CHAPTER 4.1.

ROLE OF POPULATION GENETICS IN THE STERILE INSECT TECHNIQUE

E. S. KRAFSUR¹ AND J. O. OUMA²

¹*Department of Entomology, Iowa State University, Ames, Iowa 50011, USA*

Email: ekrafsur@iastate.edu

²*Africa Technical Research Centre, Vector Health International, P.O. Box 15500,
Arusha, Tanzania*

TABLE OF CONTENTS

1.	INTRODUCTION.....	530
2.	WHY POPULATION GENETICS?.....	531
3.	CHIEF KINDS OF GENETIC VARIATION	532
4.	RELATIVE ADVANTAGES AND DISADVANTAGES OF PRINCIPAL METHODS	533
4.1.	Obsolete Methods	533
4.2.	Mitochondrial Variations.....	534
4.3.	Chromosomal Polymorphisms	534
4.4.	Microsatellites and Single Nucleotide Polymorphisms	535
4.5.	DNA Sequencing	536
5.	BASIC PRINCIPLES OF POPULATION GENETICS	536
6.	PROCESSING AND INTERPRETING POPULATION GENETIC DATA	537
7.	POPULATION GENETICS AND CONSTRUCTION OF RELEASE STRAINS	540
8.	APPLICATIONS AND EXAMPLES	543
9.	REFERENCES.....	544

SUMMARY

The detection and analysis of genetic variation in natural and laboratory populations of pest insects are reviewed. The application of population genetic methods and theory can help to plan and evaluate the implementation of area-wide integrated pest management (AW-IPM) programmes that use the sterile insect technique (SIT). Population genetic studies can play an important role in estimating dispersal rates and thus the degree of isolation or gene flow among target populations, determining if sibling species exist, establishing the origin of outbreaks or reintroductions, and supporting the quality control of mass-reared colonies. The target's population history may be examined in terms of "bottlenecks", range

Pages 529–548

V. A. Dyck, J. Hendrichs and A. S. Robinson (eds.), Sterile Insect Technique. Principles and Practice in Area-Wide Integrated Pest Management. Second Edition.

© 2021 IAEA. CRC Press, Boca Raton, Florida, USA.

fragmentations, and expansions. Genetic methods can be helpful in distinguishing wild insects from released sterile or semi-sterile ones, and in ascertaining, together with mating compatibility studies, the compatibility of mass-reared colonies with target wild insects.

1. INTRODUCTION

Population genetics, in the broad sense, is the study of gene frequencies in and among subdivided populations. It encompasses estimations of variation in terms of allelic and genotypic frequencies. Given appropriate data, the application of population genetic theory allows estimates of rates of gene flow within and among populations, estimates of effective, i.e. reproductive, population sizes, tests for past “bottlenecks” and rapid expansions of population size, and pairwise genetic distances. Since all evolutionary change is accompanied by change in gene frequencies, population and evolutionary genetics overlap.

Methods of population genetics can be used to support the application of the sterile insect technique (SIT) for the integrated control of target pest populations in nature. Briefly, the degree of genetic isolation of target populations from each other, and from untargeted populations, can be estimated. The existence of sibling species may be investigated, in combination with behavioural, cytogenetic and other approaches, by careful sampling and genetic analysis of pest populations. Immigration rates from unchallenged populations may be estimated, in terms of reproducing females per generation. The origin of pest outbreaks in treated areas may thereby be investigated. The genetic relatedness of mass-reared strains and target populations may be estimated in terms of allelic composition and gene frequencies. The question of laboratory “degradation” of strains destined for release can be examined objectively in terms of inbreeding and genetic distances from founding and target populations.

Given what may be done by using modern techniques and adequate sampling of natural populations, it should nevertheless be noted that the New World screwworm *Cochliomyia hominivorax* (Coquerel), tsetse fly *Glossina austeni* Newstead and Mediterranean fruit fly *Ceratitis capitata* (Wiedemann) area-wide integrated pest management (AW-IPM) programmes were executed successfully without any substantial knowledge of the population genetics (section 8) (Enkerlin, this volume; Feldmann et al., this volume; Vargas-Terán et al., this volume).

Here we outline the kinds of genetic variation that may be examined efficiently, and briefly review tests of hypothesis and analytical methods that may be applied to genetic data. In addition, the causes of genetic differentiation are discussed.

Population genetics is built on well-developed and sophisticated mathematical and statistical theory (but not reviewed here). Crow and Kimura (1970) and Wright (1969a, b; 1978a, b) covered the theoretical basis of population genetics. The analysis of population genetic data was examined by Nei (1987), Weir (1996), Rousset and Raymond (1997), and Nei and Kumar (2000). Luikart and England (1999) reviewed the methods of data analysis, and tabulated the necessary software. Wright (1969a, b; 1978a, b), Avise (2004), Gillespie (2004), Hartl and Clark (2007), Hedrick (2011), and Nielsen and Slatkin (2013) provided general treatments of population and evolutionary genetics. We recommend Hartl and Clark as the

“friendliest” treatment. Roderick (1996) and Black et al. (2001) have written particularly useful reviews of insect population genetics.

2. WHY POPULATION GENETICS?

- Sibling or cryptic species are reproductively isolated forms that are morphologically undifferentiated. If pre-mating isolating mechanisms exist, the application of sterile males of one member of a species complex may be totally ineffective against other members. Sibling species are common, e.g. among tsetse flies *Glossina morsitans* Westwood, *G. fusca* Walker, *G. palpalis* Robineau-Desvoidy and *G. fuscipes* Newstead (Gooding and Krafur 2005), the *Drosophila willistoni* Sturtevant species complex, the *Anopheles maculipennis* Meigen complex (Bates 1949), the *Anopheles gambiae* Giles complex (Coluzzi et al. 2002; Dorn et al. 2011), the *Aedes scutellaris* (Walker) complex (Kambhampati et al. 1992), the *Simulium damnosum* Theobald complex (Vajime and Gregory 1990), the sand fly *Lutzomyia longipalpis* (Lutz and Neiva) (Lanzaro et al. 1993) and the ladybird beetle *Coleomegilla maculata* (De Geer) complex (Krafur and Obrycki 2000). Most complexes uncovered to date are medically important Diptera or genetically interesting *Drosophila* species, but this is probably sampling bias. Complexes important to agriculture include the tephritid flies *Anastrepha fraterculus* (Wiedemann), where the existence of at least seven morphotypes has been confirmed recently, while for *Bactrocera dorsalis* (Hendel) the putative complex was shown to represent intra-specific variation, resulting in the synonymization of four pest species (Hendrichs et al. 2015). Colonies of medically important Diptera are routinely established to study insecticide resistance, vector-pathogen transmission, etc., and intercrosses among strains may produce sterile progeny. Examination of polytene chromosomes, well developed in many Diptera, may reveal reproductively isolated forms, as demonstrated in some *Drosophila* species (Dobzhansky 1970). Many insect pest species are not easily reared or studied genetically, particularly herbivorous beetles and moths, hence the question of cryptic speciation may not arise and cannot be investigated easily.
- Natural populations tend to be discontinuously distributed. Gene flow varies among them as per many factors, including the historical, geographical, and environmental. There may be local selection regimes that cause one or more populations to differ biologically in ways that could make the SIT less effective due to mating barriers. Such barriers might be ecological, temporal, or behavioural, and their strengths, in principle, could vary greatly. It should be noted, however, that no such mating barriers were detected in the New World screwworm (Krafur 1985, 1994) and Mediterranean fruit fly programmes (Cayol et al. 2002).
- Estimating rates of gene flow among populations can provide an estimate of the rate of invasion or reinvasion of eradicated areas from unchallenged populations. The rates can be described in terms of the average numbers of reproducing flies per generation. Moreover, the likely source(s) of pests in an outbreak can be examined and, in many instances, specified. For easily transported pests such as

the Mediterranean fruit fly, the geographic origins of new infestations can, in principle, be specified. Determining the source of new infestations is an important undertaking to determine and close pathways, up to establishing embargos on produce originating from areas suspected of having shipped infested commodities (Aguirre-Ramirez et al. 2017).

