Population genetic analysis of *Lasthenia conjugens* (Asteraceae) (Contra Costa Goldfields) populations within Solano County, CA.

Jennifer M. Ramp Neale, Research Associate, University of Colorado at Boulder, Museum of Natural History, Boulder, Colorado, 80309. Current address: Associate Director of Research and Conservation, Denver Botanic Gardens, 909 York Street, Denver, CO. 80206. nealejr@gmail.com. 720-865-3562.

Introduction

Long-term persistence of a species depends on the maintenance of sufficient variation within and among populations to preserve evolutionary potential (Falk and Holsinger 1991 and references therein, Hedrick 2001, Frankham 2003). Evolutionary potential of a species refers to its ability to adapt to future environmental change (Falk and Holsinger 1991). Knowledge of genetic variation within and among populations of endangered plant species is essential for future management strategies, especially those that involve restoration or mitigation of the focal species (Hamrick et al. 1991). For species of high conservation priority, knowledge of genetic variation within and among populations may allow managers to capture the highest levels of variation available in nature for founding new populations or restoring populations that have suffered extensive damage or extirpation (Fleishman et al. 2001).

Contra Costa Goldfields (*Lasthenia conjugens* Greene) is a federally endangered self-incompatible annual plant. It is associated with vernal pool habitats, has an extremely limited distribution and is under intense development pressures. Previous genetic studies of *L. conjugens* found moderate levels of genetic differentiation among sampled populations spanning the range of the species. These results differ from those found in studies of other vernal pool organisms in that populations were only moderately divergent from one another as opposed to the high divergence found in other species (Ramp 2004, Ramp Neale et al. 2008). Previous studies examining genetic variation within and among populations of other vernal pool plants found high levels of divergence and low levels of diversity among populations (Dole and Sun 1992, Eakins 1995). However, these studies were conducted on self-compatible species using protein markers which tend to capture less diversity than DNA-level markers. Davies et al.

(1997) performed a similar study on a vernal pool fairy shrimp, *Branchinecta sandiegonensis* (Branchinectidae), and also found high levels of genetic differentiation among pools.

Objective

The present study is an expansion of previous work examining DNA-level genetic variation in *Lasthenia conjugens*. This study is only the second to examine DNA-level variation within a vernal pool plant. Previous work utilized the dominant marker method of Intersimple Sequence Repeat (ISSR) analysis. The present study was conducted using microsatellite markers. Microsatellites are neutral codominant markers, meaning levels of heterozygosity can be obtained (Schlotterer and Pemberton 1998). They provide a fine-scale means of examining genomic DNA that is greater than that of allozymes or ISSRs. The work conducted here focuses on *Lasthenia conjugens* populations found solely within Solano county and thus represents only a portion of the populations sampled for the previous work (Ramp Neale et al. 2008). Moderate to high levels of diversity were detected among the populations in Solano County (8.35% of the variation distributed among 4 populations, 91.65% of the variation distributed within populations) through previous work (Ramp, unpublished data). The specific questions addressed here are: 1) What levels of genetic diversity are present within and across the focal populations? Based on the results of previous genetic work, it is hypothesized that there will be moderate levels of genetic diversity present within and among populations. 2) How is genetic diversity structured among populations? Due to the spatially discrete nature of vernal pools and geographic barriers (developed areas) separating some populations, it is hypothesized that populations farthest away from one another will be moderately divergent, whereas populations adjacent to one another will be less divergent. 3) How would the loss of populations affect the overall genetic diversity within the species? 4) What recommendations can be made regarding future conservation and restoration projects involving *L. conjugens* based on genetic data?

Materials and Methods

Sampling procedure

Lasthenia conjugens leaf tissue was collected from vernal pools found on fifteen properties (populations) within Solano County, CA. Collections from fourteen properties were conducted by staff or associates of LSA Associates, Inc. in the springs of 2006 and 2007. A fifteenth

property (Travis Air Force Base) was sampled by Sharon Collinge (University of Colorado Boulder) in April of 2006. Leaf tissue was placed in small plastic bags with silica gel for preservation prior to DNA extraction. Tissue was received by the author in the late summer of 2006 and 2007. Sampling from most properties consisted of sampling individuals within several pools. For some properties, individuals were collected within a single area. The fifteen properties sampled for this study include: Barnfield Property, Biggs Property, Director's Guild Site, Dobles Site, Edenbridge, Jehovah's Witness Complex, McCoy Basin, Noonan Ranch (2007), North Suisun Mitigation Bank (2007), Peterson and Johnson Trust Lands West (2 parcels: one sampled in 2006, one sampled in 2007), Pullin Property (2007), Strassberger Industrial Park, Travis Air Force Base Aero Club, and Villages (not included in genetic study because only 1 individual was collected) (Table 1). For the purposes of conducting analyses by property, the two collections from the Johnson and Peterson Trust Lands were combined for property-level analyses. A total of 341 individuals are included in this study with a range of 8-50 individuals per population (Table 1). The number of individuals included varies based on the number collected and the substructure within the property. For many of the properties, more individuals were collected than could be included due to budget restrictions. When not all individuals could be included, a broad geographic sample from several pools was included.

