically acquainted with scientific work is aware that those who refuse to go beyond fact rarely get as far as fact [...]’

T.H. Huxley (1898)
The male function hypothesis for OP flowers in CLOP yields four important predictions: 1. Chasmogamous fruit set is limited by factors other than pollinator service; 2. Non-fruiting or excess OP flowers contribute to the male reproductive success; 3. The increase in OP flower proportion with plants size is accompanied by increased attractiveness to pollinators; 4. This increased attractiveness to pollinators causes greater male reproductive success but has little or no effect on the proportion of OP fruit, or fruit set in OP flowers.

**Evidence for the male function hypothesis for chasmogamy in G. clandestina**

*Weak advantage to chasmogamy based on female function in G. clandestina*

One reason for considering that the gains via male function offer an important advantage to chasmogamy in *G. clandestina* is that gains in the quality of OP progeny from avoidance of ID are unlikely to overcome the fecundity gained from seed production in CL flowers. This is apparent even without assessment of total lifetime ID in the species. Female reproductive success in CL flowers was shown to exceed that in OP flowers by 4 to 7-fold (Chapter 2). Cleistogamous seedlings also benefited from the preferential provisioning of CL seed that gave them a 20% advantage in size (Chapter 4). When combined, this amounted to a 5 to 8-fold greater female reproductive success for cleistogamy over chasmogamy. Given that any benefits from avoidance of ID in OP flowers are discounted by an average of 55% selfing in OP flowers (Chapter 3), the lifetime ID must exceed 80% in order to overcome the advantage to cleistogamy in *G. clandestina* if based solely on OP female function (OP/CL: [1-s*ID]/[1-ID]; see equation 1 in Schoen and Lloyd, 1984). However, this study found little evidence for early ID in *G. clandestina* in common with studies of other selfers (Chapter 4). Moreover, while this study underestimated lifetime ID in *G. clandestina*, other studies of both selfers and CLOP suggest lifetime ID rarely reaches 50% (Husband and Schemske 1996). Thus, if the avoidance of ID is the only short-term benefit from chasmogamy, a strategy of seed production via complete cleistogamy should prevail.

*Patterns of chasmogamy not explained by female reproductive success*

Additional evidence that suggests chasmogamy in *G. clandestina* provides gains other than through seed quality emerges from the patterns of variation in flower and fruit production with plant size. The relative investment in OP flowers increased while the proportion of OP fruit remained constant with increase in plant size (Chapter 2). Thus
the adaptive value of OP reproduction in large plants appears not to depend on the production or properties of seeds (or female function in chasmogamy), be it ID-based or otherwise. On the other hand, this pattern is consistent with the male function hypothesis for chasmogamy since it reflects an increase in the proportion of the non-fruiting OP flowers with plant size.

In *G. clandestina* there is also evidence for other predictions of the male function hypothesis. Despite significant pollen limitation at fruit initiation, fruit set was further limited at fruit maturation, with one third of OP flowers aborting after fruit initiation. Since little autonomous selfing was found in *G. clandestina*, the aborted flowers most likely acted as pollen donors, as predicted by the male function hypothesis for excess flowers. Also consistent with this hypothesis was the finding that the increased OP display size in larger plants did not affect the female reproductive success as measured by the probability of OP fruit set. Constraints on female fecundity were shared between cleistogamy and chasmogamy, as indicated by a similar rate of increase in fruit production with plant size in CL and OP flowers.

While these patterns of size-dependent chasmogamy in *G. clandestina* are broadly consistent with the male function hypothesis for chasmogamy, further study is needed. In particular it will be important to confirm that non-fruiting OP flowers contribute to male reproductive success and that display size enhances it. The lack of correlation between floral display size and fruit set may not be fully indicative of fitness gains via seed production because of the effects of display size on seed quality (Wilson 1983; Wilson et al. 1994). Therefore it will also be necessary to demonstrate that increase in plant size causes gains from male function to increase at a greater rate than those from female function.

**Male function hypothesis for chasmogamy in other CLOP and in other plants**

Similarly to *G. clandestina*, other studies of CLOP have found that chasmogamy provides only a weak counterbalance to the benefits of cleistogamy because ID levels are low (Wilken 1982; Schmitt and Ehrhardt 1987; Schmitt and Gamble 1990; Culley 2000; Lu 2002) and selfing in OP flowers is often substantial (Wilken 1982; Mitchell-Olds and Waller 1985; Waller and Knight 1989a; Stewart 1994; Lu 2000; Culley 2002). Like *G. clandestina*, many of these CLOP species showed an increase in the proportion
of OP flowers with plant size (Waller 1980; Wilken 1982; Schnee and Waller 1986; Bell and Quinn 1987; Schmitt et al. 1987; Diaz and Macnair 1998). The change in the relative advantage to chasmogamy and cleistogamy with plant size has been discussed as an important element in maintaining the balance between the two reproductive modes. However, in contrast to the findings in *G. clandestina*, the increasing investment in OP flowers in other CLOP does not appear to be decoupled from the proportion of chasmogamy at fruiting. Therefore, the discussion has been couched in terms of the properties of OP seeds beyond the ID-based advantage, such as their wider dispersal (Schoen and Lloyd 1984) or better performance in sib-competition (Waller 1980; Schmitt and Ehrhardt 1987). Only those female function-based alternatives to ID-based OP advantage have been subject to empirical tests (e.g. Waller 1984; Mitchell-Olds and Waller 1985; Schmitt et al. 1985; Schmitt and Ehrhardt 1987; Trapp 1988; Gara and Muenchov 1990; Schmitt and Ehrhardt 1990; Le Corff 1996; Berg 2000a).

Similarly to the OP flowers in CLOP species, many floral features and reproductive patterns in plants have been primarily seen as mechanisms to prevent selfing and inbreeding depression, acting to enhance the seed quality and the female function (Willson 1983; Richards 1986; Barrett 2003). Reformulation of the sexual selection theory for plants in the late 1970s highlighted the potential importance of selection on these features through the male function (Willson 1979; Charnov 1982). However, theoretical predictions, including the male function hypothesis for the excess flowers (Broyles and Wyatt 1995; Queller 1997) and the size dependent sex allocation (Lloyd and Bawa 1984; Klinkhamer et al. 1997), remain little tested and controversial. It is partially because, as in CLOP, hypotheses based on male function often predict similar outcomes to the more traditional explanations based on the quality of progeny produced as seeds (Barrett 2003). More importantly, empirical tests were prevented by the practical difficulties in direct measurement of fitness gains from pollen export.

Recent developments in methodology, including the availability of powerful genetic markers for tracking paternity and new methods for analysis of selection via male and female function (e.g. path analysis Burd and Callahan 2000; selection gradient analysis Morgan and Conner 2001) now make empirical investigations of male function feasible (Klinkhamer et al. 1997; Burd and Callahan 2000; Barrett 2003). Latest reviews of empirical studies show growing evidence for the importance of selection via male
function in plants and it is becoming accepted as complementary to the more traditional hypotheses based around the selfing-outcrossing paradigm (Klinkhamer et al. 1997; Burd and Callahan 2000; Barrett 2003). Importantly, the male function hypothesis for excess flowers has been profitably explored in species where, like _G. clandestina_, plant size or display size did not correlate with female gains (e.g. _Asclepias syriaca_, Willson and Rathcke 1974; Broyles and Wyatt 1995).