- Construction of laboratory strains for sterilization and release can be improved. If adequate numbers of insects are to be sterilized and released, it is necessary to establish thriving colonies. How much of the existing natural genetic diversity should be incorporated into a release strain? The geographic range of many pest species is very large, and locally adapted populations may exist. How then should the genetic make-up of release strains be formulated?
- An analysis of genetic variation may also reveal genetic markers in laboratory colonies that can be used to differentiate released insects or the offspring of semi-sterile insects from their wild counterparts during the monitoring of programmes that release sterile insects.

3. CHIEF KINDS OF GENETIC VARIATION

Chromosome morphologies and structural rearrangements may vary within species and provide markers that distinguish populations (Robinson, this volume). Examples include paracentric inversions, easily detected in stained squashes of polytene chromosomes found in *Drosophila*, *Anopheles* mosquitoes, and the supernumerary chromosomes of some *Glossina morsitans* group tsetse flies (Gooding and Krafsur 2005). These kinds of variation rarely provide the high degree of resolution afforded by molecular variation in DNA or its transcription and translation products such as allozymes. Moreover, chromosome rearrangements are not selectively neutral as has been demonstrated in *Drosophila* species and *Anopheles gambiae* (Coluzzi et al. 2002). They may nevertheless provide important information about reproductive isolation and the operation of natural selection.

Allozymes are allelic forms of an enzyme. Isozymes are enzymes that act on a common substrate. Alleles are codominant. They are readily demonstrated by electrophoresis of tissue homogenates on starch, cellulose acetate, or polyacrylamide followed by histochemical staining that involves a dye or UV-fluorescence (Richardson et al. 1986; Murphey et al. 1996). Allozyme banding patterns correspond to genotypes, i.e. one band for a homozygote, two bands for a heterozygote at a monomeric locus, and three bands for a heterozygote at a dimeric locus. Most allozyme loci occur on the nuclear genome, but some loci, e.g. nicotinamide adenine dinucleotide phosphate (NADP)-dependent malate dehydrogenase, are mitochondrial. Allozyme analysis requires that tissue samples be fresh or frozen to preserve enzyme activities. This restriction does not apply to analysis of DNA.

Mitochondrial DNA (mtDNA) is single-copy, hence haploid, self-replicating, and is maternally transmitted as a circular molecule of about 16 kilobases. It encodes 22 transfer RNA genes and 13 proteins, including the cytochrome oxidases, NADH, RNAase, and ribosomal RNAs. There are many highly conserved regions from which oligonucleotide primers have been designed, e.g. Simon et al. (1994), which

are commercially available in kits. These primers allow the amplification of regions of the mitochondrial chromosome by the polymerase chain reaction (PCR). Mitochondrial variation can be assessed by several methods. Nucleotide sequencing of amplified regions can be followed by the comparison of sequences among individuals. Sequencing is recommended now that its costs have greatly decreased.

Microsatellites are simple sequence repeats (SSRs) of nucleotide motifs that are repeated n times: $(CA)_n$, $(GA)_n$, $(TA)_n$, $(CAG)_n$, etc. (where C is cytosine, A is adenine, G is guanine, T is thymine) that may occur throughout the nuclear genome. Each locus may vary among individuals in the number of repeats. Alleles are codominant. Conserved flanking sequences allow the design of oligonucleotide primers with which to amplify via PCR the microsatellite region. Comparatively sophisticated procedures are necessary to find microsatellites in the genome, and to develop oligonucleotide primers to amplify them. A substantial number of protocols are available to do so, e.g. Abdelkrim et al. (2009), Malausa et al. (2011). Some protocols involve enrichment of DNA for repeat sequences to maximize the yield of desired repeats (Hamilton et al. 1999). Alternatively, one may choose to hire a commercial organization to isolate and characterize microsatellite loci, and for some laboratories this may be the least expensive and most efficient way to proceed.

Single Nucleotide Polymorphisms (SNPs) are sites in the genome where DNA sequences differ by a single base when two or more individuals are compared. Sequence differences occur due to single nucleotide substitutions or insertions/deletions. SNPs occur in coding and non-coding regions of nuclear DNA, and constitute the most abundant markers in eukaryotic genomes. They are widely distributed and most loci are biallelic. Improved sequencing technology and available genomic sequence data have made it possible to analyse directly genetic variation in DNA. Signatures of adaptive genetic variation in wild populations can be detected without prior identification of candidate genes (Elmer and Meyer 2011). This is done by conducting genome scans of SNPs by using “next generation” DNA sequencing (section 4). Gloria-Soria et al. (2016) developed a panel of ~73,000 SNPs distributed across the genome of *G. f. fuscipes* (*Gff*) that was used to estimate genome-wide genetic diversity, differentiation and linkage disequilibrium in wild *Gff* populations in Uganda.

4. RELATIVE ADVANTAGES AND DISADVANTAGES OF PRINCIPAL METHODS

4.1. *Obsolete Methods*

Several methods have become obsolete because much more efficient techniques have become available. Falling into disuse are studies of allozymes, used to assess enzymatic protein variation. Also, obsolete are the SSCP (single strand conformational polymorphism) technique to examine mitochondrial variation, RFLPs (restriction fragment length polymorphisms), RAPDs (randomly amplified polymorphic DNA), and AFLPs (amplified fragment length polymorphisms), all of which can be used to examine elements of genomic DNA.

4.2. Mitochondrial Variations

Regions of the mitochondrial genome are highly variable, and some may be useful as population-specific markers. The mitochondrion is inherited as a single locus, and the effective population size (Box 1) is about a quarter of the size estimated by nuclear loci. This makes mitochondrial variation more sensitive to the effects of historical reductions in population size ("bottlenecks") and colonizing episodes. The reason is that the mitochondrial genome is single copy (haploid) and maternally inherited.

Box 1. Basic Formulae Used in Population Genetics

Gene heterozygosity (diversity) at a locus is one minus the frequency of homozygotes expected by Hardy-Weinberg criteria, corrected for sampling bias (Nei 1987):

$$h_e = 2n(1 - \sum x_i^2)/(2n - 1)$$

where x is the frequency of allele i , and n is the sample size. For measuring diversity at haplotypes, use $n/(n-1)$ as the correction factor. The average heterozygosity H_e over s loci is,

$$H_e = \sum h_e/s$$

with variance

$$\text{Var} = \sum (H_e - h_e)^2/[s(s-1)]$$

Under Hardy-Weinberg assumptions, the relationship between diversity H_e , mutation rate ν , and effective population size N_e is:

$$H_e = 4N_e\nu/(4N_e\nu + 1)$$

Effective population size is, roughly, the number of successfully reproducing individuals. Note that $4N_e$ is likely to be a very large number and the mutation rate a very small number, of the order 10^{-6} for allozyme loci; both are difficult to measure.

The index F describes departures from random mating by taking the difference between expected heterozygosity h_e and the observed h_o as a fraction of the heterozygosity expected on Hardy-Weinberg criteria:

$$F = (h_e - h_o)/h_e, \text{ or } F = 1 - (h_o/h_e)$$

It gives the equilibrium value of the proportion of homozygous genotypes. F is a sensitive index of breeding structure, and may be expressed for hierarchical levels in subdivided populations. (It is described further in Box 3.)

