Problems in the Analysis of Genetic Differentiation Among Populations – a Case Study in *Quercus robur*

By H.-R. Gregorius[1], B. Degen[2] and A. König[2]

Abstract

The conclusions drawn from studies of genetic differentiation among populations largely determine our understanding of ecological and population genetic processes. These conclusions basically depend on the applied type of genetic marker and the method of measuring and estimating genetic differentation. However, concerns have been raised about the conceptual appropriateness of common methods of measuring genetic differentiation. The present paper contributes to the clarification of the problems involved by recalling the conceptual characteristics of F_{ST} (= G_{ST}), by specifying basic tests of the major causal factors of genetic differentiation with the help of permutation analysis, by comparing F_{ST} and Hedrick's new normalization F'_{ST} with the basic index δ of differentiation for data on allozymes and microsatellites obtained from 6 oak stands. All three descriptors display small values, among which δ is largest and closely followed by F'_{ST}, while F_{ST} is distinctly smaller than both across all loci. Degrees of covariation of δ with F_{ST} and F'_{ST} differ distinctly between allozymes and microsatellites as a probable consequence of confounding aspects of differentiation with aspects of fixation in the F_{ST} descriptors. Permutation analysis reveals that the boundary conditions provided by the number of populations and their (sample) sizes as well as the overall genetic variation across population samples determine the order of magnitude of differentiation. This mathematical artefact undermines the widely held opinion that small degrees of differentiation at many loci are the result of extensive gene flow or recent joint history. Differentiation patterns vary considerably among allozyme loci (indicating the action of homogenizing and diversifying selection). In contrast, microsatellite loci consistently display significant differentiation as can be explained by mechanisms of non-recurrent mutation. These observations apply to all three descriptors for the relatively high within population polymorphism observed in the studied stands. At least for low within population polymorphism close to fixation, however, it is shown theoretically that the predictions may diverge distinctly among the three descriptors.

Key words: population differentiation, homogenizing forces, diversifying forces, permutation analysis, analysis of covariation, oak, microsatellites, allozymes.

Introduction

Populations derived from a common ancestor are generally expected to become genetically differentiated if migration or gene flow between them is small. This ten-dency can be inhibited at particular loci, at which selection tends to widely favor the same allelic variants (see e.g. Gregorius and Bergmann, 1995) or at selectively neutral loci for persistently large populations. The fact that low genetic differentiation among populations may result either from extensive gene flow or uniformly acting selection gave rise to some debate with special reference to isozyme gene markers (see e.g. the introduction and discussion in Latta and Mitton, 1997). This situation explains the high level of interest in experimental and theoretical studies of genetic differentiation among populations with particular reference to effects of large scale gene flow. Since in most cases large scale gene flow is extremely difficult to observe directly, methods of indirect (model-based) estimation of gene flow are very popular (see e.g. Heuertz et al., 2004 or Whitlock and McCauley, 1999). Such estimates are usually based on observations of differences in genetic characteristics among collections of individuals. Most of them are indexed $_{ST}$ in order to emphasize the relation to Wright's famous measure F_{ST}.

Problems arising with the model-dependence of indirect estimates of large scale gene flow based on F_{ST} and its relatives (e.g. R_{ST} by Slatkin, 1995) are well recognized. Among the most frequently mentioned concerns are unrealistic assumptions about stationary of population size, absence of selection and mutation, or form of migration or gene flow (for a review see e.g. Whitlock and McCauley, 1999; Neigel, 1997, 2002). Therefore, indirect estimation needs careful analysis of its assumptions (see e.g. Weir and Cockerham, 1984; Neigel, 1997, p.107). Basically, indirect estimation rests on testable models, and only models that have passed a test can legitimately be used for indirect estimation of its parameters (see Gregorius, 1998, 2002). Models are not testable, for example, if they allow for perfect calibration (fit) to all possible observations.