**Conclusions**

Selection based on male function in hermaphroditic flowers provides an explanation, complementary to the avoidance of inbreeding depression, for many features of floral display, such as non-fruiting or excess flowers. The persistence of chasmogamy, and the common pattern of the increase in chasmogamy with plant size may be best explained based on the male function of OP flowers. Under this hypothesis chasmogamy primarily provides excess or non-fruiting flowers to enhance male function beyond the limits that resource availability may impose on the invariably more costly seed production in cleistogamy and chasmogamy.

The male function hypothesis appears an attractive complement to the hypothesis of an ID-based advantage to chasmogamy, as in many CLOP species the ID-based advantage is not strongly evident. However, so far CLOP studies have overlooked the male function hypothesis in favour of other female function-based hypotheses. _G. clandestina_ may be an ideal model species in which to redress this imbalance, especially given that the size-dependence of OP flowering is uncoupled from OP fruit production, and ready availability of powerful microsatellite markers for determining paternity. Indeed, further studies of _G. clandestina_ are certain to illuminate our understanding of the importance of selection via male function in plants generally.
Chapter 6

Summary and further directions
Dimorphic strategies for sexual reproduction are common in plants and their evolutionary dynamics is one of the central themes in plant evolutionary biology. True cleistogamy or CLOP is one such strategy where individuals use two types of flowers for reproduction: normal chasmogamous (or OP) flowers, which are open for insect pollination and can outcross; and reduced, closed cleistogamous (or CL) flowers specifically modified for self-pollination. This thesis has examined the relative advantages of cleistogamous and chasmogamous reproduction in an Australian perennial species, *Glycine clandestina*. The factors examined included inbreeding depression, which has the potential to favour chasmogamy, and resource and pollinator limitation as factors that may favour cleistogamy. The effect of these factors on female reproductive success or seed production was identified as central to the balance between cleistogamy and chasmogamy in an influential model by Schoen and Lloyd (1984) and have been emphasised in empirical studies to date.

This study confirmed the two classic advantages of cleistogamy in *G. clandestina* – the greater reliability and more economical seed production compared to chasmogamy. The greater reliability of seed production in CL flowers was due to pollinator limitation on fruit initiation in OP flowers, and their inability to autonomously self. This finding contradicted previous anecdotal evidence in this species. The production of CL seeds in *G. clandestina* required fewer resources because of the lower biomass of CL flowers. However, the overall savings on CL seed production were reduced because CL seeds also had a greater biomass than OP seeds. This type of seed size dimorphism is unusual for a CLOP species with aerial cleistogamy but is consistent with the prediction of Schoen and Lloyd’s model (1984) that investment in CL female function should be enhanced. As in other CLOP species, the conditional employment of cleistogamy was an important aspect of CLOP strategy in *G. clandestina*. In *G. clandestina* cleistogamy was the favoured mode of reproduction in smaller plants and in plants with lower chasmogamous fruit set, which is consistent with the greater economy and reliability of cleistogamous seed production, respectively.

Apart from the effect on relative economy of seed production, the greater maternal investment in CL seeds also enhanced CL progeny quality in *G. clandestina*. Seedlings derived from CL selfing were larger than those from OP flowers, irrespective of whether the latter were derived from outcrossing or selfing. The high selfing rates and
lack of evidence for early inbreeding depression suggest that the overall level of ID may be reduced in *G. clandestina* due to purging of dominance-based ID. Irrespective of the level of ID, any advantage to OP from outcrossing is likely reduced because of the high level of inbreeding in OP flowers from geitonogamous selfing and biparental inbreeding. Overall, the collective evidence suggested that the demonstrated benefits of CL on female function (fecundity and seed quality) likely exceeded any gains from the avoidance of inbreeding depression in *G. clandestina*. Similar findings have been common in other CLOP species.

Fitness gains afforded by the pollen export or male function in OP flowers may better explain both the maintenance of OP and the plant size-dependent increase in proportion of OP flowers in *G. clandestina*. Under the male function hypothesis, chasmogamy primarily provides excess or non-fruiting flowers to enhance male function beyond the limits that resource availability may pose on the costly seed production. In agreement with this hypothesis in *G. clandestina* there was no increase in the investment in OP female function (proportion of OP fruit) and female reproductive success (fruit set) with plant size, and fruit production appeared strongly limited post-pollination.

Most other studies of CLOP to date have focused on annual species, although recent investigations of perennials have suggested that life history does shape the way CL is conditionally employed in response to pollinator and resource limitation. However, the case of *G. clandestina* does not support the hypothesis that perennials employ a unique strategy of size-dependent investment in cleistogamy (where the proportional investment in cleistogamy increases with plant size). Rather, it is now apparent that perennials exhibit a mixture of size-dependent responses, with the investment in cleistogamy either increasing or decreasing with size, depending on the species. Nonetheless, this study of *G. clandestina* supports the hypothesis that the regulation of cleistogamy based on OP fruit set evolved in response to the perennial life history.

Cross-species amplification of microsatellite loci from the large set available in soybean (*G. max*), the crop relative of the wild *G. clandestina*, proved an efficient alternative to de novo development. It provided a powerful set of highly polymorphic markers, rarely available in selfers, which facilitated many aspects of this study including mating system estimates, assessment of long-term inbreeding across multiple populations, and a
*posteriori* identification of OP selfed and outcrossed progeny. Most importantly, microsatellites in combination with novel methods of analysis of small-scale population genetic structure allowed for the first comprehensive study of population genetic structure at multiple scales in a cleistogamous species. As expected, population genetic structure in *G. clandestina* was typical of a plant where gene flow is restricted due to intermediate selfing: there was strong differentiation among populations and significant small-scale genetic structure within populations. Strong population genetic structure may form an important background for selection. In *G. clandestina* the strong small-scale population structure likely caused the significant biparental inbreeding that was detected among outcrossed OP progeny. Such inbreeding in OP progeny reduces any advantage derived from the avoidance of inbreeding depression.

As a key element of the experimental design, this study included an investigation of two distinctive morphotypes - high elevation broad leaf morphotype (BRO) and the low elevation narrow leaf morphotype (NAR). Neither reproductive strategy nor population genetic structure differed between these morphotypes. The expectation, based on an earlier study by Schoen and Brown (1991), that the BRO morphotype may experience higher levels of autonomous selfing in OP, high levels of autonomous selfing in OP, higher inbreeding and stronger population genetic structure than the NAR morphotype was not supported. The two morphotypes also shared the differences in mass between cleistogamous and chasmogamous seed, and the conditional employment of cleistogamy in response to plant size and OP reproductive success. Despite the absence of inter-morphotypic differences, there were some important differences among populations. For example, the shorter growth season at higher elevations appeared to curtail the production of CL flowers thus reducing the value of CL for assurance of seed production under OP failure. While most of the populations shared a similar level of self-fertilisation and inbreeding, some experienced substantially higher levels. There also appeared to be some variation in scale and strength of small-scale population genetic structure, some of it unrelated to the differences in overall inbreeding but rather correlated with patterns of spatial aggregation of plants. Further investigations into evolutionary dynamics of reproductive strategy of CLOP in *G. clandestina* may be able to exploit this natural variation among populations.
Several aspects of the dynamics of the reproductive strategy of CLOP in *G. clandestina* need further investigation. Based on the findings of this study, I suggest that the following hypotheses will offer fruitful lines of inquiry:

1. Juvenile fitness of CL progeny is enhanced by the greater CL seed mass but this effect disappears at later life stages. In turn, the beneficial effects of OP outcrossing become more evident with age and substantially enhance the quality of outcrossed OP progeny over the long term. Under this hypothesis CL-self will outperform OP-self progeny as juveniles (e.g. establishment success) but their performance will be similar once they mature (e.g. fecundity) in the field. At the same time, OP-outcross will perform similarly to OP-self progeny at early life stages but will outperform OP-self progeny at later life stages.