4.3. Chromosomal Polymorphisms

Chromosomal polymorphisms occur in many species. Cytological examination of pest taxa can provide important evidence of population subdivisions, variation in sex-determining mechanisms, and indicators of natural selection. There are a number of common structural variations in chromosome morphology that can be observed by microscopic examination of appropriate tissues. Such tissues include brain, testis, ovarian and trichogen cells, and salivary glands. Giant multi-banded

“polytene” chromosomes can be found in tsetse trichogen and mosquito salivary glands and ovarian nurse cells; polytenes arise from a large number of chromosome reduplications without cell separation -- thus giant cells with greatly enlarged chromosomes. They are particularly well-developed in nematoceros Diptera, and can also be found in higher Diptera, e.g. *Glossina* spp.

Sex chromosomes (“heterosomes”) in many taxa indicate the phenotypic sex, e.g. in most Diptera, homogametic XX for females and heterogametic XY for males, but male Lepidoptera are homogametic and the females heterogametic. High frequencies of sex chromosome polymorphisms were detected in *G. p. palpalis* (Maudlin 1979) indicating that further cytological work on the *G. palpalis* complex is warranted.

Among the most common chromosomal polymorphisms are paracentric inversions, where two breaks occur in a chromosome arm followed by a 180-degree rotation of the interstitial piece and a rejoin of the three pieces. Most, if not all, seem to respond to selective forces, and therefore do not provide unbiased estimators of gene flow. Nevertheless, they may provide striking indicators of genetic differentiation of populations. Pericentric inversions include the centromere within the interstitial central part of the chromosome; such an inversion can greatly change the appearance of a chromosome.

Cytological analysis can be time-consuming and requires much skill, and therefore adequate and representative sampling of natural populations can be a problem (Augustinos et al. 2014).

4.4. Microsatellites and Single Nucleotide Polymorphisms

SSRs can be well distributed throughout the genome and highly polymorphic, which has both advantages and disadvantages. Microsatellite alleles can prove useful in pinpointing the source or sources of immigrant insects in a treated area. Automated scanning and scoring of acrylamide gels allow high throughputs [by using Genemarker (SoftGenetics, USA) or Genescan®, Genotyper® and GeneMapper® software (Applied Biosystems, ThermoFisher Scientific, USA)]. Null alleles, usually caused by mutations in primer annealing sites, lead to underestimates of heterozygotes. Some taxa have few SSRs, including some culicine mosquitoes and ticks (Fagerberg et al. 2001). High mutation rates occur at some loci, violating a key Hardy-Weinberg assumption, and leading to homoplasy and biased estimates of gene flow and genetic differentiation of populations. Homoplasy, the convergent development of genotypically different but phenotypically similar repeats, is common in microsatellites. Highly variable loci with many alleles provide downwardly biased estimates of gene flow and genetic differentiation (because as diversities $h_o/h_e \rightarrow 1$, $F \rightarrow 0$). Since many microsatellite loci do not conform well to the infinite allele model (the stepwise mutation model would be more appropriate) (Box 2), some analytical procedures, e.g. F statistics estimation, may be inappropriate, and it then becomes necessary to use other methods. Some problems related to SSRs are avoided by using SNPs. Nevertheless, microsatellite loci and SNPs have become the principal tools with which to examine the genetics of

populations, leading to an explosion of analytical methods to examine the plethora of data now available (Luikart and England 1999; Excoffier and Heckel 2006).

Box 2. Stepwise Mutation Model (SMM)

Change in microsatellite repeat number is thought to occur principally in a stepwise fashion by unequal crossing over and replication slippage. Two or more independently arisen alleles can have the same molecular weight, and therefore appear as one on gels (homoplasy). Clearly, there are upper and lower limits to the number of repeats a polymorphic locus can have, so the infinite alleles model of mutation is inappropriate. Thus, in theory, the stepwise mutation model (SMM) is most appropriate when dealing with microsatellite loci. Slatkin (1995) developed a model based on the SMM analogous to Wright's F statistics. He termed the index of gene flow R_{ST} . The basic difference is that F statistics are based on variance in allele frequencies, and R_{ST} is based on variance in repeat number. However, simulations have shown that, when sample sizes are less than 50 and the number of loci scored are less than about 20, F_{ST} provides less biased estimates of gene flow than R_{ST} (Gaggiotti et al. 1999).

4.5. DNA Sequencing

Next generation (high throughput) sequencing describes several advanced, rapid and efficient sequencing methods now available (Mardis 2008; Schuster 2008; Seeb et al. 2011; EMBL-EBI 2017). Refer to Danecek et al. (2011) for formatting high throughput DNA sequencing for further use in population and phylogenetic work. Also, refer to Frías-López et al. (2016) for appropriate software for the same purpose. High throughput sequencing now enables a more comprehensive view of how genomes vary spatially and temporally than do the older methods.

5. BASIC PRINCIPLES OF POPULATION GENETICS

As cited earlier, several academic texts offer comprehensive treatments of population and ecological genetics. Here we offer only an elementary treatment of the subject.

The Hardy-Weinberg theorem, or rule, is the fundamental template on which genetic data are tested. The rule specifies that in diploid, sexually reproducing organisms, genotypic frequencies remain unchanged from generation to generation if the locus under consideration is selectively neutral, mating is random, generations are discrete, the mutation rate negligible, and population size infinite, so that no sampling errors occur in the transmission of gametes from one generation to the next. The foregoing assumptions, however, apply to few, if any, natural populations. Violations of Hardy-Weinberg assumptions may cause departures from random mating within and among populations. The principle cause is genetic drift — the random change in small gene pools due to sampling errors that occurs when the gametes of a parental generation unite to form the individuals of the next generation.

For diploid genomes, the probability of fixation of a neutral allele, i.e. attains a frequency of 1, is $1/2N_e$, where N_e is the “effective” population size (Box 1). For uniparentally inherited loci, the probability is $1/N_e$. Thus, fixation, i.e. loss of diversity, is inversely proportional to the effective population size N_e . It should be

noted that the effective population size may be much less than the census number because not all individuals reproduce, and some generate more progeny than others.

Most species are distributed discontinuously in space and time, so individuals in a population will have a greater opportunity to interbreed with receptive individuals in their own population than with individuals in other populations. Thus, matings will not ordinarily be random among populations separated by distance, altitude, or age structure. The degree of genetic differentiation among populations provides a measure of gene flow.

Selection acting on a locus can also cause departures from Hardy-Weinberg expectations. Most genetic markers used in population genetics research are assumed to be selectively neutral. However, many allozyme loci are not selectively neutral, but respond to balancing selection. An interesting example of balancing selection is that of the tsetse fly *Glossina pallidipes* Austen. Populations in Zambia and Zimbabwe show about the same level of allozyme heterozygosity as populations in Kenya, but with much reduced levels of microsatellite and mitochondrial variation (Krafsur 2002).

6. PROCESSING AND INTERPRETING POPULATION GENETIC DATA

Having scored haplotype or genotypic frequencies, it is necessary to test hypotheses about their distributions within individuals (if dealing with nuclear variation) and at hierarchical levels of population structure. Much software is available for these tasks. (A website maintained by Joseph Felsenstein at the University of Washington lists, with annotations, 392 free software packages and 54 free web servers.) Excoffier and Heckel (2006) offer a helpful list of programmes, while Lischer and Excoffier (2012) provide an automated data-conversion tool for connecting genetic data with analytic programmes.

The first step is to examine gene diversities for each population, and record the observed heterozygosity h_o at each locus. Next, calculate the expected single locus heterozygosities h_e , together with the expected heterozygosity averaged over loci H_E (Box 1). Allele or haplotype frequencies can be tested for homogeneity over sampling units by using the Chi-square contingency tests (Rice 1989) (these are incorporated into most software packages). If gene frequencies at all or most loci are homogenous among populations, then there is little likelihood of genetic differentiation among populations, assuming one is dealing with selectively neutral variation. It is then likely that the sampled populations behave as a single randomly breeding unit, and the SIT would likely have to contend with high immigration rates. Such a result does not necessarily argue against the successful application of the SIT, as is discussed below. However, it is most likely that allele frequencies among populations will differ, a consequence of population structure and genetic drift. Over time, drift and genetic differentiation can lead to the development of reproductive isolation and, eventually, to speciation (Dobzhansky 1970; Futuyma 1998; Avise 2004).