All of the above-cited papers which review the limits of F_{ST} in the estimation of gene flow emphasize the importance of the role that F_{ST} has in a more general context as a model-independent measure of genetic differentiation among populations. This view addresses F_{ST} as a descriptor of genetic variation with certain conceptual characteristics. Concerning its concept, Wright characterized F_{ST} as a "fixation index" that is "not a measure of degree of differentiation in the sense implied in the extreme case by absence of any common allele" (Wright, 1978, p.82). In fact F_{ST} = 1 only if all populations are genetically fixed though not neccessarily to different alleles (in the following F_{ST} is understood to apply to multiple alleles and is thus identical to the G_{ST} of Nei, 1973). This includes the possibility that all popula-

tions with the exception of one are fixed for the same allele thus implying arbitrarily small degrees of differentiation among populations (GREGORIUS and ROBERDS, 1986). The implied conflict between the common notions of fixation and differentiation has become more intense with the utilization of highly polymorphic genetic markers (CHARLESWORTH, 1998; NAGYLAKI, 1998; HEDRICK, 1999, 2005).

The latter three authors provide several examples demonstrating the failure of F_{ST} and its relatives to reflect basic concepts of genetic difference both when applied to models and to data from experimental studies. They warn against ignorance of this aspect particularly when applying highly variable genetic markers. In fact, the amount of genetic difference may strongly depend on the resolving power of the developed gene markers and their screening methods. Particularly moderately polymorphic genetic markers like isoenzymes may gain allelic variability solely by re-electrophoreses or identification of sequences of the enzyme controlling genes. The resulting additional potential for genetic differentiation can thus be expected to be accounted for by an appropriate measure. In particular, such a measure would be required to (not necessarily strictly) increase with increasing resolution of the genetic trait. (In a strict sense, increasing the resolution defines a new trait; whether a trait B, say, constitutes a higher resolution of a trait A, say, can be determined by ensuring that any two individuals differing in their states of trait A also differ in their states of trait B.)

Yet, as HEDRICK (1999) pointed out, the opposite may be suggested by F_{ST}-values. Obviously, if a gene marker displayed monomorphy in each of several populations such that $F_{ST} = 1$, a higher resolution of the marker may reveal populations to be polymorphic and thus result in a decrease of F_{ST} below 1. Relating F_{ST}-values to amounts of gene flow or to bottleneck effects may thus lead to erroneous comparative conclusions. Therefore, HEDRICK (2005) recently proposed a normalized version F'_{ST} of F_{ST} that assumes its maximum value of 1 for genetically disjoint polymorphic populations. By this it is guaranteed at least that genetically disjoint populations remain so after increasing the resolution of the genetic trait.

Many if not most of the above-mentioned problems can be traced back to the preference for measures of genetic variation which either directly result from special population genetic models (like F_{ST} with respect to the "island model") or whose concepts can be justified through incorporation (probably as a parameter) into such models (like R_{ST}). Their interpretation is thus largely confined to the model context. This is different for purely conceptually reasoned descriptors of genetic variation, since rather than being characterized by a model, they themselves serve to characterize models as well as experimental data for independently specified features of variation. One such descriptor of genetic differentiation among populations is δ and its components D_j (GREGORIUS and ROBERDS, 1986). The descriptor has a number of conceptually desirable properties among which are the above-mentioned two basic requirements

that the maximum value is assumed only if all populations are genetically disjoint (including polymorphic populations) and that with increasing trait resolution the measure of differentiation increases.

It appears that the discussion concerning critical and concious choice of measures of genetic variation had only little effect on their application to problems associated with genetic differentiation among populations including the determination of spatialgenetic patterns. For this reason and because of the considerable importance in several fields including the management of genetic resources, we will try to assess some of the practical consequences in interpretation of F_{ST} and F'_{ST} and compare them with the results obtainable from δ as a model-independent and conceptually argued descriptor of differentiation. This will be done for data sets selected from an oak species (*Quercus robur*) and for gene markers of different overall degrees of polymorphism and supposed function (allozymes and microsatellites).