DNA extraction

Leaf tissue was stored in silica gel until DNA was extracted at the University of Colorado, Boulder. Genomic DNA was extracted from leaf tissue following the hexadecyltrimethyl ammonium bromide CTAB protocol of Ramp et al. (2006). All genomic DNA was eluted in TE buffer and sent to Genetic Identification Systems (GIS) in California for genotyping.

Microsatellite development and genotyping

Microsatellite development was initiated from genomic DNA isolated from several *L. conjugens* individuals included in a previous study (Ramp Neale et al 2008). Initially, DNA was sent to GIS for microsatellite development. GIS was unable to develop microsatellite markers with the DNA originally sent due to problems with DNA quality. In order to acquire high quantities of DNA, several plants were grown from seed collected at Travis AFB in 2001. When at least 2 cm tall, entire plants were collected and placed in silica gel then sent to GIS for DNA

extraction. Because microsatellite development is based on genomic DNA, source location of the DNA is not important. Staff at GIS then extracted DNA using the PureGene DNA Extraction Kit® kit (Gentra Systems, Minneapolis, MN, USA) following the manufacturers instructions.

Microsatellite development proceeded according to the methods of Jones et al. (2002). Genomic DNA was partially restricted with a cocktail of seven blunt-end cutting enzymes (Rsa I, Hae III, Bsr B1, Pvu II, Stu I, Sca I, Eco RV). Fragments in the size range of 300 to 750 bp were adapted and subjected to magnetic bead capture (CPG, Inc., Lincoln Park, New Jersey), using biotinylated capture molecules. Libraries were prepared in parallel using Biotin-CA(15), Biotin-GA(15), and Biotin-ATG(8) (Integrated DNA Technologies, Inc., Coralville, IA, USA) as capture molecules in a protocol provided by the manufacturer. Captured molecules were amplified and restricted with HindIII to remove the adapters. The resulting fragments were ligated into the HindIII site of pUC19. Recombinant molecules were electroporated into E. coli DH5alpha. Recombinant clones were selected at random for sequencing, and enrichment levels were expressed as the fraction of sequences that contained a microsatellite. Sequences were obtained on an ABI 377, using ABI Prism Taq dye terminator cycle sequencing methodology.

Initial genotyping reactions amplified 11 microsatellite loci. Due to many failed reactions, all loci were re-run two to three times in order to achieve scoreable amplification. The eleven microsatellite loci were amplified (A12, B1, B101, B104, B106, B117, B119, C106, C191, C2-149, C2-189) in 10 µl reactions in the following reaction mix: $MgCl_2$, 2 mM; dNTPs (premixed), 0.2 mM each; primers, 0.3 µM each; BioTaq DNA Polymerase® (Bioline USA, Canton, MA, USA), 0.025 U/µl; template DNA, 0.2 ng/µl. PCR was conducted in a RoboCycler Gradient 96® thermocycler (Stratagene, Inc., La Jolla, CA, USA) using the following conditions: initial denaturation, 94ºC (3 min), followed by 35 cycles of 94ºC (40 sec), 55ºC (40 sec), and 72ºC (30 sec), and terminating with a final extension at 72ºC for 4 min. PCR products were labeled using one of the conventional sequencing dyes NED, HEX or FAM (Applied Biosystems, Inc.). Amplification products were separated an ABI 3730 DNA sequencer and sized using Genemapper software (Applied Biosystems, Inc., Foster City, CA USA). A total of 7 loci produced enough product to be included in this analysis. For some loci, binning of peaks into alleles was necessary to account for stutter in the results.

Population genetic diversity

As mentioned above, the two collections from the Peterson and Johnson Trust lands were combined for analyses and the sample from the Villages property was not included resulting in all property-level analyses being conducted with 13 properties. Data analysis to assess genetic diversity estimated the following parameters: the average sample size (n), proportion of polymorphic loci (P), average number of alleles per locus (A) and per polymorphic locus (Ap), expected proportion of heterozygotes (He), observed proportion of heterozygotes (Ho), and the inbreeding coefficient (f) utilizing the Genetic Data Analysis Program (Lewis and Zaykin 1999). Descriptive statistics were calculated for the 13 properties sampled as well as for all vernal pool populations sampled when sub-population division is included. Properties with sub-structure include: Barnfield Property (4 sub-populations), Director's Guild Site (2 sub-populations), Edenbridge (4 sub-populations), North Suisun Mitigation Bank (4 sub-populations), Noonan Ranch (3 sub-populations), Peterson and Johnson Trust Lands West (7 sub-populations total, 3 from the West parcel and 4 from the South parcel), Strassberger Industrial Park (5 sub-populations), and Travis Air Force Base Aero Club (4 sub-populations). When considering all pools as sub-populations, a total of 42 populations were included in analyses. The descriptive statistics were also calculated for each locus.