2. Male function via pollen export in OP flowers is an important advantage to chasmogamy and is responsible for the production of excess OP flowers and increased relative allocation to OP flowers with plant size. Under this hypothesis the OP flowers that fail to set fruit contribute to plant reproductive success by siring seeds on other plants. Furthermore the male reproductive success in OP flowers is increased in larger plants and this increase exceeds any gains to the OP female function due to increased plant size.

3. The scale and extent of small-scale population genetic structure influences the relative advantage of OP due to avoidance of inbreeding depression (ID). Under this hypothesis the relative inbreeding of outcrossed progeny will decrease with the distance between the parents leading to the increased ID. Additionally, the expression of ID will be enhanced with increased dispersal distance of progeny.

Beyond this study of *G. clandestina*, it will be of interest to further investigate the interspecific differences in perennial strategy of size-dependent allocation to cleistogamy as well as the role of perennial-specific factors, like delayed costs of reproduction, on the conditional investment in cleistogamy. More population genetic studies of CLOP species are also needed to better understand the interplay between small-scale population genetic structure and inbreeding in OP progeny. The findings in *G. clandestina* emphasise that small differences in seed mass between cleistogamy and chasmogamy may have important implications for juvenile fitness and should not be
ignored. Most importantly, the male function hypothesis for OP needs urgent and thorough experimental evaluation in CLOP species. *G. clandestina* may well represent an ideal model system for such an investigation.
References


Landry, P.-A., Koskinen, M.T. and Primmer, C.R. (2002) Deriving evolutionary relationships among populations using microsatellites and (delta mu)^2: all loci are equal, but some are equal than others... *Genetics* 161: 1339-1347.


179
Waller, D.M. (1984) Differences in fitness between seedlings derived from
cleistogamous and chasmogamous flowers in *Impatiens capensis*. *Evolution* 38:
427-440.

Waller, D.M. (1985) The genesis of size hierarchies in seedling populations of


N.W., *The Natural History of Inbreeding and Outbreeding - Theoretical and

Waller, D.M. and Knight, S.E. (1989a) Genetic consequences of outcrossing in the
cleistogamous annual *Impatiens capensis*. II. Outcrossing rates and genotypic

Waller, D.M. and Knight, S.E. (1989b) Genetic consequences of outcrossing in the
cleistogamous annual, *Impatiens capensis*.III. Interlocus associations. *Heredity*

in angiosperms. In Thornhill, N.W., *The Natural History of Inbreeding and
Outbreeding - Theoretical and Empirical Perspectives*. The University of


Weatherwax, P. (1928) Cleistogamy in two species of *Danthonia*. *Botanical Gazette*

J. and Lovett-Doust, L., *Plant reproductive ecology*. Oxford University Press,
New York.228-245.


effects on seedling characteristics of *Solidago altissima* Compositeae. *American


Appendix

Cross-species amplification from crop soybean *Glycine max* provides informative microsatellite markers for the study of inbreeding wild relatives

*Note:* Acronyms for the two morphotypes included in the study differ between the thesis and the published paper. The broad leaf morphotype (BRO) corresponds to the high altitude morphotype (HAM) in the paper, and the narrow leaf morphotype (NAR) corresponds to the high altitude morphotype (HAM).
Cross-species amplification from crop soybean Glycine max provides informative microsatellite markers for the study of inbreeding wild relatives

K. Hempel and R. Peakall

Abstract: The development of microsatellite markers through transfer of primers from related species (cross-species amplification) remains a little-explored alternative to the de novo method in plants. In this study of 100 microsatellite loci from Glycine max, we examined two aspects of primer transfer. First, we tested if source locus properties can predict primer transfer and polymorphism in Glycine cyrtoloba and Glycine clandestina. We transferred 23 primers to G. cyrtoloba and 42 to G. clandestina, with 19 loci polymorphic within G. clandestina. However, we could not predict transfer or polymorphism from the source locus properties. Second, we evaluated the subset of 11 polymorphic loci for study in G. clandestina populations representing two local morphotypes. All loci were informative within populations (population mean \( H_s \pm SE = 0.58 \pm 0.04 \)). We directly sequenced 28 alleles at 4 representative loci. The allelic patterns and sequencing results established that 8 of 11 loci were typical microsatellites, confirming the utility of primer transfer as an alternative to de novo development. Additionally, we found that morphotypic differentiation between populations was paralleled by changes in polymorphism level at six loci and size homoplasy at one locus. We interpret these patterns as being a product of selfing in G. clandestina. Our results demonstrate the value of allele sequence knowledge for the most effective use of microsatellites.

Key words: microsatellite transfer predictability, cross-species amplification, Glycine, selfing, size homoplasy.

Introduction

The extensive polymorphism of microsatellites makes these genetic markers an ideal choice for studies in population genetics. This is particularly true for species that are naturally depauperate in genetic variation, like habitually selfing plants, where the traditional population genetics markers, allozymes, often fail to be informative (Jarne and...
Lagoda 1996; Cruzan 1998; Green et al. 2001). However, the widespread application of microsatellites is limited by a requirement for species-specific primers (Ashley and Dow 1994), the development of which remains a costly and lengthy process despite the continuing improvements in its efficiency (Zane et al. 2002). Transfer of primers (cross-species amplification) can offer an alternative to de novo development in plants (Peil et al. 1998, Rossetto 2001), with transfer rates ranging from 90% within subgenera to 35% within the family, and potential polymorphism from 78 to 58%, respectively (Rossetto 2001).

In an early recognition of the promise of cross-species transferability, Ashley and Dow (1994) noted that a search for the correlates of success might serve as a guide for more efficient transfer of microsatellite primers. Indeed, recent reviews of the patterns in microsatellite evolution predict that locus properties, for example repeat unit type or array length, should correlate with increased locus conservation and polymorphism (Chambers and MacAvoy 2000; Schlüterer 2000). Some studies in animals attempted to find the predictors of transfer success (Moore et al. 1991; Valsecchi and Amos 1996; Primmer et al. 1996), but there appear to be no similar attempts in plants, perhaps owing to the smaller size of most plant primer sets.

Even though a growing number of plant studies report some information on cross-species amplification tests, the tests of polymorphism at transferred loci are usually limited (Rossetto 2001). There are also presently few studies that have used cross-transferred microsatellite loci to address questions within plant populations (e.g., Nagamitsu et al. 2001; Bakker et al. 2001). Consequently, we know little about the behaviour and utility of transferred loci within and across populations at regional scales.

In this study, we exploit large numbers of microsatellite loci available in the commercial crop soybean (*Glycine max*) to examine the correlates of successful cross-species amplification of informative microsatellite loci. We also investigate the properties of 11 polymorphic loci within and among populations at a regional scale, including patterns of polymorphism as well as their underlying sequence variation at a selection of four representative loci. This follows an earlier study by Peakall et al. (1998) of soybean primer transfer among legumes which confirmed potential for their utility within the genus *Glycine*.