The magnitude of genetic drift is inversely proportional to the effective population size (Box 1), i.e. the geometric mean number of individuals that contributes progeny to a subsequent generation, and more formally, the size of an

ideal population experiencing the same rates of random genetic change as the real population being considered. Genetic drift is most expeditiously quantified by estimating F_{ST} or its analogues (Box 3).

Box 3. *F Statistics*

Wright's F statistics and analogues are used to describe the breeding structure of subdivided populations. The original statistical theory was formulated in terms of two alleles segregating at a single locus. F can be defined in terms of inbreeding, as departures from random mating, and as correlations between uniting gametes.

F_{IS} describes departures from random mating within populations; it is the correlation of random alleles within individuals averaged over populations, relative to the correlation of two random alleles from the population as a whole. Thus, the heterozygote deficiency (increase of homozygosity or inbreeding) in individuals is (Hartl and Clark 2007): $F_{IS} = (H_S - H_I)/H_S = 1 - (H_I/H_S)$. Here H_I is estimated as the average observed heterozygosity among populations, and H_S is the expected heterozygosity.

F_{ST} indicates departures from random mating among populations; it is the correlation of two random alleles in a population relative to the correlation of two random alleles chosen from the entire population ($F_{ST} = 1 - [H_S/H_T]$), where H_T is the heterozygosity averaged over all populations. It was termed by Wright as the standardized measure of genetic variance among populations, and also as the "fixation index" that measures reduction of heterozygosity, or heterozygote deficit compared with Hardy-Weinberg.

F_{IT} describes departures from random mating accruing from all causes ("inbreeding"); it is the correlation of alleles in individuals relative to that in the entire population ($F_{IT} = 1 - [H_I/H_T]$). Important extensions of F -statistics include G_{ST} (Nei 1987) and Weir and Cockerham's (1984) theta (θ).

F_{ST} , G_{ST} , and θ underestimate differentiation of subdivided populations progressively as gene heterozygosity h_e increases because as diversity $h \rightarrow 1$ in Hardy-Weinberg demes, F_{ST} , G_{ST} , and θ will approach zero. Indeed, they cannot take a value equal to or greater than the mean level of homozygosity (Hedrick 1999). Slatkin (1985), Barton and Slatkin (1986) and Slatkin and Barton (1989) offered a way around the problem of high diversities. Slatkin observed a nearly linear relationship between the mean frequency of private alleles p_I and the mean number of reproducing migrants per generation $N_e m$: $\log p_I = a \log_{10} N_e m + b$. Constants a and b vary with mean sample size. The method was found to be robust.

F_{ST} , G_{ST} , and θ are variance statistics. θ , for example, may be used in an analysis of variance for both diploid and haploid loci (Weir 1996). Thus, one can obtain variance statistics and estimates of θ for each level of hierarchy (Rice 1989).

The inverse of F_{ST} is proportional to the level of migration that satisfies the equation, $F_{ST} = (4N_e m + 1)^{-1}$, where the term $N_e m$ is, very roughly, the harmonic mean number of reproducing migrants per generation (Box 4). Fig. 1 shows the relationship between $N_e m$ and F_{ST} . Wright (1978a, b) suggested that a critical level of migrants is about one reproducing organism per generation. A lower sustained migration rate among equilibrium populations would lead to further genetic differentiation by drift.

For discontinuously distributed, i.e. subdivided, populations, it is helpful to make estimates of F_{ST} between all possible pairs of populations. In this way, one can test hypotheses that F_{ST} increases with geographic distance or other physical measure of interest. Highly differentiated populations can easily be identified in this way. Such local populations could, in principle, have a measure of pre-mating reproductive isolation that might work against the SIT. If deemed to be of interest, colonization,

and laboratory and field-cage testing for pre-mating isolation, could then be carried out (Cayol et al. 2002).

Box 4. Gene Flow

According to Wright's island model of gene flow, F_{ST} can be defined in relation to the effective population size N_e , mutation rate ν , and migration rate m . $F_{ST} = [4N_e m + \nu + 1]^{-1}$ for populations at mutation-drift-migration equilibrium. This equation allows the estimation of a migration rate in terms of the mean number of reproducing organisms per generation $N_e m$. Thus, where the mutation rate ν is negligible (i.e. $\rightarrow 0$), $N_e m \approx (1 - F_{ST}) / 4F_{ST}$ for diploid loci. For mitochondrial haplotypes, $N_e m \approx (1 - F_{ST}) / 2F_{ST}$ (where the sex ratio approaches unity).

F_{ST} is non-linear because it is a reciprocal function of $N_e m$ (Fig. 1). Therefore, mean values of $N_e m$ can be quite meaningless, and caution is required in their interpretation (Whitlock and McCauley 1999). Pairwise estimates of F_{ST} and its linear transformations are useful, however, and can be obtained by using "Arlequin" software (Excoffier and Lischer 2010).

Some assumptions of the island model of gene flow include populations of equal size in which drift does not occur, all of which exchange alleles with equal probability.

Other models have been formulated, including the stepping-stone model, in which only adjacent populations exchange alleles.

The island model, however, has proved sufficiently robust for many purposes, and it remains the most generally used. Theoretical developments have led to new applications, for example, the cladistic nested analysis of phylogeographical data (Templeton 1998).

An additional caution is necessary when using microsatellite loci to estimate F_{ST} . Mutation rates can be substantial at some microsatellite loci, leading to downwardly biased estimates of F_{ST} .

Figure 1. Mean numbers of reproducing migrants per generation exchanged among populations ($N_e m$) as a function of the mean departures from random mating (F_{ST}) according to the island model of gene flow. $N_e m = (1 - F_{ST}) / 4F_{ST}$.

Do high rates of dispersal argue definitively against the SIT? History indicates that they do not. Epizootiological and experimental studies have shown that the New World screwworm has a great capacity for dispersal, as might be expected in such a colonizing species. High dispersal rates of the screwworm flies did not fatally

compromise the effective application of the SIT to this species wherever it was applied.

7. POPULATION GENETICS AND CONSTRUCTION OF RELEASE STRAINS

How may the rules of population genetics be used to construct strains for eventual release? Thriving mass-reared colonies of the target insect are required, the progeny of which, when radiosterilized, must be competitive with wild males in seeking to mate with wild females (Lance and McInnis, this volume; Parker, Mamai et al., this volume; Parker, Vreysen et al., this volume).

It is often assumed that the “degradation” of laboratory stocks occurs via adaptation through selection and inbreeding, with a correlating loss of field competitiveness. Work on *Drosophila* showed a great loss of gene diversity (Briscoe et al. 1992) and fitness that was proportional to the duration of laboratory colonization (Jungen and Hartl 1979). Also, Cioci et al. (2014) detected a significant loss of microsatellite diversity arising from a genetic “bottleneck” in a ca. 38-year-old *Glossina pallidipes* colony. On the other hand, no substantial loss of mitochondrial variation was detected in *Glossina pallidipes* (Krafsur et al. 2016) or *Glossina morsitans* Westwood (Wohlford et al. 1999). Thus, each case must be taken on its own merits.