Population genetic structure

To examine population genetic structure, genetic differentiation among all sampled populations was analyzed using F_{ST}, as well as the Bayesian approach of Holsinger *et al.* (2002) as implemented in Hickory. This approach calculates direct estimates of population divergence (θ^B, an analogue of F_{ST}).

Hierarchical genetic structure was examined through an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) as implemented in Arlequin 3.1.1 (Schneider et al. 2000). Two analyses were conducted: all properties independently, and populations (sub-populations). An analysis of molecular variance examines all the diversity present within the sampled individuals and partitions it among the pre-assigned groups.

Geographic structure of the genetic data was examined through a Mantel test, a principal coordinate analysis (PCoA), and a cluster analysis. These analyses were performed for all of the 42 sub-populations sampled because exact geographic information on pool location was

available. Using all 42 sampled pools for these analyses provides a more detailed examination of genetic structure than looking at properties alone. The Mantel test examining isolation by distance (Mantel 1967) was conducted using the program PASSAGE (Rosenberg 2001). The Mantel test examined the relationship between geographic distance and Nei's genetic distance as calculated in POPGENE v1.31 (Yeh et al. 1999). The geographic distance matrix was constructed by pulling latitude and longitude locations for each pool sampled from the ArcGIS map provided by LSA. The program Geographic Distance Matrix Generator (Peter J. Ersts available online at http://biodiversityinformatics.amnh.org/open_source/gdmg/) was used to generate the geographic distance matrix.

The principal coordinates analysis (PCoA) was performed on all 42 pools using Jaccard's similarity index in the program PAST (Hammer et al. 2001). A neighbor-joining cluster analysis based on Nei's genetic distance (1978) was performed in GDA (Lewis and Zaykin 1999) and visualized in TREEVIEW (Page 1996) for both the 13 properties and all 42 populations.

Results

Genetic Diversity

The descriptive statistics calculated for each property include the average sample size (n), proportion of loci polymorphic (P), average number of alleles per locus (A) and per polymorphic locus (Ap), expected proportion of heterozygotes (He), observed proportion of heterozygotes (Ho), and inbreeding coefficient (f) (Table 2). These same statistics are calculated for all 42 populations sampled and for each locus (Table 3 and Table 4, respectively). The seven microsatellite loci employed in this study were polymorphic across all 13 properties. A high number of alleles were detected across all loci (11 - 47) with an average of 25.86 alleles per locus. Inbreeding was detected in all properties, sub-populations, and in all loci except locus C2-154.

Genetic Structure

Population differentiation, or among population geographic structure was determined using F_{ST} , the Bayesian approach of Holsinger *et al.* (2002) (θ^B), and an AMOVA. In examining the 13 properties independently, F_{ST} was calculated to be 0.02215 ($p < 0.0001$) and the average Bayesian estimate of θ^B was 0.0111. The examination of among population differentiation using

the AMOVA for all properties, showed the majority of the variation to be partitioned within populations (97.78%) with the remaining variation distributed among properties (2.22%).

For the analyses including all sampled pools (n = 42), F_{ST} was calculated to be 0.04003 ($p < 0.0001$) and the Bayesian estimate of θ^B was 0.007. In the AMOVA examining genetic variation with the inclusion of sub-populations, the among property variation was 1.14% of the total variation with 2.86% of the variation found among sub-populations within properties and the remaining 96% of the variation found within populations.

The examination of the relationship between geographic and Nei's genetic distances using a Mantel test was not significant ($z = 0.05911$, $p > 0.05$), indicating no isolation by distance among the populations (42 populations). The principal coordinates analysis (PCoA) was performed on all 42 populations using Jaccard's similarity index in the program PAST (Hammer et al. 2001). The PCoA shows the populations to be scattered throughout the coordinate space (Figure 1). The ellipses enclose 95% of the samples for each of the 13 properties. There is no strong clustering of populations by property or geographic proximity (labels not shown for ease of reading the figure). The first axis explains 25.106% of the variation, the second axis explains 19.242%, with the third axis explaining 13.699% of the overall variation. A neighbor-joining cluster analysis based on Nei's genetic distance (1978) was performed in GDA (Lewis and Zaykin 1999) and visualized in TREEVIEW (Page 1996) for both the 13 properties and all 42 populations. There is some geographic structuring to the genetic variation when examined at the property level (Figure 2) in that some populations that are in close geographic proximity to one another group together in the analysis (McCoy, Strassberger and Edenbridge). When conducting the analysis on all 42 populations, no complete set of pools (sub-populations) group together by property (Figure 3). Despite some pools grouping together by property (North Suisun Mitigation Bank pools D, B, C; Barnfield C, A, D), there is very little geographic structure to the data (Figure 3).