**Materials and methods**

**Study species**

*Glycine clandestina* and *Glycine cytlooba* are both endemic Australian members of the perennial soybean subgenus *Glycine*, closely related to the annual subgenus *Soja*, containing the crop *Glycine max* (Tindale 1984). The two perennial species are diploid representatives of the genome groups at the extremes of subgeneric differentiation (Doyle et al. 1990) — *G. clandestina* belongs to group A and *G. cytlooba* to group C (Hymowitz et al. 1998). There is some evidence that *G. cytlooba* is more genetically diverged from *Soja* than *G. clandestina* (Zhu et al. 1995). Both species employ a cleistogamous reproductive strategy that combines normal open-pollinated flowers with cleistogamous flowers specially modified for self fertilization (Tindale 1991). Selfing in *G. clandestina* populations is estimated to range from about 70 to 90%, with both types of flowers contributing to these high rates of inbreeding (Schoen and Brown 1991a; K. Hempel, unpublished data).

*Glycine clandestina* is a widespread species of southeastern Australia and exhibits a number of prominent regional leaf morphotypes throughout its range (Peil et al. 2001). We focused on seven populations over a narrow geographic range of approximately 35 km representative of two morphotypes within the Canberra region, Australian Capital Territory. Four populations were characteristic of the morphotype of higher altitudes within the Great Dividing Range (Peil et al. 2001) and three of the lower altitude morphotype (see Table 1). In the remainder of the paper we refer to these morphotypes as HAM and LAM, respectively.

**Study design**

**Primer transfer**

We selected 100 soybean primer pairs from among the (AT), and (ATT), motif-based set developed by Cregan et al. (1999) in *Glycine max* ‘Williams’ (Cregan et al. 1999), and purchased them commercially from Research Genetics Inc. (Huntsville, Ala.). Using a primer set from a single source and amplifying the same type of loci avoided confounding effects in the analysis of locus properties in relation to primer transfer. The loci were spread across the soybean linkage groups and were known to be variable among soybean cultivars (Cregan et al. 1999). The trinucleotide microsatellites were preferentially selected, representing 87 of microsatellites tested, because they tend to produce more easily interpretable profiles with fewer multiple-b and artifacts (Smith and Dewey 1994; Rongwen et al. 1995).

The transfer test array included a single individual of each of the target species (*G. clandestina* and *G. cytlooba*), and a representative of the source species (*Glycine max* ‘Stephens’) as a positive control. Only primers successfully transferred to *G. clandestina* were further subjected to a preliminary test of polymorphism at the species level by screening eight individuals, four from each of one high- and one low-altitude population (B1 and M, respectively).

**Locus utility within populations of G. clandestina**

**Levels and patterns of polymorphism**

We explored levels and patterns of polymorphism at 11 of the loci initially diagnosed as polymorphic in *G. clandestina* to assess their utility for within-population study in Canberra region. The loci were chosen so that their products had non-overlapping size ranges to facilitate combining them at electrophoresis within a single lane for the routine genotyping. First, we genotyped 12 individuals from each of the two high-altitude (F and B2) and low-altitude (A and T) populations. Sampling was then extended to other populations only for the loci polymorphic in all four populations. Final sample numbers ranged from 15 to 60 individuals/population/locus.

Progeny arrays of three to six seeds (either open pollinated or cleistogamously selfed) from known heterozygous maternal plants were screened for some allelic combinations at all of the polymorphic loci to confirm that patterns of polymorphism corresponded to codominant alleles segregating in Mendelian fashion.
Molecular basis of polymorphism

Based on the outcomes of the survey of polymorphism in *G. clandestina*, we selected 4 of the 11 loci for the investigation of array sequence at selected alleles within and among populations. The four loci were chosen because they provided good representation of the observed range in both overall levels of polymorphism and the patterns of variation in levels of polymorphism among the populations. We sampled with three objectives in mind. First, to confirm that the amplified fragments were homologous to *G. max* and that they contained microsatellite arrays. Second, to examine the molecular basis of allele size variation. Third, to exclude the possibility that variation in polymorphism levels across populations had basis in array interruptions (sensu Chambers and MacAvoy 2000), which affect inherent locus mutability through polymerase slippage (Schlötterer 2000).

Molecular methods

DNA extraction

Leaf samples were collected from 20 to 60 individuals sampled from the full length of each population and stored at −70°C before DNA extraction. For the larger quantities of high quality DNA required for the primer transfer tests, the sodium dodecyl sulphate (SDS)-based DNA extraction method described in Peakall et al. (1998) was used. The more rapid mini-prep cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1990) was used for processing the remaining samples.

PCR conditions and product visualization

In the primer transfer and initial screening for polymorphism at a regional level, we used 20-μL PCRs modified from Peakall et al. (1998) containing 0.2 μM of each primer, 200 μM of each dNTP, 1.5 mM MgCl₂, 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 0.2 μg bovine serum albumin, 1.6 U AmpliTag (Applied Biosystems, Foster City, Calif.), and approximately 40 ng of template DNA. PCR was performed on a Corbett Research (Sydney, Australia) FTS-960 thermal sequencer using the touch-down profile of Peakall et al. (1998) with the annealing temperature being reduced by 1°C every two cycles from 55 to 50°C.

In tests for primer transfer, the PCR products were electrophoresed through 2% w/v agarose gels in a 0.5x TBE (Tris–borate, EDTA (pH 8.0)) buffer and stained with ethidium bromide. We classified products into three quality classes as follows: successful transfer, when one or two clear strong bands were produced for both source and target species; sub-optimal transfer, weak or multiple and (or) complex banding patterns; transfer failure, no product. In the screening for polymorphism presence at the transferred loci in *G. clandestina*, the PCR products were resolved using vertical electrophoresis at 4°C through 3–4% w/v Super Fine agarose (Amresco, Solon, Ohio) in 1x TBE and visualized with ethidium bromide staining. Those primers producing simple fragments of different sizes, within individuals and (or) between them, were deemed to be polymorphic.

To assess polymorphism levels and patterns within populations of *G. clandestina* at a subset of 11 of the polymorphic loci, fluorescently labeled fragments were visualized and sized on 4.25% w/v denaturing polyacrylamide gels containing 6 M urea and run with a 1x TBE buffer on an ABI Prism 377 automated sequencer (Applied Biosystems). In this case, the DNA fragments were labeled during PCR of 10-μL reactions with one of three fluorescent dUTPs (TAMRA, R6G, or R110 (Perkin-Elmer)) by adding the [3H]dUTPs to the standard PCR mixture at a final concentration of 2 μM for TAMRA and 0.5 μM for R6G and R110. ABI Prism Genescan and Genotyper software (Applied Biosystems) were used to aid scoring.

In preparation for sequencing, alleles of known size either in homozygous or in heterozygous state where alleles were widely separated were electrophoresed on 2% w/v agarose gels with 1x TAE (Tris–acetate, EDTA (pH 8.0)) buffer and visualized with ethidium bromide. The alleles were excised from the gel and the DNA and agarose separated by homemade filters before DNA precipitation and washing in 70% v/v ethanol. The clean PCR products were quantified and directly sequenced in forward and reverse direction using the respective original PCR primers and the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech, Piscataway, N.J.) on a Corbett Research FTS-960 thermal sequencer following the manufacturers instructions. Electrophoresis was performed on 5% w/v denaturing polyacrylamide gels containing 6 M urea and run with a 1x TBE buffer on an ABI Prism 377 automated sequencer. We aligned *G. clandestina* allele sequences with those of 'Williams' (from GeneBank) or 'Stephens' (our sequence) by hand using Sequencher software (Gene Codes Corp., Ann Arbor, Mich.).