Because of the above-mentioned problems with model-based estimates and their tests in causal analyses and, more importantly, since conceptually different descriptors are to be compared, we had to take a less specifically model-dependent (non-parametric) approach to the analysis of data on genetic differentiation. We therefore rely on general characterizations of potential causes of genetic differentiation in combination with randomization techniques for assessment of observed descriptor values. In order to be able to compare δ with F_{ST} and F'_{ST} with respect to fulfillment of the above two basic requirements on measures of differentiation, a more detailed consideration of these two descriptors for effects of polymorphism, fixation and trait resolution is provided in an Appendix.

Material and Methods

Study stands

The six stands of *Quercus robur* analyzed in the following are located in North Germany (*Figure 1*). They are named according to their neighboring villages, or the forest district they belong to (*Table 1*). The distances between the stands varied from 15 km to 290 km. All stands were regenerated artificially, very probably by using seeds or saplings of local origin, but may also comprise some natural regeneration. There are thus no indications of recent joint history of the stands (reaching back one or two generations). The age of the stands varied between 148 and 259 years, and the oaks are mixed with other tree species, mainly with beech (*Fagus sylvatica*). In Escherode and Rantzau both *Q. robur* and *Q. petraea* occur. For all stands species identity has been verified by leaf characteristics (KREMER et al., 2002). Whereas in Behlendorf, Steinhorst, Rantzau and Escherode samples were collected from all adult trees in Lübeck and Krummesse only randomly selected trees were sampled. For allozyme analysis winter buds were collected, and for microsatellite analysis as a rule leaves were used. Results on genetic differences between the stands "Behlendorf" and "Steinhorst" have been reported before (DEGEN et al., 1999).
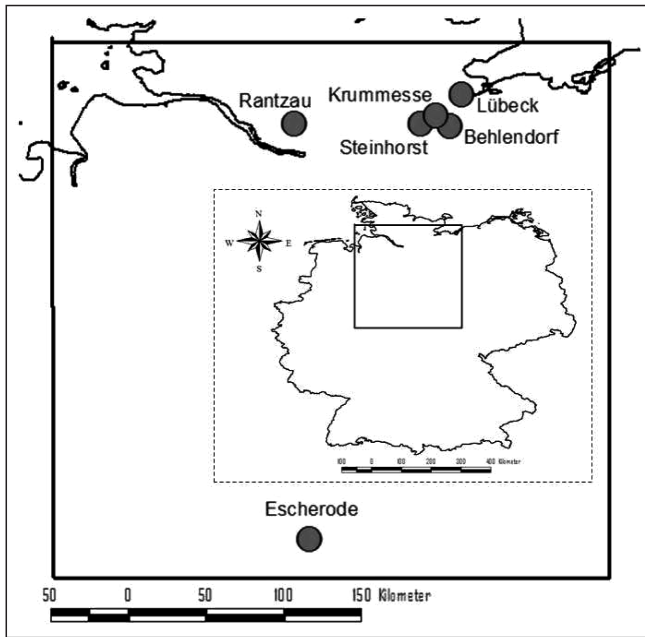
Figure 1. – Location of six sampled oak stands in Northern Germany.

Microsatellites

DNA extraction

Total DNA was isolated from the buds after removing the scales. Five to six buds were prefrozen in liquid nitrogen and the DNA extracted according to the minipreparation of DUMOLIN et al. (1995) with slight modifications, including a final treatment with 0.5 µg RNaseA (Boehringer Mannheim, Germany) at 37 °C for 30 min.