Discussion

The present study is the second to examine DNA-level within- and among-population levels of genetic variation in natural populations of a vernal pool plant. It is the first study to use microsatellite markers to examine vernal pool plant diversity. As hypothesized for this study, moderate to high levels of genetic diversity were detected within the sampled populations

through the use of microsatellite markers. Although genetic diversity estimates measured here are not directly comparable to previous estimates due to the use of a different marker system, high levels of diversity were detected with a high level of alleles per locus and high levels of polymorphism in alleles across all properties. Low to moderate levels of inbreeding (f) were detected both within properties and within pools. The detected inbreeding suggests a limitation on gene flow among pools and among properties. Although *L. conjugens* is a self-incompatible annual, the discrete nature of vernal pool habitat lends itself to inbreeding among individuals within a single pool, but see below for further discussion.

The hypothesis of moderate genetic structure for the sampled properties was not upheld in that there was less geographic structure to the data than hypothesized. The calculations of F_{ST}, θ^B, and the AMOVA all demonstrate weak population structure among either properties or all populations.

The Mantel test examining isolation by distance was not significant indicating there is not a significant relationship between geographic and genetic distances. However, given the close proximity of the sampled populations (range of approximately 150 meters – 15 kilometers) and the high levels of genetic diversity found within the populations, it is not surprising that there is no strong geographic structure to the data. The PCoA and the neighbor-joining phenograms based on Nei's genetic distance also support very little geographic structure to the data.

Broadly, these results indicate that there are high levels of genetic diversity within *L. conjugens* populations within Solano County, CA. The variation is widely distributed both among pools within a single property and across properties within the county. As outlined in Ramp Neale et al. (2008) there are two possible hypotheses to explain the levels and patterns of diversity seen in this species. The first hypothesis is that historical geologic factors have shaped the floral patterns of the Central Valley with a profound influence on vernal pool habitat distribution. It is presumed that vernal pool habitat within the Central Valley would have only recently been broken-up due to agricultural development (Raven and Axelrod 1978, Bell and Patterson 2000). Given the life-history traits of *L. conjugens*, self-incompatible annual with high reproductive output and large population sizes, it is likely that not enough generations have passed in order for genetic drift to have had a large impact on genetic structure since geographic isolation occurred (Ramp Neale et al. 2008). The second hypothesis explaining genetic levels and patterns in *L. conjugens* is related to contemporary gene flow among existing populations

(Ramp Neale et al. 2008). It is possible that gene flow is occurring among vernal pool complexes found within the Central Valley and within Solano County in particular. However, given the detection of inbreeding within pools, the short flight distances of pollinators (Leong et al. 1995, Thorp and Leong 1998, Ramp 2004), gravity dispersed seeds, and little opportunity for hydrologic connections among vernal pool complexes, current gene flow is unlikely.

Due to the high number of alleles (181) detected in this study, and the large number of individuals included here (341), it was not possible to conduct analyses examining the probability of loss of rare alleles as conducted in Ramp Neale et al. (2008). However, the issue of how the loss of populations would affect the overall genetic diversity within the species can still be addressed. Given the intense sampling conducted for this study, a robust measure of genetic diversity levels and patterns for Solano County, CA is presented. The high level of diversity detected here indicates that either the recent (approximately one hundred and fifty years) isolation of populations has not led to the development of strong genetic structure among populations, or that gene flow is still occurring among populations within the region. The lack of geographic structure to the data indicates that no property or pool is genetically distinct from other properties or pools. The AMOVA examining hierarchical genetic structure of all pools sampled indicates that there is greater diversity among pools within a single property than among properties. This indicates that the genetic makeup of each pool is unique in that each pool is composed of a unique allele composition across the sampled individuals. Even though there may not be unique alleles in each population (not assessed here), the composition (number and distribution) of alleles sampled in the populations is unique for each population. This implies that property boundaries are biologically irrelevant and that the loss of individual pools will have a negative impact on genetic composition of the species. However, the results also indicate that no pool or set of pools is more distinct than any others. The loss of any *Lasthenia conjugens* populations in Solano County will likely reduce the overall level of diversity within the species, but is unlikely to have a large negative impact on the evolutionary potential of the species given the high diversity and low population differentiation detected here.

Given the results of the present study, conservation efforts should aim to conserve as many individual pools as possible. However, given the development pressures on the land supporting vernal pool habitat in Solano County, it is understood that this is not always possible. The results indicate that although there is widespread diversity and minimal structure, no pool or

group of pools is more diverse, or more distinct than any others. Any restoration or reintroduction efforts should aim to sample the existing populations widely so as to capture the diversity found in all populations. Seed collection efforts should not be focused on single populations but rather on all populations for which collection is possible. Restoration efforts could then utilize seed from either a single or multiple sources when necessary. It is recommended that pool initiation is done in a conservative manner with seeds from a single source pool or small group of closely located pools. When necessary, multiple pools could be used as source material for new populations. The geographic proximity of restored sites to seed sources should be considered in restoration efforts, but does not need to dictate where restoration occurs given the lack of geographic structure to the genetic data. In all cases, careful documentation of both source and restored pools should be kept so that genetic diversity and structure could be tracked over time.