Data analysis

Primer transfer

Based on the known locus properties in *G. max* (sourced from Cregan et al. 1999, and Cregan unpublished), we examined whether there were any predictive variables in the...
target species for quality of product (three levels – sub-optimal transfer, transfer success, and failure), presence of polymorphism (two levels – polymorphic or invariant in preliminary screen in *G. clandestina*), and universality of transfer to both target species (four levels – success in both species, failure in both species, sub-optimal product in both species, and inconsistent product quality across the two species).

We included three source locus properties in the analysis:

(i) Repeat unit (motif) type. Many studies suggest unit length influences mutation rates, presence, and levels of polymorphism (Weber and Wong 1993; Chakraborty et al. 1997; Rossetto et al. 1999; Butcher et al. 2000) and transfer success (Kutil and Williams 2001). Therefore, we needed to confirm the two types of loci (di- and tri-repeat based) did not differ in transfer success and (or) polymorphism presence before further analysis.

(ii) Total length of the flanking region. The total length of the flanking region in the source was included because it has been shown that mutation rates may be enhanced in the immediate vicinity of microsatellite regions (e.g., Brohede and Ellegren 1999). Therefore, we tested if the conservation of priming sites between source and target was more likely for loci where primers are placed further from the array (total flanking region was longer).

(iii) Repeat unit number (array length). Array length is known to correlate well with observed polymorphism at loci amplified by de novo primers (e.g., Weber 1990; Grist et al. 1993; Thomas and Scott 1993). We confirmed this was the case for *G. max* with a significant positive relationship detected between the array length and the ascertainment of expected heterozygosity reported for soybean by Cregan et al. (1999) ($F = 4.7174$, $p = 0.0385$). We tested if a greater number of repeats in the source would be associated with presence of polymorphism in the target.

All data were analyzed using JMP v. 3.0.1 (SAS Institute Inc.), with alpha value set to 0.05. A $\chi^2$ analysis was used to test for differences in proportions of the two types of repeat unit among classes of primer transfer (sub-optimal transfer and transfer failure pooled together to prevent low cell counts), transfer consistency, and polymorphism presence. To test for significance among means we used a one-way ANOVA (analysis of variance) for length of flanking region (normally distributed) and a non-parametric Wilcoxon – Kruskal – Wallis test for number of repeats that were not normally distributed.

**Locus utility within populations of *G. clandestina***

In *G. clandestina*, standard population genetic statistics including the number of alleles, allele frequency, the effective number of alleles, and observed and expected heterozygosity were calculated using GenAlEx version 5.0 (Peakall and Smouse 2001). Estimates excluded loci and populations that were invariant and were calculated at two levels: for the total sample and within each population.

**Results**

**Primer transfer from *G. max* yields good marker loci in *G. clandestina***

Of the 100 *G. max* microsatellites tested, 42 primer pairs produced successful amplification in *G. clandestina* and 28 in *G. cytisola*. Within *G. clandestina*, 19 (44%) of these amplified polymorphic products among the 8 individuals tested (Table 2).

The subset of 11 loci surveyed in detail within *G. clandestina* was highly variable with a total of 112 alleles detected. There was variation among the loci in levels of polymorphism: three loci exhibited just two alleles each, whereas at the remaining loci a total of between 7 and 27 alleles were detected (Mean $\pm SE = 10.25 \pm 2.2$) among all of the individuals analysed (Table 3). Both the di- and multi-allelic loci were informative within populations (mean $H_s \pm SE = 0.44 \pm 0.02$ and $0.63 \pm 0.04$, respectively).

The patterns of allelic variation at 10 of the loci were consistent with the presence of microsatellite arrays, with repeat unit size as expected from *G. max* — the dinucleotide loci showed allelic size variation in units of two and the trinucleotide loci in units of three base pairs. Three of these loci (Sat373, Sat064, and Sat1155) were diallelic, with both alleles considerably shorter than the minimum-sized allele observed in *G. max*, suggesting degeneration of the array through a decrease in the number of repeats. At Sat199, contrary to expectations, alleles differed in size by two base pairs, consistent with a dinucleotide repeat, whereas in *G. max* the locus was described as a trinucleotide repeat, suggesting a change in the length of the repeat unit.

In our selection for sequencing, we represented one diallelic locus (Sat373), two loci with typical microsatellite variability patterns (Sat131 and Sat040), and Sat199 because of the apparent change in its unit type. Comparisons of the flanking regions (not shown in full) clearly confirm the homology of three loci in *G. clandestina* with *G. max* (Fig. 1). All of the four loci were also confirmed to contain microsatellite arrays (Table 4).

At Sat373 the long trinucleotide array ([AAT]$_{23}$ in ‘Williams’) was reduced to (AAT)$_3$ in *G. clandestina* and the allele variation proved to be a consequence of a 7-bp insertion–deletion (indel) in the flanking region, not variation within the array (Fig. 1; Table 4). The two other diallelic loci may contain similarly contracted arrays; although, in contrast to Sat373, allele size differed by one repeat unit, suggesting the variation in array length was responsible.

At the other three sequenced loci, variation in allele size in *G. clandestina* was due only to variation in the number of repeat units (Table 4), but all three loci exhibited modifications in array structure compared with *G. max* (Fig. 1). The Sat199 locus in *G. max* contained a compound (ATT)$_2$(AT)$_n$ array dominated by trinucleotides, whereas in *G. clandestina*, the trinucleotide repeats were absent and the dinucleotide array was expanded (Fig. 1). The switch in unit length clearly explains the unexpected allele size pattern in *G. clandestina*. At Sat040 the array in *G. max* again was compound (AT)$_2$(AG)$_2$ in contrast to perfect (simple) (AT)$_n$ in *G. clandestina*. Thus, at these two loci the array was simpler in structure in the target as opposed to the source species. At Sat131 the array of the sequenced *G. clandestina* alleles consisted of the same dinucleotide TA unit as in *G. max*. However, as described in more detail below, some of the *G. clandestina* alleles contained a compound (TA)$_n$(T)$_n$ array (Table 4; Fig. 1).
Irrespective of the nature of the loci, the patterns of allelic segregation within progeny confirmed inheritance and codominance at all of the 11 loci, proving them to be reliable molecular markers in G. clandestina.

Source locus characteristics do not predict product quality and polymorphism in target

The type of repeat unit in the source influenced neither primer transfer nor presence of polymorphism. There was no significant difference between primers amplifying loci with (AT)n and (ATT)n motifs in the proportion of primers successfully transferred to each of the target species (G. cyrtoloba \( \chi^2 = 1.941, p = 0.163 \); G. clandestina \( \chi^2 = 1.082, p = 0.298 \)), nor the presence of polymorphism in G. clandestina \( \chi^2 = 0.085, p > 0.5 \). The two types of loci were thus combined for further analysis.

We found no evidence that the longer flanking regions in the source facilitated primer transfer success or transfer consistency. Nor was there any significant difference in the mean flanking region length among primers with the three classes of product quality in both of the target species (G. clandestina \( F = 0.329, \text{d.f.} = 2, p > 0.5 \); G. cyrtoloba \( F = 2.708, \text{d.f.} = 2, p = 0.07 \)) or among the four primer classes of consistency in product quality \( F = 1.168, \text{d.f.} = 3, p > 0.33 \).

There was also no evidence that polymorphism presence in the source was affected by the number of repeats in the target, because there was no significant difference in array unit number in G. max between primers with invariant and polymorphic products in G. clandestina (Kruskal–Wallis normal approximation \( Z = -0.458, p > 0.5 \)).