During laboratory adaptation, predation pressure is absent in colonies. Equally, sexual selection that operates in nature may be relaxed because of the proximity of mates in cages. For example, contact cuticular hydrocarbon pheromones in higher Diptera (Carlson et al. 1993; Brown et al. 1998), courtship behaviour in tephritid fruit flies, and swarming behaviour in nematoceros Diptera, could, in principle, inadvertently become modified during laboratory adaptation (Eberhard 2000; Briceño and Eberhard 2002; Briceño et al. 2002). The genetics and evolution of pheromone signalling and response behaviour have been studied in moths (Löfstedt 1993; Phelan 1997). Inbreeding causes a loss of genetic variation that could, in principle, reduce competitiveness. Are these realistic concerns? Is competitiveness in the field inversely proportional to laboratory adaptation? Do laboratory strains/colonies “deteriorate” over time?

The theory of inbreeding is well known; a simple account of elementary inbreeding theory is given in Box 5. A substantial loss of heterozygosity may occur if a release strain is formed from too few founding insects, followed by a prolonged “bottleneck” in colony size. However, it is by no means certain that even a great loss in diversity will have pronounced effects on laboratory fitness or mating competitiveness in the field. Many colonizing species undergo periodic genetic bottlenecks in nature, with little adverse effect. The question of laboratory adaptation and competitiveness must be addressed on a case-by-case basis. In the New World screwworm, for example, estimates of sterile mating rates among geographically diverse target populations were found to be unrelated to the strain of sterile flies released (Krafsur 1985, 1994). Experience with *Glossina morsitans* suggested that colonized and irradiated flies dispersed and mated with their wild cousins at approximately the expected frequencies (Vale et al. 1976; Dame 1979; Vreysen et al. 2011).

Box 5. Inbreeding

Some consequences of inbreeding and drift include high frequencies of homozygosis and fixation of mildly deleterious alleles. Correlations between uniting gametes in “inbred” populations are greater than the correlations drawn from reference populations higher in the hierarchy. Loss of genetic diversity through drift is to be expected in closed populations. The inbreeding coefficient F (probability of identity by descent) in generation t is related to the effective population size N_e as follows:

$$F_t = 1 - (1 - 1/2N_e)^t \quad \text{As } t \text{ increases, } F \rightarrow 1, \text{ hence diversity} \rightarrow 0$$

Manipulating this equation shows that about half the diversity at selectively neutral loci is lost in $1.4N_e$ generations. For a breeding colony established from 2, 5, 10, 25, and 100 pairs the initial loss of diversity in a single generation is 22, 10, 5, 2, and 0.5%, respectively. These are trivial decreases. In a constantly expanding insect culture there would be little additional loss of variation. After 10 generations of “inbreeding” at the foregoing population sizes, however, heterozygosity losses become far greater, at 92, 63, 39, 12, and 5 %, respectively. Thus, the duration of a “bottleneck” has profound effects on genetic variation. It is well to remember that these formulae apply to selectively neutral loci, but insects in culture are subject to various forms of selection, some of which may be for traits not advantageous in the wild. It should be noted, however, that weak selection intensity in very small populations can be overridden by drift (Black et al. 2001).

In some cases, genetic phenomena have been invoked to explain pest outbreaks during sterile insect release programmes (section 8), but few genetic data were available to support the contentions. Nowadays, however, it is easy to obtain genetic data for target populations, and for laboratory colonies and their source populations. Inbreeding coefficients (F) and genetic distances can (and should) be estimated continuously as part of quality control programmes. A progressive loss of diversity in laboratory colonies, coupled with a progressive decay in physiological quality control indices, could provide *prima facie* evidence of degradation, and possibly predict a decline in field competitiveness. If no data are collected, little of substance can later be said about genetic causes of, or correlations with, the failure of a particular programme.

The rationale for constructing strains for propagation and release should include the number of founding organisms and their geographical origins. If there is good reason to believe that a target pest population may constitute a species complex, then it is necessary to ensure that only the pest species is cultured. Isfemale strains can be constructed for insects that reproduce readily in colony (an isfemale strain is based upon the progeny of a single mated pair). Many such lines can be constructed, and each evaluated for laboratory fitness and field competitiveness. If it can be shown that geographically adapted populations exist (Vera et al. 2006), then it would be prudent to conduct mating preference tests under semi-natural conditions (Cayol et al. 2002). Such tests have strong environmental correlations, and thus experimental designs must be adequate to detect unambiguously a true genetic component. Where strong, genetically determined mating preferences exist, of course it would be advisable to rear, sterilize, and release the most compatible strains.

The question of loss of diversity and fixation of deleterious alleles because of small effective population size can, in principle, become acute when constructing

strains by classical or modern molecular tools; the progenies are descended from a unique genome that may or may not be subject to recombination with wild-type alleles (Franz et al., this volume).

Finally, it is most important to put in perspective the issue of “strain degradation”. It is difficult and expensive to conduct field experiments that estimate the genetic and environmental components of competitiveness. In projects that release sterile insects, however, sterile mating rates in target pest populations should be monitored routinely (Parker, Vreysen et al., this volume; Vreysen, this volume). Experience has shown that rearing, handling, irradiation, and release procedures are much more important for competitiveness than the laboratory “degradation” of release strains. A good example is the North and Central American New World screwworm eradication programme (see below).

There are new, more efficient, and faster computer programmes suitable for the analysis of population genetic data. Principal component analysis (PCA) has become a standard analytical tool in population genetics. Many applications employ Bayesian methods that allow complex and subtle inferences of such data. Although beyond the scope of this chapter, Box 6 provides a brief introduction to PCA and Bayesian applications.

Box 6. Principal Components and Bayes Theorem

Principal component analysis is convenient and easy to use. Graphic presentations can be colorful and informative. In PCA, eigenvectors and eigenvalues are calculated from allele frequency covariance matrices of all pairwise comparisons. Covariance is the mean value of the product of deviations of two variates from their respective means. Patterson et al. (2006) offer further information and cite convenient computer programmes.

The application of Bayes theorem or rule offers a useful tool to estimate probabilities of events based on prior knowledge related to the event of interest and to compare gene flow models. Bayesian methods can be used to infer phylogenies (Felsenstein 2004), i.e. evolutionary relationships among taxa or geographically diverse populations.

At its mathematically simplest, Bayes theorem is given by $P(A|B) = [P(B|A) \times P(A)] / P(B)$ where $P(A)$ is the probability of A and $P(B)$ is the probability of B; $P(A|B)$ is the probability of A given B, and $P(B|A)$ is the probability of B given A. Here, $P(A)$ is the “prior” probability, and $P(A|B)$ is the “posterior” probability.

Approximate Bayesian Computation (ABC) bypasses some difficult mathematical procedures (Beaumont 2010), and provides a flexible platform to carry out complex models. ABC is useful in investigating demographic histories and evolutionary questions. ABC is incorporated into a growing list of statistical procedures and genetical software packages, e.g. DIYABC v2.1.0 (“Do It Yourself ABC”) (Cornuet et al. 2014) and MrBayes 3 (Ronquist and Huelsenbeck 2003). DIYABC is software for comprehensive analysis of population history using approximate Bayesian computation (ABC) on DNA polymorphism data. MrBayes is a programme for Bayesian inference and model choice across a wide range of phylogenetic and evolutionary models. MrBayes uses a Markov chain Monte Carlo simulation method to estimate the expectation of a statistic in a complex model. Another software is ABCtoolbox (Wegmann et al. 2010).

8. APPLICATIONS AND EXAMPLES

Experience suggests that studies of gene flow throughout a species' range are not necessary to implement and successfully carry out programmes that release sterile insects. Mediterranean fruit fly and New World screwworm programmes were conducted successfully for years without any substantial population genetic data available. However, the spectre of reproductively isolated screwworm populations was raised repeatedly, particularly when progress was slow or releases deemed ineffective (Krafsur et al. 1987). Some investigators were so certain of the existence of assortative matings between wild and released flies, and of reproductively isolated forms, that numerous claims were made in anticipation of their certain discovery. Indeed, a few geneticists played a controversial and unhelpful role (in the New World screwworm eradication programme in the USA and Mexico) by claiming evidence of assortative matings (Bush et al. 1976), or later (Richardson 1978; Richardson et al. 1982), by claiming evidence that screwworms constituted a hitherto unrecognized complex of reproductively isolated forms.