References

Bell, C. D., and R. W. Patterson. 2000. Molecular phylogeny and biogeography of *Linanthus* (Polemoniaceae). American Journal of Botany **87**:1857-1870.

Dole, J. A., and M. Sun. 1992. Field and genetic survey of the endangered Butte County Meadowfoam - *Limnanthes floccosa* subsp. *californica* (Limnanthaceae). Conservation Biology **6**:549-558.

Eakins, D. G. 1995. Intraspecific genetic diversity in *Blennosperma bakeri* Heiser (Asteraceae). Master of Arts. Sonoma State University, Sonoma.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131**:479-491.

Falk, D. A., and K. E. Holsinger. 1991. Genetics and Conservation of Rare Plants. Oxford University Press, New York.

Fleishman, E., A. E. Launer, K. R. Switky, U. Yandell, J. Heywood, and D. D. Murphy. 2001. Rules and exceptions in conservation genetics: genetic assessment of the endangered plant *Cordylanthus palmatus* and its implications for management planning. Biological Conservation **98**:45-53.

Frankham, R. 2003. Genetics and conservation biology. Comptes Rendues Biologies **326**:S22-S29.

Hammer, O., D. A. T. Harper, and P. D. Ryan. 2001. PAST: Palaenotological Statistics Software Package for education and data analysis. Palaeontologia Electronica **4**:9.

Hamrick, J. L., M. J. W. Godt, D. A. Murawski, and M. D. Loveless. 1991. Correlations between species traits and allozyme diversity: implications for conservation biology. Pages 75-86 *in* D. A. Falk and K. E. Holsinger, editors. Genetics and Conservation of Rare Plants. Oxford University Press, New York.

Hedrick, P. W. 2001. Conservation genetics: where are we now? Trends in Ecology and Evolution **16**:629-636.
Holsinger, K. E., P. O. Lewis, and D. Dey. 2002. A Bayesian approach to inferring population structure from dominant markers. Molecular Ecology **11**:1157-1164.
Jones, K. C., K. F. Levine, and J. D. Banks. 2002. Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). Molecular Ecology Notes **2**:425-427.
Leong, J. M., R. P. Randolph, and R. W. Thorp. 1995. Observations of the foraging patterns of *Andrena* (*Diandrena*) *Blennospermatis* Thorp (Hymenoptera: Andrenidae). Pan-Pacific Entomologist **71**:68-71.
Lewis, P. O., and D. Zaykin. 1999. Genetic Data Analysis: computer program for the analysis of allelic data. .
Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research **27**:209-220.
Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**:583-590.
Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences **12**:357-358.
Ramp, J. M. 2004. Restoration genetics and pollination of the endangered vernal pool endemic *Lasthenia conjugens* (Asteraceae). Dissertation. University of Colorado, Boulder.
Ramp, J. M., S. K. Collinge, and T. A. Ranker. 2006. Restoration genetics of the vernal pool endemic *Lasthenia conjugens* (Asteraceae). Conservation Genetics **7**:631-649.
Ramp Neale, J. M., T. A. Ranker, and S. K. Collinge. 2008. Conservation of rare species with island-like distributions: A case study of *Lasthenia conjugens*
(Asteraceae) using population genetic structure and the distribution of rare markers. Plant Species Biology **23**:97-110.
Raven, P. H., and D. I. Axelrod. 1978. Origin and relationships of the California flora. University of California Press, Berkeley, CA.
Rosenberg, M. S. 2001. PASSAGE. Pattern analysis, spatial statistics, and geographic exegesis. Department of Biology, Arizona State University, Tempe, AZ.
Schlotterer, C., and J. Pemberton. 1998. The use of microsatellites for genetic analysis of natural populations - a critical review. Pages 71-86 *in* R. DeSalle and B. Schierwater, editors. Molecular approaches to ecology and evolution. Birkhauser Verlag Basel.
Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin ver. 2.000: A software program for population genetic data analysis. Genetics and Biometry Laboratory, Geneva.
Thorp, R. W., and J. M. Leong. 1998. Specialist bee pollinators of showy vernal pool flowers. Pages 169-179 *in* C. W. Witham, E. T. Bauder, D. Belk, W. R. J. Ferren, and R. Ornduff, editors. Ecology, Conservation, and Management of Vernal Pool Ecosystems - Proceedings from a 1996 conference. California Native Plant Society, Sacramento, CA.
Yeh, F. C., R.-C. Yang, T. Boyle, Z. Ye, and J. X. Mao. 1999. POPGENE Version 1.31. Edmonton, Alberta.

Table 1: Populations sampled for genetic analysis within Solano County, CA.