Locus utility in G. clandestina varies among populations

Variation in levels of polymorphism

The utility of any given locus was heterogenous across the populations of G. clandestina. Some population by primer combinations either resulted in amplification difficulties (Satt350 and Satt228), invariant products (all of the di-allelic loci Sat064, Satt373, Satt155; and multi-allelic Satt163, Satt199), or reduction in number of alleles (e.g., Satt040) (Table 3). Monomorphism and poor amplification meant that only some of the 11 polymorphic loci were identified as informative within any given population, ranging from 3 to 10 per population. Interestingly, the changes in allele numbers across populations were not synchronized among loci, but largely corresponded to the split of populations between LAM and HAM. In LAM populations, Satt199 and Sat064 were monomorphic and Sat040 was much less polymorphic, whereas in HAM populations, Satt373 and Satt155 were monomorphic. Also, the total diversity at any of the loci on average exceeded the diversity recorded within populations by at least two fold (Table 3), indicating a considerable lack of overlap in allele sizes and numbers among populations.

Variation in allele sequence

We sequenced alleles from at least two populations per morphotype (Table 4) at Satt199 and Sat040 as representatives of loci with polymorphism restriction and Sat131 as a representative of uniform polymorphism levels across the two morphotypes. At Sat040 and Sat199, variation in size of sequenced fragments within populations, among populations within morphotypes, and between morphotypes was due to variation in number of units in the microsatellite arrays (Table 4). Therefore, the possibility that changes in locus polymorphism across morphotypes arose because of inherent differences in mutability owing to array interruptions at these loci was rejected. At Sat131, we found difference in the array sequences between the two morphotypes (Table 4; Fig. 2) that was not apparent from the polymorphism patterns. The pure \( (\text{AT})_n \) arrays were characteristic of the LAM, and the compound \( (\text{AT})_n(\text{T})_n \) arrays were characteristic of the HAM populations. For the LAM, within- and between-population variation in fragment size resulted from variation in the number of AT units in the pure array. In contrast, for the HAM populations size, variation was a result of variation in the number of AT and T units in the compound array (Table 4). Retrospectively, we noted some evidence for the differences in array structure between morphotypes in the DNA profiles (Fig. 2). Because of the complexity of the array structure at Sat131, we observed a case of the same size variants of 210 and 212 bp having divergent sequence between the two morphotypes. Such size homoplasies was also observed for a 212-bp allele between populations of HAM where different numbers of AT and T units combined to produce the same length of array (Table 4; Fig. 2).

Discussion

Our findings confirm the utility of primer transfer as an alternative source to de novo development when microsatellites are available in related species, but we were unable to find simple predictors of successful transfer and polymorphism. Primer transfer from G. max resulted in suc-
cessful development of a set of informative microsatellite loci for study of a selfing wild relative of the soybean, *G. clandestina*. Morphotypic differentiation between *G. clandestina* populations was paralleled by changes in the polymorphism level at six loci and phylogenetically informative size homoplasy at one locus at a relatively small geographic scale of 40 km. We interpret this complexity as an example of genetic differentiation patterns generated by in-breeding in this species.

**Transfer as marker source**

Our tests of transfer for 100 Glycine *max* (subgenus Soja) microsatellites succeeded for 42 primer pairs in *G. clandestina* and 28 in *G. cyrtoloba* (subgenus Glycine) without any additional optimization of PCR conditions. These transfer rates were within the range reported for transfer between subgenera in other plants (e.g., 17–100% in *Pinus* (Echt et al. 1999); 65-76% in *Pruinus* (Cipriani et al. 1999); 22–78% in *Quercus* (Isagi and Suhandono 1997)). However, the transfer success rate for *G. clandestina* was lower than the 61% observed for another *G. max* primer set (Peakall et al. 1998). This difference may be partially attributed to different criteria of successful transfer. Unlike Peakall et al. (1998) who focused on homology of loci between source and target irrespective of amplification quality, our criteria rejected weakly amplifying primers to avoid optimization of PCR conditions before application. The discrepancy disappears when we count weakly amplifying primers as successful, because it raises the transfer success rate to 55%.

For *G. clandestina*, 19 of the transferred microsatellite markers showed potential for polymorphism. When we examined 11 of these loci in detail across 7 *G. clandestina* populations, they proved highly informative for within population study. In fact, the mean per locus within population heterozygosity at polymorphic loci (mean ± SE = 0.58 ± 0.04) placed them above the known maximum for sets of de novo microsatellites in other predominant selfers — 0.33–0.51 (Viard et al. 1997; Viard et al. 1996; Kuitinen et al. 1997; Sun et al. 1998; Bonnin et al. 2001; Green et al. 2001; Awadalla and Ritland 1997; Kelly and Willis 1998) — and closer to the range of 0.68 to 0.84 observed in predominant outcrossers (Terauchi and Konuma 1994; Heurert et al. 2001; Ueno et al. 2000; Dow and Ashley 1998; Nagamitsu et al. 2001), indicating that microsatellites derived from transfer can be as informative as those derived de novo.

Of the 11 putative microsatellite loci in *G. clandestina*, 8 were likely to be typical microsatellites as suggested by multiallelic patterns with variation in allele length by the multiples of the repeat unit. The microsatellite character was confirmed by sequencing at three of these loci. At the further three loci, diatypism atypical of microsatellites suggested degeneration of the repeat region, which was confirmed by sequencing in one of them. We predict that consistently diatypic microsatellite loci as commonly reported in other studies both using de novo and transferred markers (Awadalla and Ritland 1997; van Treuren et al. 1997; Viard et al. 1996; Sun et al. 1998) are likely to contain similarly degenerated arrays. These atypical microsatellites are usually retained in studies of selfers because they provide good information within populations (mean $H_e$ ± SE = 0.44 ± 0.02), and fit well in the multiproduct combination within ABI Genescan lanes for visualization. Their application might be restricted, however, because they are more susceptible to loss of polymorphism across populations and are likely to violate the step-wise mutation assumption of microsatellite-based statistics. Irrespective of their character, all of the loci had codominant inheritance essential for their utility as markers.

When considering both the allelic patterns and sequencing results, all of the loci were likely to be homologous to those in *G. max*. At the multi-allelic loci, sequencing showed preservation of uninterrupted arrays for three of them, even though the array structure changed from compound in the source to pure in the target.

**Predictability of transfer**

Despite the large set of microsatellite loci tested, we found neither predictors for primer transfer success and consistency, nor for the polymorphism present in the source species. While there are relatively few similar tests, other attempts to find simple predictors of transfer (Moore et al. 1991; Primmer et al. 1996) and polymorphism in target (Moore et al. 1991; Valsecchi and Amos 1996; Primmer et al. 1996) have also failed. In contrast to other studies we used flanking region properties rather than properties of the repeat region itself to predict transfer success and its consistency, under the assumption that they are more directly related to priming site conservation. Contrary to expectations, we did not find the shorter flanking regions in *G. max* to be associated with transfer success in either *G. clandestina* or *G. cyrtoloba*. We also did not find that polymorphism potential within the source species (repeat number) predicted the presence of polymorphism in *G. clandestina*. The universal lack of any pattern across studies and organisms suggests that diverse factors influence the transferability of informative microsatellite loci; for example, the variation in genome-wide or locus-specific mutation rates in neutral regions among taxa (Brohede and Ellegren 1999; Kruglyak et al. 1998).