This situation, precipitated by screwworm epizootics in 1972–1976 in the southwestern USA, was exacerbated by the practice of presenting and evaluating eradication programme results solely on screwworm case incidence. Programme officials invoked screwworm migration across the border from Mexico and rainfall to explain the outbreaks, but no gene flow or meteorological data were ever provided. The possibility of genetic changes in target populations was also acknowledged. In the absence of published meaningful data to support programme claims, interested commentators offered their own explanations, attracting much notice in the scientific community, and leading to attempts to stop the development of the Mexico screwworm eradication programme until critics could solve the very problems they claimed to have identified.

There was very little evidence that the screwworm outbreaks could be explained by powerful genetic phenomena or the existence of cryptic species. There was much evidence to falsify the claims, had it been recognized and invoked by programme officials (LaChance et al. 1982; McInnis 1983; Krafsur et al. 1987; Krafsur 1998). One of the lessons to be learned from this episode is that periodic estimates of gene frequencies from wild and mass-reared insects of a targeted species, together with predictive quality control measures, can be used to track changes that could, in principle, point to a decline in strain effectiveness for SIT implementation. The availability of such data, particularly when published in peer-reviewed scientific journals, can do much to defuse potentially damaging and confounding criticism from the scientific community. It would also be necessary, in those species where it is practicable, to maintain estimates of sterile mating frequencies -- to show a relationship between sterility and sterile insect dose rates and target population densities (Krafsur and Hightower 1979; Krafsur 1994; Vreysen, this volume).

Pest outbreaks during a sterile insect release programme can reasonably be expected, and it will be necessary to specify the most likely causes. It must be anticipated that natural events, such as weather or the evolution of assortative mating, can, in principle, occur, but experience has shown that it is much more likely that failures in insect production, the distribution of sterile insects or the

implementation of complementary suppression measures will have caused this breakdown in control.

9. REFERENCES

- Abdelkrim, J., B. C. Roberston, J. L. Stanton, and N. J. Gemmell. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46: 185–192.
<https://www.biotechniques.com/BiotechniquesJournal/2009/March/Fast-cost-effective-development-of-species-specific-microsatellite-markers-by-genomic-sequencing/biotechniques-118341.html>
- Aguirre-Ramirez, E. J., S. M. Velasco-Cuervo, J. J. Gallo-Franco, R. Gonzles, N. S. Carrejo, and N. Toro-Perea. 2017. Genetic diversity and population structure of *Anastrepha obliqua* in southwestern Colombia. *Entomologia Experimentalis et Applicata* 164(Issue 3): 291–304.
<https://doi.org/10.1111/eea.12613>
- Augustinos, A. A., E. Drosopoulou, A. Gariou-Papalexiou, K. Bourtzis, P. Mavragani-Tsipidou, and A. Zacharopoulou. 2014. The *Bactrocera dorsalis* species complex: comparative cytogenetic analysis in support of sterile insect technique applications. *BMC Genetics* 15 (Suppl. 2): S16.
<https://doi.org/10.1186/1471-2156-15-S2-S16>
- Avise, J. C. 2004. Molecular markers, natural history and evolution, 2nd edn. Sinauer Associates, Sunderland, MA, USA.
- Barton, N. H., and M. Slatkin. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56: 409–415.
- Bates, M. 1949. The natural history of mosquitoes. Macmillan, London, UK.
- Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* 41: 379–406.
<https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Black, W. C. IV, C. F. Baer, M. F. Antolin, and N. M. DuTeau. 2001. Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology* 46: 441–469.
<https://doi.org/10.1146/annurev-ento.46.1.441>
- Briceo, R. D., and W. G. Eberhard. 2002. Decisions during courtship by male and female medflies (Diptera, Tephritidae): correlated changes in male behavior and female acceptance criteria in mass-reared flies. *Florida Entomologist* 85: 14–31. <http://www.fcla.edu/FlaEnt/fe85p14.pdf>
- Briceo, R. D., W. G. Eberhard, J. C. Vilardi, P. Liedo, and T. E. Shelly. 2002. Variation in the intermittent buzzing songs of male medflies (Diptera: Tephritidae) associated with geography, mass-rearing, and courtship success. *Florida Entomologist* 85: 32–40.
<http://www.fcla.edu/FlaEnt/fe85p32.pdf>
- Briscoe, D. A., J. M. Malpica, A. Robertson, G. J. Smith, R. Frankham, R. G. Banks, and J. S. F. Barker. 1992. Rapid loss of genetic variation in large captive populations of *Drosophila* flies: implications for the genetic management of captive populations. *Conservation Biology* 6: 416–425.
<http://onlinelibrary.wiley.com/doi/10.1046/j.1523-1739.1992.06030416.x/full>
- Brown, W. V., R. Morton, M. J. Lacey, J. P. Spradbery, and R. J. Mahon. 1998. Identification of the geographical source of adults of the Old World screw-worm fly, *Chrysomya bezziana* Villeneuve (Diptera: Calliphoridae), by multivariate analysis of cuticular hydrocarbons. *Comparative Biochemistry and Physiology – Part B: Biochemistry and Molecular Biology* 119: 391–399.
[https://doi.org/10.1016/S0305-0491\(97\)00365-9](https://doi.org/10.1016/S0305-0491(97)00365-9)
- Bush, G. L., R. W. Neck, and G. B. Kitto. 1976. Screwworm eradication: inadvertent selection for noncompetitive ecotypes during mass rearing. *Science* 193: 491–493.
- Carlson, D. A., S. K. Milstrey, and S. K. Narang. 1993. Classification of tsetse flies *Glossina* spp. (Diptera: Glossinidae) by gas chromatographic analysis of cuticular components. *Bulletin of Entomological Research* 83: 507–515.
- Cayol, J. P., P. Coronado, and M. Taher. 2002. Sexual compatibility in medfly (Diptera: Tephritidae) from different origins. *Florida Entomologist* 85: 51–57. <http://www.fcla.edu/FlaEnt/fe85p51.pdf>
- Ciosi, M., D. K. Masiga, and C. M. R. Turner. 2014. Laboratory colonisation and genetic bottlenecks in the tsetse fly *Glossina pallidipes*. *PLOS Neglected Tropical Diseases* 8(2): e2697.
<https://doi.org/10.1371/journal.pntd.0002697>
- Colluzi, M., A. Sabatini, A. della Torre, M. A. Di Deco, and V. Petrarca. 2002. A polytene