Property	Year of collection	Number of individuals included in genetic analysis	Number of populations (pools) included in analyses	Included in previous genetic analyses?
Barnfield Property	2006	40	4	
Biggs Property	2006	10	1	
Director's Guild Site	2006	30	2	Yes (Potrero Hills)
Dobles Site	2006	50	5	
Edenbridge	2006	35	4	
Jehovah's Witness Complex	2006	8	1	
McCoy Basin	2006	24	1	Yes
Noonan Ranch	2007	16	3	
North Suisun Mitigation Bank	2007	24	4	
Peterson and Johnson Trust Lands West	2006, 2007	46	7	
Pullin Property	2007	10	1	
Strassberger Industrial Park	2006	24	5	
Travis Air Force Base Aero Club	2006	24	4	Yes
Villages	2006	0	0	
Total		341	42	

Table 2: Descriptive statistics for the 13 sample properties. The descriptive statistics calculated for each property include the average sample size (n), proportion of loci polymorphic (P), average number of alleles per locus (A) and polymorphic locus (Ap), expected proportion of heterozygotes (He), observed proportion of heterozygotes (Ho), and inbreeding coefficient (f).

Population	n	P	A	Ap	He	Ho	f
Barnfield	35.28571	1	13.28571	13.28571	0.798954	0.579391	0.27746
Dobles	41.57143	1	14.28571	14.28571	0.843128	0.589288	0.303197
Jehovah witness	7.285714	1	7.571429	7.571429	0.830041	0.618367	0.268366
Eddy Biggs	8.857143	1	6.285714	6.285714	0.745587	0.632143	0.160688
Edenbridge	33	1	15.57143	15.57143	0.840057	0.650413	0.22835
Peterson Johnson Trust W and S	41.42857	1	15.42857	15.42857	0.841075	0.60325	0.284945
McCoy	20.71429	1	11.85714	11.85714	0.817262	0.536017	0.348401
Noonan	14.42857	1	8.571429	8.571429	0.800359	0.595833	0.259597
Directors Guild	27.28571	1	15	15	0.880184	0.643242	0.272243
Pullin	9.285714	1	7.428571	7.428571	0.831436	0.644048	0.234082
Strassberger	21.71429	1	11.71429	11.71429	0.793296	0.605556	0.240212
Travis AFB	22.14286	1	13.71429	13.71429	0.873354	0.731027	0.165772
North Suisun Mitigation Bank	20.42857	1	11.28571	11.28571	0.845151	0.582727	0.314486
Mean	23.34066	1	11.69231	11.69231	0.826145	0.616254	0.259425

Table 3: Descriptive statistics for all 42 populations. The descriptive statistics calculated for each property include the average sample size (n), proportion of loci polymorphic (P), average number of alleles per locus (A) and polymorphic locus (Ap), expected proportion of heterozygotes (He), observed proportion of heterozygotes (Ho), and inbreeding coefficient (f).

Population	n	P	A	Ap	He	Ho	f
Barnfield A	7.714286	1	6.285714	6.285714	0.79338	0.520408	0.360669
Barnfield B	9.428571	1	7.142857	7.142857	0.752226	0.543537	0.288567
Barnfield C	9.285714	1	7.571429	7.571429	0.80503	0.605669	0.257455
Barnfield D	8.857143	1	6.571429	6.571429	0.777286	0.624603	0.200062
Dobles D	8.285714	1	7.857143	7.857143	0.865549	0.619728	0.29225
Dobles J	8.142857	1	6.571429	6.571429	0.768344	0.568254	0.273099
Dobles P	7.857143	1	7	7	0.827666	0.585034	0.304626
Dobles Q	7.714286	1	5.142857	5.142857	0.742169	0.440816	0.401264
Dobles R	9.571429	1	6.571429	6.571429	0.835874	0.638889	0.244489
Jehovah Witness	7.285714	1	7.571429	7.571429	0.830041	0.618367	0.268366
Eddy Biggs	8.857143	1	6.285714	6.285714	0.745587	0.632143	0.160688
Edenbridge G	4.714286	1	5	5	0.774603	0.571429	0.277108
Edenridge I	9.714286	1	8.285714	8.285714	0.826566	0.646429	0.2274
Edenbridge J	9.428571	1	7.571429	7.571429	0.807957	0.628571	0.231492
Edenbridge O	9.142857	1	7.571429	7.571429	0.840422	0.701587	0.173276
Peterson Johnson Trust W and S NWA	8.571429	1	6.285714	6.285714	0.789864	0.620408	0.221586
Peterson Johnson Trust W and S NWD	9.714286	1	6	6	0.74354	0.58254	0.225443
Peterson Johnson Trust W and S NWE	8.714286	1	7	7	0.84384	0.689569	0.190756
Peterson Johnson Trust W and S S1	4.428571	1	4.142857	4.142857	0.714626	0.485714	0.339994
Peterson Johnson Trust W and S S2	2.428571	0.857143	2.428571	2.666667	0.538095	0.5	0.107143
Peterson Johnson Trust W and S S5	5	1	4.857143	4.857143	0.755556	0.571429	0.266055
Peterson Johnson Trust W and S S6	2.571429	0.857143	3.428571	3.833333	0.719048	0.595238	0.204918
McCoy	20.71429	1	11.85714	11.85714	0.817262	0.536017	0.348401
Noonan 1	8.857143	1	6.714286	6.714286	0.80674	0.592857	0.267391
Noonan 31	4.571429	1	3.857143	3.857143	0.688095	0.57619	0.175511
Noonan 33	1	0.571429	1.571429	2	0.571429	0.571429	0
Directors Guild North	14.14286	1	12	12	0.880946	0.655861	0.262055
Directors Guild South	13.14286	1	10.28571	10.28571	0.867297	0.635003	0.27401
Pullin	9.285714	1	7.428571	7.428571	0.831436	0.644048	0.234082
Strassberger A	4.285714	0.857143	3.714286	4.166667	0.651814	0.4	0.401844
Strassberger C	4.714286	1	4.285714	4.285714	0.664739	0.578571	0.139907
Strassberger F	4.571429	1	4.428571	4.428571	0.722222	0.585714	0.205584
Strassberger G	4.571429	1	5.714286	5.714286	0.812018	0.695238	0.162526
Strassberger I	3.571429	1	4	4	0.798299	0.797619	0.001736
Travis AFB N1	5.285714	1	5.428571	5.428571	0.820594	0.557143	0.339231
Travis AFB N27	5.571429	1	6.428571	6.428571	0.850639	0.769048	0.104846
Travis AFB N36	5.571429	1	6.142857	6.142857	0.824026	0.728571	0.124873
Travis AFB N76	5.714286	1	6.428571	6.428571	0.856834	0.833333	0.028011
North Suisun Mitigation Bank A	7.428571	1	6.857143	6.857143	0.846429	0.636905	0.264441
North Suisun Mitigation Bank B	6.142857	1	5.571429	5.571429	0.783716	0.503401	0.373227
North Suisun Mitigation Bank C	4.166667	1	4.833333	4.833333	0.823413	0.663889	0.214193
North Suisun Mitigation Bank D	3.285714	0.857143	4.142857	4.666667	0.676871	0.571429	0.174853
Mean	7.238662	0.97619	6.162698	6.211451	0.778383	0.607682	0.23786