PCR amplification

PCR amplification was carried out in a total volume of 25 µL containing about 20 ng of template DNA, 2.5 MM MgCl$_2$, 100 µM of each dNTP, 0.2 µM of each primer and 0.25 U of *Taq* polymerase with the respective 1x PCR buffer (*Taq* polymerase and 10 x PCR buffer were purchased from Eurogentec, Ougree, Belgium), following the cycle profile described by STREIFF et al. (1998). PCR was run in the TouchDown™ Thermal System (Hybaid Limited, Teddington, UK).

Separation and staining of PCR products

The PCR products were pretreated according to STREIFF et al. (1998) and run in a 6% denaturing poly-

acrylamide gel (Rotiphor 40, 38:2 alcrylamide:bisacrylamide, Pharmacia, Freiburg, Germany), using sequencing gel apparatus (52 Gibco BRL, Life Technologies, Eggenstein, Germany). The gels were run in 1xTBE buffer adjusted to pH 8.3 at 2000 V for 2.5 to 3 h. Silver staining of the gels was performed according to STREIFF et al. (1998). A standard was constructed using the DNA of probes with different alleles. This standard was run every 5 or 10 lanes on each gel. The allelic variation of each sample was assessed by comparison to this standard.

Microsatellite analysis

Six microsatellite loci were analysed and the sequence information of the relevant primer pairs was taken from DOW et al. (1995) and STEINKELLNER et al. (1997). According to DOW et al. (1995) the loci are coded MSQ4 and MSQ13 and following STEINKELLNER et al. (1997) the other loci are coded AG36, AG1/5, AG9 and AG1O4. All six are characterized by variable numbers of (AG)-repeats. Furthermore they follow simple codominant Mendelian inheritance. Mapping showed that the loci AG36 and AG1O4 belong to the same linkage group (BARRENECHE et al., 2004), also AG9 and AG1/5 were mapped on the same linkage group but the mapped distance is rather large. In the present study, however, effects of associations among loci on population differentiation will not be analyzed further as explained later on.

Allozymes

For allozyme analysis, crude proteins were extracted from winter buds. Extraction procedures and the composition of the electrode and gel buffers followed those of HERTEL and ZASPEL (1996). From the enzymatic systems studied seven polymorphic loci, which showed simple Mendelian inheritance in controlled crosses (ZANETTO et al., 1996) were used: *Aap-b* (alanine aminopeptidase), *Pgi-b* (phosphoglucoisomerase), *Mnr* (menadione reductase), *6pgdh-b* (6-phosphogluconate dehydrogenase), *Idh-b* (isocitrate dehydrogenase), *Pgin* (phosphoglucomutase) and *Acp-c* (acid phosphatase). *Aap-b*, *Acp-c* and *Pgin* were monomeric, *6pgdh-b*, *Idh-b* and *Pgi-b* were dimeric and *Mnr* was tetrameric.

Assessing causes of differentiation among populations

With respect to genetic differentiation among populations it is meaningful to basically distinguish homogenizing from diversifying forces. Among the diversifying forces are random drift, differential selection due to

Table 1. – Studied *Quercus robur* populations, size of the stand, age of trees, sample size (*N*) and analysed type of gene markers (allozymes and/or Microsatellites = SSRs).

Name	Size (ha)	Age	N	Allozymes	SSRs
Behlendorf	5.5	180-210	228	X	X
Steinhorst	4.0	150-259	84	X	X
Lübeck	10	166	82	X	
Krummesse	5	161	62	X	
Rantzau	10.4	148	159		X
Escherode	5	150	210		X

adaptive variation of environmental factors among populations, and non-recurrent mutation. Homogenizing forces are exerted by migration and uniformly acting selection (similar selection regimes in all populations). In addition, lack of differentiation may be due to joint descent in the absence of diversifying forces. From this perspective, it is desirable to distinguish amounts of differentiation that can typically be produced by purely randomly acting forces from those that are due to non-random forces. Obviously, each observable degree of differentiation can be realized at random. Thus one depends on consideration of frequency distributions of descriptors of differentiation as they result from randomness.