- chromosome analysis of the *Anopheles gambiae* species complex. *Science* 298: 1415–1418.
DOI: 10.1126/science.1077769 <http://science.sciencemag.org/content/298/5597/1415>
- Cornuet, J.-M., P. Pudlo, J. Veyssier, A. Dehne-Garcia, M. Gautier, R. Leblois, J.-M. Marin, and A. Estoup. 2014.** DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30(8): 1187–1189. <https://doi.org/10.1093/bioinformatics/btt763>
- Crow, J. F., and M. Kimura. 1970.** An introduction to population genetics theory. Burgess, Minneapolis, MN, USA.
- Dame, D. A. 1979.** The sterile insect technique against tsetse flies, *Glossina* spp. *Bulletin of the Entomological Society of America* 16: 24–30.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011.** The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.
<https://doi.org/10.1093/bioinformatics/btr330>
- Dobzhansky, T. 1970.** Genetics of the evolutionary process. Columbia University Press, NY, USA.
- Dorn, P. L., F. C. Noireau, E. S. Krafur, G. C. Lanzaro, and A. J. Cornel. 2011.** Genetics of major insect vectors, pp. 411–472. *In* M. Tibayrenc (ed.), Genetics and evolution of infectious diseases. Elsevier, Amsterdam, The Netherlands. DOI: 10.1016/B978-0-12-384890-1.00015-7
<https://ucdavis.pure.elsevier.com/en/publications/genetics-of-major-insect-vectors>
- Eberhard, W. G. 2000.** Sexual behavior and sexual selection in the medfly, *Ceratitis capitata*, pp. 459–489. *In* M. Aluja and A. L. Norrbom (eds.), Fruit flies (Tephritidae): phylogeny and evolution of behavior. CRC Press, Boca Raton, FL, USA.
- Elmer, K. R., and A. Meyer. 2011.** Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology and Evolution* 26: 298–306.
<https://dx.doi.org/10.1016/j.tree.2011.02.008> <http://nbn-resolving.de/urn:nbn:de:bsz:352-195140>
- (EMBL-EBI) European Bioinformatics Institute. 2017.** <http://www.ebi.ac.uk/training>
- Excoffier, L., and G. Heckel. 2006.** Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics* 7: 745–758. DOI: 10.1038/nrg1904
<http://www.nature.com/nrg/journal/v7/n10/full/nrg1904.html>
- Excoffier, L., and H. E. L. Lischer. 2010.** Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1755-0998.2010.02847.x/abstract>
- Fagerberg, A. J., R. E. Fulton, and W. C. Black IV. 2001.** Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. *Insect Molecular Biology* 10: 225–236. <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2583.2001.00260.x/abstract>
- Felsenstein, J. 2004.** Inferring phylogenies. Sinauer Assoc. Inc., Sunderland, MA, USA.
<http://evolution.genetics.washington.edu/phylip/software.html>
- Frías-López, C., J. F. Sánchez-Herrero, S. Guirao-Rico, E. Mora, M. A. Arnedo, A. Sánchez-Gracia, and J. Rozas. 2016.** DOMINO: development of informative molecular markers for phylogenetic and genome-wide population genetic studies in non-model organisms. *Bioinformatics* 32: 3753–3759. <https://doi.org/10.1093/bioinformatics/btw534>
- Futuyma, D. J. 1998.** Evolutionary biology, 3rd edn. Sinauer Associates, Sunderland, MA, USA.
- Gaggiotti, O. E., O. Lange, K. Rassmann, and C. Gliddon. 1999.** A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8: 1513–1520. <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-294x.1999.00730.x/full>
- Gillespie, J. H. 2004.** Population genetics: a concise guide. Johns Hopkins Univ. Press, Baltimore, MD.
- Gloria-Soria, A., W. A. Dunn, E. L. Telleria, B. R. Evans, L. Okedi, R. Echodu, W. C. Warren, M. J. Montague, S. Aksoy, and A. Caccone. 2016.** Patterns of genome-wide variation in *Glossina fuscipes fuscipes* tsetse flies from Uganda. *G3: Genes, Genomes, Genetics* 6(6): 1573–1584.
<https://doi.org/10.1534/g3.116.027235>
- Gooding, R. H., and E. S. Krafur. 2005.** Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Annual Review of Entomology* 50: 101–123.
<https://doi.org/10.1146/annurev.ento.50.071803.130443>
- Hamilton, M. B., E. L. Pincus, A. DiFiore, and R. C. Fleischer. 1999.** Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques* 27: 500–507. https://www.biotechniques.com/multimedia/archive/00011/99273st03_11745a.pdf
- Hartl, D. L., and A. G. Clark. 2007.** Principles of population genetics, 4th edn. Sinauer Associates,

- Sunderland, MA, USA.
- Hedrick, P. W. 1999.** Highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313–318. <http://dx.doi.org/10.2307/2640768>
- Hedrick, P. W. 2011.** Genetics of populations, 4th edn. Jones and Bartlett Publishers, Sudbury, MA.
- Hendrichs, J., T. Vera, M. de Meyer, and A. Clarke. 2015.** Resolving cryptic species complexes of major tephritid pests. *ZooKeys* 540: 5–39. <https://doi.org/10.3897/zookeys.540.9656>
http://zookeys.pensoft.net/browse_journal_issue_documents.php?issue_id=763
- Jungen, H., and D. L. Hartl. 1979.** Average fitness of populations of *Drosophila melanogaster* as estimated by using compound chromosomes. *Evolution* 33: 359–370.
- Kambhampati, S., W. C. Black, and K. S. Rai. 1992.** Random amplified polymorphic DNA of mosquito species and populations (Diptera: Culicidae): techniques, statistical analysis, and applications. *Journal of Medical Entomology* 29: 939–945. <https://doi.org/10.1093/jmedent/29.6.939>
- Krafsur, E. S. 1985.** Screw worm flies (Diptera: Calliphoridae): analysis of sterile mating frequencies and covariates. *Bulletin Entomol. Soc. America* 31(3): 36–40. <https://doi.org/10.1093/besa/31.3.36>
- Krafsur, E. S. 1994.** Application of models to screw worm eradication programs, pp 299–308. *In* B. Perry and J. Hansen (eds.), *Proceedings, Workshop: Modelling Vector-Borne and Other Parasitic Diseases*. ILRAD/FAO, 1992, Nairobi, Kenya.
- Krafsur, E. S. 1998.** Sterile insect technique for suppressing and eradicating insect populations: 55 years and counting. *Journal of Agricultural Entomology* 15: 303–317.
<http://scentoc.org/Volumes/JAE/v15/4/00154303.pdf>
- Krafsur, E. S. 2002.** Population structure of the tsetse fly *Glossina pallidipes* estimated by allozyme, microsatellite and mitochondrial gene diversities. *Insect Molecular Biology* 11: 37–45.
<https://doi.org/10.1046/j.0962-1075.2001.00307.x>
- Krafsur, E. S., and B. G. Hightower. 1979.** Field tests of sterile screw worm flies, *Cochliomyia hominivorax* (Diptera: Calliphoridae), against natural populations in three coastal areas of Mexico. *Journal of Medical Entomology* 16: 33–42. <https://doi.org/10.1093/jmedent/16.1.33>
- Krafsur, E. S., and J. J. Obrycki. 2000.** *Coleomegilla maculata* (Coleoptera: Coccinellidae) is a species complex. *Annals of the Entomological Society of America* 93(5): 1156–1163.
[https://doi.org/10.1603/0013-8746\(2000\)093\[1156:CMCCIA\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2000)093[1156:CMCCIA]2.0.CO;2)
- Krafsur, E. S., C. J. Whitten, and J. E. Novy. 1987.** Screw worm eradication in North and Central America. *Parasitology Today* 3(5): 131–137. [https://doi.org/10.1016/0169-4758\(87\)90196-7](https://doi.org/10.1016/0169-4758(87)90196-7)
- Krafsur, E. S., J. G. Marquez, and J. O. Ouma. 2016.** Phylogeography and genealogy of the tsetse fly *Glossina pallidipes* (Diptera: Glossinidae). *International Journal of Tropical Insect Science* 36: 32–47. <https://doi.org/10.1017/S1742758415000223>
- LaChance, L. E., A. C. Bartlett, R. A. Bram, R. J. Gagne, O. H. Graham, D. O. McInnis, C. J. Whitten, and J. A. Seawright. 1982.** Mating types in screw worm populations? *Science* 218: 1142–1143.
- Lanzaro, G., K. Ostrovskaya, M. Herrero, P. Lawyer, and A. Warburg. 1993.** *Lutzomyia longipalpis* is a species complex: genetic divergence and inter-specific hybrid sterility among three populations. *American Journal of Tropical Medicine and Hygiene* 48: 839–847.
- Lischer, H. E. L., and L. Excoffier. 2012.** PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.
<http://cmpg.unibe.ch/software/PGDSpider/>
- Löfstedt, C. 1993.** Moth pheromone genetics and evolution. *Philosophical Transactions of the Royal Society of London B* 340: 167–177.
- Luikart, G., and P. R. England. 1999.** Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution* 14(7): 253–256. [https://doi.org/10.1016/S0169-5347\(99\)01632-8](https://doi.org/10.1016/S0169-5347(99)01632-8)
- Malausa, T., A. Gilles, E. Megléc, H. Blanquart, S. Duthoy, C. Costedoat, V. Debut, N. Pech, P. Castagnone-Sereno, C. Délye, N. Feau, P. Frey, P. Gauthier, T. Guillemaud, L. Hazard, V. Le Corre, B. Lung-Escarment, P.-J. G. Malé, S. Ferreira, and J.-F. Martin. 2011.** High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Molecular Ecology Resources* 11: 638–644.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1755-0998.2011.02992.x/full>
- Mardis, E. R. 2008.** The impact of next-generation sequencing technology on genetics. *Trends in Genetics* 24: 133–141. <https://doi.org/10.1016/j.tig.2007.12.007>
- Maudlin, I. 1979.** Chromosome polymorphism and sex determination in a wild population of tsetse. *Nature* 277: 300–301.
- McInnis, D. O. 1983.** Chromosomal variation in the screw worm: polymorphism or cryptic species.