Table 4: Descriptive statistics for all 42 populations. The descriptive statistics calculated for each property include the average sample size (n), proportion of loci polymorphic (P), average number of alleles per locus (A) and polymorphic locus (Ap), expected proportion of heterozygotes (He), observed proportion of heterozygotes (Ho), and inbreeding coefficient (f).

Locus	n	P	A	Ap	He	Ho	f
B1	329	1	24	24	0.881963	0.817629	0.073047
B101	268	1	29	29	0.900558	0.455224	0.494977
B103	311	1	33	33	0.953146	0.549839	0.423526
B119	239	1	21	21	0.931484	0.435146	0.533369
C191	338	1	16	16	0.670642	0.532544	0.206161
C2-149	309	1	47	47	0.889779	0.673139	0.243775
C2-157	330	1	11	11	0.796188	0.863636	-0.08485
All	303.4286	1	25.85714	25.85714	0.860537	0.618166	0.281948

Figure 1: Principle Coordinate Analysis based on Jaccard similarity index. Forty two sub-populations represented. Ellipses represent 95% inclusion of samples from each of the thirteen properties. Labels not included for ease of reading.

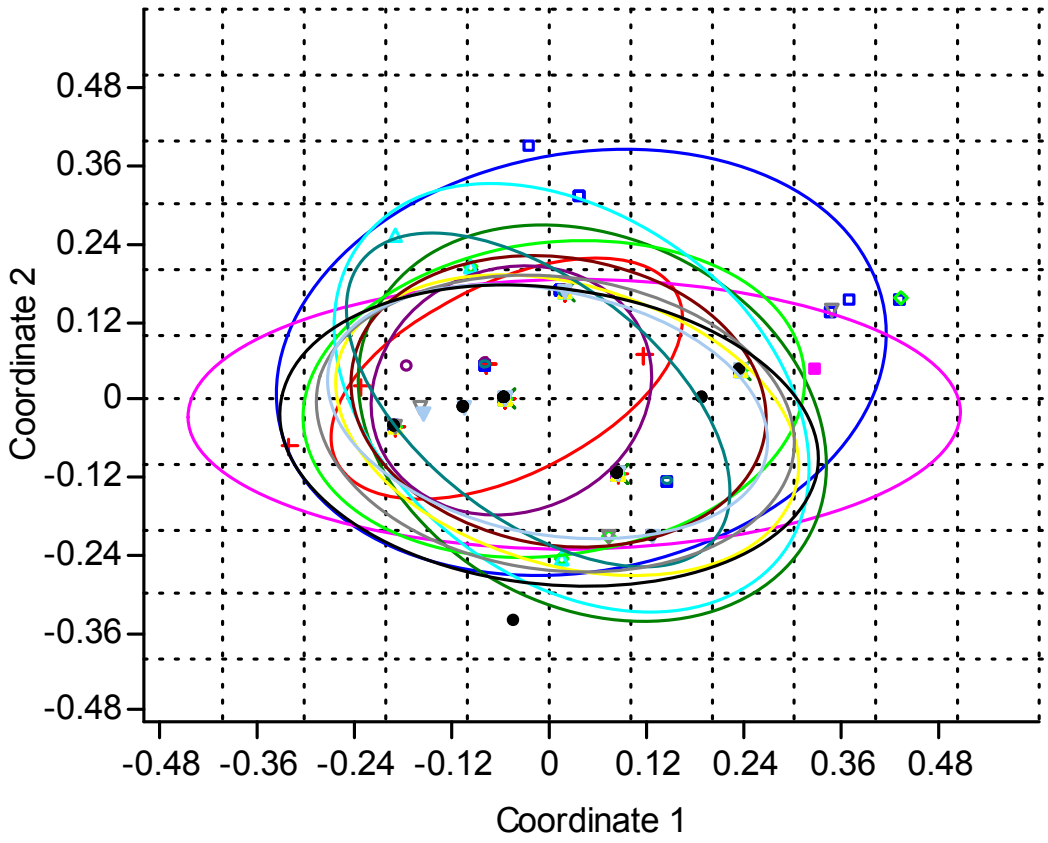

Figure 2: Neighbor-joining tree based on Nei's (1978) genetic distance for the 13 properties.

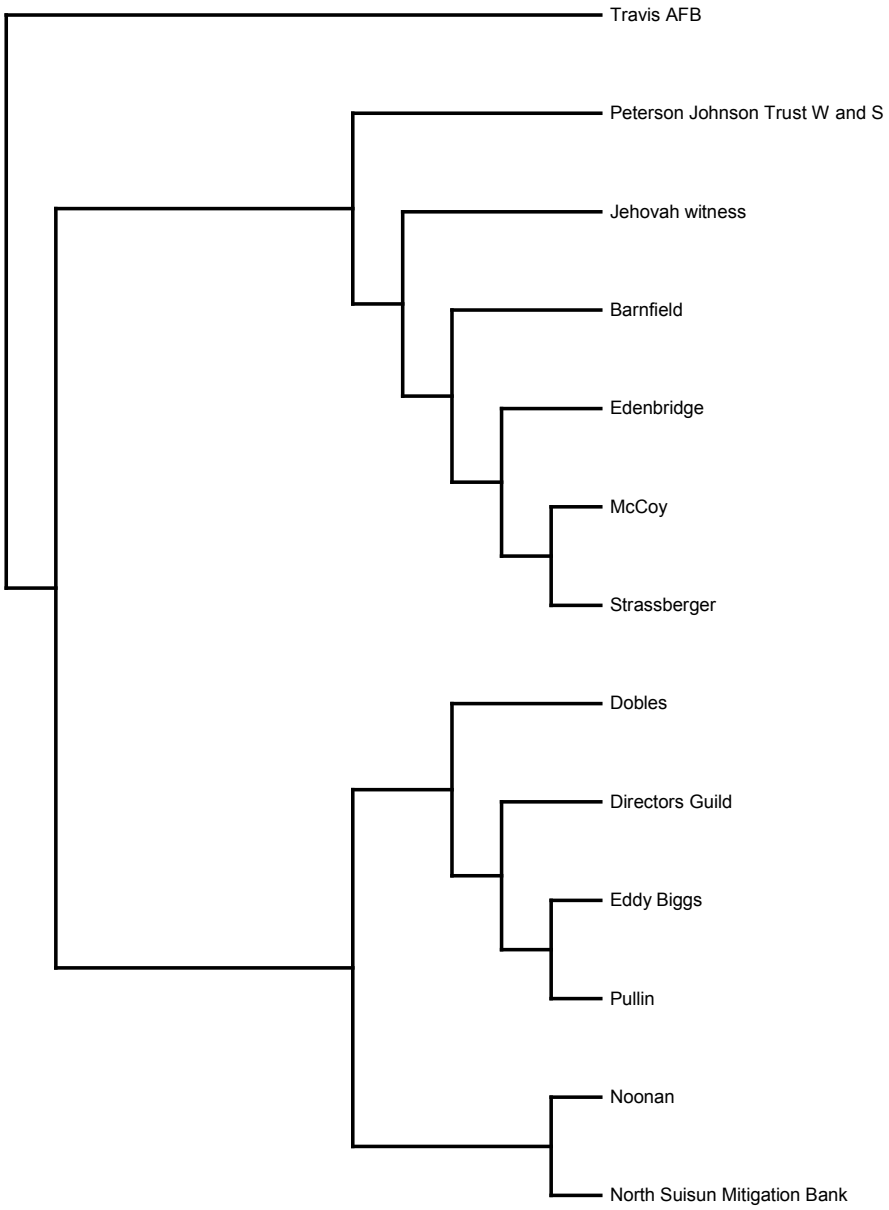

Figure 3: Neighbor-joining tree based on Nei's (1978) genetic distance for the 42 populations.