The only consistent pattern evident in studies to date is the decline in transfer rates and polymorphism with increasing evolutionary distance from the source species (e.g., Dayanandan et al. 1997; Peakall et al. 1998; Roa et al. 2000). Our finding of lower transfer success in *G. cyrtoloba* compared with *G. clandestina* is consistent with this general pattern since it reflects the relative genetic distance of the two targets from *G. max* (Zhu et al. 1995). Thus, apart from the consideration of evolutionary distance, the streamlining of a test for primer transfer may be unachievable.

**Complexity of allelic patterns and sequence variation in selfers**

Of the 11 informative loci we examined in detail in *G. clandestina*, 5 were monomorphic and 1 experienced three-fold reduction in allele numbers in some of the 7 populations. These changes in polymorphism occurred mostly between the LAM and HAM populations, but the patterns across loci were not strictly synchronous across the morphotypes. Sequencing of two of these loci confirmed the loss of polymorphism was not likely to be due to interruptions in the arrays, which can diminish their inherent muta-
Table 3. Patterns of genetic variation recorded in 7 *Glycine clandestina* populations at 11 polymorphic microsatellite loci derived by primer transfer from *G. max*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Habitat morphotype</th>
<th>Population</th>
<th>N</th>
<th>$N_s$</th>
<th>$N_B$</th>
<th>I</th>
<th>$H_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt478</td>
<td>LAM</td>
<td>T</td>
<td>53</td>
<td>4</td>
<td>2.69</td>
<td>1.10</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>30</td>
<td>3</td>
<td>2.20</td>
<td>0.86</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>36</td>
<td>3</td>
<td>2.38</td>
<td>0.98</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B1</td>
<td>25</td>
<td>6</td>
<td>2.87</td>
<td>1.35</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE</td>
<td>18</td>
<td>2</td>
<td>1.60</td>
<td>0.56</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B2</td>
<td>24</td>
<td>3</td>
<td>1.40</td>
<td>0.54</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>50</td>
<td>6</td>
<td>1.76</td>
<td>0.91</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>33.71 (5.06)</td>
<td>3.86 (0.59)</td>
<td>2.13 (0.21)</td>
<td>0.90 (0.11)</td>
<td>0.50 (0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>236</td>
<td>4</td>
<td>4.00</td>
<td>1.60</td>
<td>0.75</td>
</tr>
<tr>
<td>Satt040</td>
<td>LAM</td>
<td>T</td>
<td>49</td>
<td>4</td>
<td>1.93</td>
<td>0.92</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>28</td>
<td>3</td>
<td>2.65</td>
<td>1.03</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>38</td>
<td>3</td>
<td>1.87</td>
<td>0.79</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B1</td>
<td>23</td>
<td>15</td>
<td>11.14</td>
<td>2.54</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE</td>
<td>21</td>
<td>11</td>
<td>6.17</td>
<td>2.06</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B2</td>
<td>30</td>
<td>13</td>
<td>7.66</td>
<td>2.25</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>42</td>
<td>18</td>
<td>10.72</td>
<td>2.58</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>33.00 (3.90)</td>
<td>9.57 (2.35)</td>
<td>6.02 (1.51)</td>
<td>1.74 (0.30)</td>
<td>0.73 (0.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>231</td>
<td>27</td>
<td>7.16</td>
<td>2.58</td>
<td>0.86</td>
</tr>
<tr>
<td>Satt462</td>
<td>LAM</td>
<td>T</td>
<td>57</td>
<td>6</td>
<td>2.24</td>
<td>1.04</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>27</td>
<td>9</td>
<td>4.30</td>
<td>1.79</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>41</td>
<td>7</td>
<td>2.61</td>
<td>1.28</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B1</td>
<td>27</td>
<td>9</td>
<td>5.70</td>
<td>1.92</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE</td>
<td>22</td>
<td>8</td>
<td>4.99</td>
<td>1.81</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B2</td>
<td>44</td>
<td>11</td>
<td>7.17</td>
<td>2.14</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>26</td>
<td>7</td>
<td>3.53</td>
<td>1.58</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>35.29 (5.16)</td>
<td>8.29 (2.35)</td>
<td>4.25 (0.72)</td>
<td>1.62 (0.17)</td>
<td>0.70 (0.07)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>247</td>
<td>18</td>
<td>5.88</td>
<td>2.25</td>
<td>0.83</td>
</tr>
<tr>
<td>Satt131</td>
<td>LAM</td>
<td>T</td>
<td>57</td>
<td>6</td>
<td>2.24</td>
<td>1.04</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>27</td>
<td>9</td>
<td>4.30</td>
<td>1.79</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>41</td>
<td>7</td>
<td>2.61</td>
<td>1.28</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B1</td>
<td>27</td>
<td>9</td>
<td>5.70</td>
<td>1.92</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE</td>
<td>22</td>
<td>8</td>
<td>4.99</td>
<td>1.81</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B2</td>
<td>44</td>
<td>11</td>
<td>7.17</td>
<td>2.14</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>26</td>
<td>7</td>
<td>3.53</td>
<td>1.58</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>35.86 (4.83)</td>
<td>8.00 (0.93)</td>
<td>4.14 (0.84)</td>
<td>1.56 (0.17)</td>
<td>0.70 (0.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>251</td>
<td>15</td>
<td>7.90</td>
<td>2.30</td>
<td>0.87</td>
</tr>
<tr>
<td>Satt350</td>
<td>LAM</td>
<td>A</td>
<td>18</td>
<td>6</td>
<td>3.43</td>
<td>1.43</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>30</td>
<td>4</td>
<td>1.61</td>
<td>0.70</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>24 (6)</td>
<td>5 (1)</td>
<td>2.52 (0.91)</td>
<td>1.07 (0.37)</td>
<td>0.54 (0.16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>48</td>
<td>7</td>
<td>2.55</td>
<td>1.21</td>
<td>0.61</td>
</tr>
<tr>
<td>Satt163</td>
<td>LAM</td>
<td>T</td>
<td>27</td>
<td>14</td>
<td>9.41</td>
<td>2.39</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F</td>
<td>24</td>
<td>6</td>
<td>2.80</td>
<td>1.23</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>25.50 (1.5)</td>
<td>10.00 (4)</td>
<td>6.10 (3.30)</td>
<td>1.81 (0.58)</td>
<td>0.77 (0.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>51</td>
<td>16</td>
<td>8.42</td>
<td>2.38</td>
<td>0.88</td>
</tr>
<tr>
<td>Satt288</td>
<td>LAM</td>
<td>A</td>
<td>16</td>
<td>6</td>
<td>3.28</td>
<td>1.38</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>28</td>
<td>4</td>
<td>1.20</td>
<td>0.40</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Population mean (SE)</td>
<td>22.00 (6)</td>
<td>5.00 (1)</td>
<td>2.24 (1.04)</td>
<td>0.89 (0.49)</td>
<td>0.43 (0.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>44</td>
<td>7</td>
<td>2.29</td>
<td>1.16</td>
<td>0.56</td>
</tr>
<tr>
<td>Satt199</td>
<td>HAM</td>
<td>F</td>
<td>30</td>
<td>9</td>
<td>5.49</td>
<td>1.86</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Produced with permission of the copyright owner. Further reproduction prohibited without permission.
tion rates and therefore allelic diversity. Concurrent morphological and genetic differentiation among populations at short geographic distances is expected in selfing species (Loveless and Hamrick 1984; Hamrick and Godt 1989; Schoen and Brown 1991b; Hamrick and Godt 1996), where interpopulation gene flow is restricted and the effects of genetic drift (Crow and Kimura 1970) and linkage (Hedrick 1985) are pronounced. Despite the homogenizing effect of their relatively high mutation rates, genetic diversity at microsatellite loci is vulnerable to those processes both in theory (Nauta and Weissing 1996) and in practice (Kelly and Willis 1998; Awadalla and Ritland 1997; Via 1996; Sun et al. 1998; Todokoro et al. 1995; Kuittinen et al. 1997; Green et al. 2001). Given the pronounced intermorphotype loss of polymorphism at a short geographic distance and the multitude of leaf morphotypes across *G. clandestina*'s range (Pfeil et al. 2001), it is likely that the variation in polymorphism levels will affect the application of microsatellites across the species range. Equally, such variation should be a consideration for development of informative markers in other selfers.