- American Naturalist 122: 840–842.
- Murphey, R. W., J. W. Sites Jr., D. G. Buth, and C. H. Haufler. 1996.** Molecular techniques, pp. 51–120. *In* D. M. Hillis, C. Moritz and B. K. Mable (eds.), Molecular systematics, 2nd edn. Sinauer Associates, Sunderland, MA, USA.
- Nei, M. 1987.** Molecular evolutionary genetics. Columbia University Press, New York, NY, USA.
- Nei, M., and S. Kumar. 2000.** Molecular evolution and phylogenetics. Oxford University Press, New York, NY, USA.
- Nielsen, R., and M. Slatkin. 2013.** An introduction to population genetics. Theory and applications. Sinauer Assoc. Inc., Sunderland, MA, USA.
- Patterson, N., A. L. Price, and D. Reich. 2006.** Population structure and eigenanalysis. PLOS Genetics 2: 2074–2093. <https://doi.org/10.1371/journal.pgen.0020190>
- Phelan, P. L. 1997.** Genetics and phylogenetics in the evolution of sex pheromones, pp 563–579. *In* R. T. Carde and A. K. Minks (eds.), Insect pheromone research: new directions. Chapman and Hall, New York, NY, USA.
- Rice, W. R. 1989.** Analyzing tables of statistical tests. Evolution 43: 223–225.
- Richardson, B. J., P. R. Baverstock, and M. Adams. 1986.** Allozyme electrophoresis: a handbook for animal systematics and population structure. Academic Press, Sydney, Australia.
- Richardson, R. H. (ed.). 1978.** The screwworm problem. University of Texas Press, Austin, TX, USA.
- Richardson, R. H., J. R. Ellison, and W. W. Averhoff. 1982.** Autocidal control of screwworms in North America. Science 215: 361–370.
- Roderick, G. K. 1996.** Geographic structure of insect populations: gene flow, phylogeography, and their uses. Annual Rev. Entomology 41: 325–352. <https://doi.org/10.1146/annurev.en.41.010196.001545>
- Ronquist, F., and J. P. Huelsenbeck. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574. <http://www.MrBayes.sourceforge.net>
- Rousset, F., and M. Raymond. 1997.** Statistical analyses of population genetic data: new tools, old concepts. Trends in Ecology and Evolution 12(8): 313–317. [https://doi.org/10.1016/S0169-5347\(97\)01104-X](https://doi.org/10.1016/S0169-5347(97)01104-X)
- Schuster, S. C. 2008.** Next-generation sequencing transforms today's biology. Nature Methods 5: 16–18. DOI: 10.1038/nmeth1156 <http://www.nature.com/nmeth/journal/v5/n1/abs/nmeth1156.html>
- Seeb, J. E., G. Carvalho, L. Hauser, K. Naish, S. Roberts, and L. W. Seeb. 2011.** Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11 (Suppl. 1): 1–8. <http://onlinelibrary.wiley.com/doi/10.1111/j.1755-0998.2010.02979.x/abstract>
- Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994.** Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87(6): 651–701. <https://doi.org/10.1093/aesa/87.6.651>
- Slatkin, M. 1985.** Rare alleles as indicators of gene flow. Evolution 39: 53–65. <https://doi.org/10.1111/j.1558-5646.1985.tb04079.x>
- Slatkin, M. 1995.** A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457–462. <http://www.genetics.org/content/139/1/457>
- Slatkin, M., and N. H. Barton. 1989.** A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43: 1349–1368. <https://doi.org/10.1111/j.1558-5646.1989.tb02587.x>
- Templeton, A. R. 1998.** Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7(4): 381–397. <https://doi.org/10.1046/j.1365-294x.1998.00308.x>
- Vajime, C. G., and W. G. Gregory. 1990.** Onchocerciasis: species complex of vectors and epidemiology. Acta Leidensia 59: 235–252.
- Vale, G. A., J. W. Hargrove, A. M. Jordan, P. A. Langley, and A. R. Mews. 1976.** Survival and behaviour of tsetse flies (Diptera, Glossinidae) released in the field: a comparison between wild flies and animal-fed and *in vitro*-fed laboratory-reared flies. Bulletin of Entomological Research 66(4): 731–744. <https://doi.org/10.1017/S0007485300010786>
- Vera, M. T., C. Cáceres, V. Wornoayporn, A. Islam, A. S. Robinson, M. H. de la Vega, J. Hendrichs, and J. P. Cayol. 2006.** Mating incompatibility among populations of the South American fruit fly *Anastrepha fraterculus* (Diptera: Tephritidae). Annals Entomological Society of America 99: 387–397. [https://doi.org/10.1603/0013-8746\(2006\)099\[0387:MIAPOT\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2006)099[0387:MIAPOT]2.0.CO;2)

- Vreysen, M. J. B., K. M. Saleh, R. Lancelot, and J. Bouyer. 2011.** Factory tsetse flies must behave like wild flies: a prerequisite for the sterile insect technique. *PLOS Neglected Tropical Diseases* 5(2): e907. <https://doi.org/10.1371/journal.pntd.0000907>
- Wegmann, D., C. Leuenberger, S. Neuenschwander, and L. Excoffier. 2010.** ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* 11: 116. <https://doi.org/10.1186/1471-2105-11-116>
- Weir, B. S. 1996.** Genetic data analysis II. Sinauer Associates, Sunderland, MA, USA.
- Weir, B. S., and C. C. Cockerham. 1984.** Estimating F -statistics for the analysis of population structure. *Evolution* 38: 1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
- Whitlock, M. C., and D. E. McCauley. 1999.** Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity* 82: 117–125. <https://www.nature.com/articles/6884960>
- Wohlford, D. L., E. S. Krafsur, N. T. Griffiths, J. G. Marquez, and M. D. Baker. 1999.** Genetic differentiation of some *Glossina morsitans morsitans* populations. *Medical Veterinary Entomology* 13: 377–385. <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2915.1999.00185.x/abstract>
- Wright, S. 1969a.** Evolution and the genetics of populations. Volume 1. Genetics and biometric foundations. University of Chicago Press, Chicago, IL, USA.
- Wright, S. 1969b.** Evolution and the genetics of populations. Volume 2. The theory of gene frequencies. University of Chicago Press, Chicago, IL, USA.
- Wright, S. 1978a.** Evolution and the genetics of populations. Volume 3. Experimental results and evolutionary deductions. University of Chicago Press, Chicago, IL, USA.
- Wright, S. 1978b.** Evolution and the genetics of populations. Volume 4. Variability within and among natural populations. University of Chicago Press, Chicago, IL, USA.