Apart from the more obvious changes in locus polymorphism, we also revealed a case of cryptic inter-morphotype differentiation on sequencing of the Sat131 locus. Despite the uniform polymorphism levels and consistency in allele size, the array structure at this locus differed between LAM and HAM, with pure array in the former and the compound array in the latter populations. The difference in array structure caused size homoplasy between morphotypes. Additionally, the compound nature of the array found in HAM populations resulted in size homoplasy among populations of this morphotype. It is likely that the compound nature of the array may make homoplasy apparent also within populations of HAM. Size homoplasy owing to variation in array sequence is an observable case of identity in state (IIS) of alleles which violates the assumption of identity by descent (IBD) implicit in many population genetic applications of microsatellites (Jarne and Lagoda 1996; Summucks 2000). The prevalence of size homoplasy at both inter- and intra-population levels has been increasingly appreciated from studies of differences in allele sequence in flanking regions and of loci with interrupted and compound arrays (e.g., Estoup et al. 1995; Via et al. 1996; Makova et al. 2000), yet little is known about its effects on bias in estimates of genetic parameters (but see Angers et al. 2000). However, Angers et al. (2000) demonstrated that knowledge of homoplasy in compound arrays does increase information content of microsatellite loci in selfers. Our results provide further evidence that array sequence contains valuable phylogenetic information and highlight the importance of investigations of homoplasy for microsatellite applications in population genetic studies, irrespective of the source of microsatellites.

**Conclusions**

Cross-species amplification of microsatellite loci, in our case involving a native Australian plant species, has saved considerable costs and time compared with de novo develop-

© 2003 NRC Canada

Table 3. (concluded).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Habitat morphotype</th>
<th>Population</th>
<th>N</th>
<th>Diversity measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$N_a$</td>
</tr>
<tr>
<td>B2</td>
<td></td>
<td>15</td>
<td>3</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>Population mean (SE)</td>
<td>22.50 (7.5)</td>
<td>6.00 (3)</td>
<td>3.74 (1.75)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>45</td>
<td>9</td>
<td>4.47</td>
</tr>
<tr>
<td>Satt064</td>
<td>HAM</td>
<td>F</td>
<td>23</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B2</td>
<td>28</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>Population mean (SE)</td>
<td>25.50 (2.5)</td>
<td>2.00 (0)</td>
<td>1.85 (0.06)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>51</td>
<td>2</td>
<td>1.86</td>
</tr>
<tr>
<td>Satt373</td>
<td>LAM</td>
<td>A</td>
<td>18</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>29</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>Population mean (SE)</td>
<td>23.50 (5.5)</td>
<td>2.00 (0)</td>
<td>1.64 (0.11)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>47</td>
<td>2</td>
<td>1.67</td>
</tr>
<tr>
<td>Satt155</td>
<td>LAM</td>
<td>A</td>
<td>18</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>30</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>Population mean (SE)</td>
<td>24.00 (6)</td>
<td>2.00 (0)</td>
<td>1.85 (0.05)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>48</td>
<td>2</td>
<td>1.84</td>
</tr>
</tbody>
</table>

All loci (mean across loci within populations)  
Mean across loci total  
Di-allelic loci (mean across loci within populations)  
Mean across loci total  
Multi-allelic loci (mean across loci within populations)  
Mean across loci total

Note: HAM, high altitude; LAM, low altitude.

© 2003 NRC Canada

produced with permission of the copyright owner. Further reproduction prohibited without permission.
Fig. 1. Comparison of the flanking sequence at the three loci between the source (Glycine max) and the target species (G. clandestina). Glycine max sequences are from 'Stephens', as sequenced in this study (*), or 'Williams', as sequenced by Cregan et al. (1999) (**).

Table 4. Comparison of repeat region sequence among alleles at four of the polymorphic loci amplified in G. clandestina by primers transferred from G. max.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele size (bp)</th>
<th>Habitat</th>
<th>Population</th>
<th>Repeat region sequence</th>
<th>Length of repeat region (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sat131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>LAM</td>
<td>M</td>
<td>(AT)<em>{10}(T)</em>{2}</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>198</td>
<td>LAM</td>
<td>A</td>
<td>(AT)<em>{11}(T)</em>{2}</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>202</td>
<td>HAM</td>
<td>TE</td>
<td>(AT)<em>{11}(T)</em>{6}</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>LAM</td>
<td>M</td>
<td>(AT)<em>{12}(T)</em>{2}</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>212</td>
<td>LAM</td>
<td>A</td>
<td>(AT)<em>{12}(T)</em>{3}</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>216</td>
<td>LAM</td>
<td>B2</td>
<td>(AT)<em>{12}(T)</em>{10}</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>LAM</td>
<td>M</td>
<td>(AT)<em>{13}(T)</em>{2}</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Satt373</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>204</td>
<td>LAM</td>
<td>T</td>
<td>(ATT)_{3}...CAGATCG</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>LAM</td>
<td>T</td>
<td>(ATT)_{3}</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>196 (invariant)</td>
<td>HAM</td>
<td>B2</td>
<td>(ATT)_{3}</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Satt199</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>163 (invariant)</td>
<td>LAM</td>
<td>A</td>
<td>(AT)_{10}</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>163</td>
<td>LAM</td>
<td>T</td>
<td>(AT)_{10}</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>HAM</td>
<td>B2</td>
<td>(AT)_{12}</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>173</td>
<td>HAM</td>
<td>F</td>
<td>(AT)_{11}</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Sat040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>194</td>
<td>LAM</td>
<td>A</td>
<td>(AT)_{3}</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>202</td>
<td>HAM</td>
<td>F</td>
<td>(AT)_{14}</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>218</td>
<td>HAM</td>
<td>F</td>
<td>(AT)_{21}</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>224</td>
<td>HAM</td>
<td>F</td>
<td>(AT)_{24}</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

Note: Leaf morphotype: HAM, high altitude; LAM, low altitude.
*Alleles share a T→C transition within the microsatellite region.
1Allele sequenced twice from different individuals within the population.
Fig. 2. An example of the differences in sequence and DNA profiles for the alleles at the Sat131 locus that were identified between populations of *G. clandestina* representing the low (A) and high (B) altitude morphotype. Populations of the allele origin are identified on the right of each lane. Asterisk indicates alleles on which size homoplasy was identified.

### Acknowledgements

We would like to thank Dr. Perry Cregan for sharing the unpublished information on properties of the microsatellite loci in the source species *G. max*; Daniel Ebert and Simon Gilmore for technical assistance with sequencing; Drs. Lyn Cook and James Procter and anonymous reviewers for constructive comments on the manuscript. This work was carried out with financial support from Australian Postgraduate Award.

### References


© 2003 NRC Canada