Within such distributions, the position of the observed descriptor value can be assessed by the frequency of situations (or system states) yielding larger and by the frequency yielding smaller values than the observed. Small frequencies of exceeding the observed differentiation indicate that the observation is larger than can be expected under purely randomly acting forces, so that diversifying forces in the form of differential selection or non-recurrent mutation can be inferred to clearly override randomly acting and homogenizing forces. Non-recurrent mutation is likely to produce high degrees of genetic diversity, and by this it may be distinguished from differential selection. Conversely, if the frequency of falling below the observation is small, homogenizing forces can be inferred to be dominant. In this case, further distinction between migration and uniformly acting selection may be difficult and may require additional information such as on the type of genetic marker. Generally, however, one expects migration to be a homogenizing force that affects many loci, while uniformly acting selection affects only a comparatively small number of loci. Joint descent in the absence of significantly diversifying forces should be distinguishable from gene flow by affecting smaller numbers of gene loci, since it is unlikely that large parts of the genome remain the same over several generations.

This argument applies equally to specifically model-dependent and model-independent (non-parametric) approaches. Because of the numerous problems arising with the assumptions in model-dependent tests of selective neutrality (see e.g. the review of NIELSEN, 2001), the present analyses will rely on randomization methods as explained in the following section (and referring to the work of MANLY, 1997, for example).

The above explanations are particularly suited to studies of effects on individual loci. Efficient inclusion of multiple loci would require explicit consideration of stochastic associations (including linkage) in their specific effects on measuring degrees of genetic differentiation among populations. Since the descriptors considered in this paper are not readily amenable to the inclusion of these effects the present paper will concentrate on analyses of individual loci.

Testing randomness of differentiation

The presence of directed forces that increase or decrease differentiation among populations beyond mere random differentiation can be analyzed efficiently with the help of randomization (permutation) methods. Essentially, these methods consist in generating all possible reassignments (permutations) of individuals to populations and compute for each reassignment the differentiation among populations. This yields a distribution of δ (or F_{ST} or F'_{ST}). Within this distribution the observed value of δ (or F_{ST} or F'_{ST}) is assessed as described in the preceeding section. Recall that this assessment differs from common practice in that excessively large values give rise to different conclusions than excessively small values. While the implied causal inference applies to δ without restrictions, it is problematic with F_{ST}, since neither large nor small values of F_{ST} are consistent indicators of differentiation. F'_{ST} is less concerned by this problem (see Appendix).

This model-independent (non-parametric) approach differs from a model-dependent approach in several respects. Most basically it is conditional in the sense that the population sizes and the overall distribution of genetic types are boundary conditions as given in the samples. Random effects are restricted to population membership of individuals. As a consequence, random loss of genotypes (and even more so genes) across populations cannot take place. For the same reason, randomization can simulate effects of distribution processes only within one generation excluding sexual phases. The processes are characterized by replacements or exchanges of individuals. Hence, if differential selection is cited as evidence for large degrees of genetic differentiation, for example, this rests on the idea that individuals are selectively replaced in populations that showed no genetic differences prior to the replacement. Similarly, when inferring gene-flow this rests on a concept of non-selective exchange of individuals among populations with the effect of reducing initial differences in genetic composition.

Covariation of δ with each of F_{ST} and F'_{ST}

Even though the conceptual ambiguity in using F_{ST} as a descriptor of differentiation has been demonstrated in several analyses (and is further detailed in the Appendix), this does not exclude the possibility that over certain frequency ranges it may come close to a true descriptor of differentiation such as δ. For F'_{ST}, this can be expected to an even larger degree. In view of the large number of studies in which F_{ST} is computed and its values interpreted in terms of genetic differentiation, this is a crucial problem to address. It will be verified for the oak data presented in this paper by looking for covariation of F_{ST} and δ across the gene loci studied. For this purpose it is useful to briefly recall the general idea of covariation.

By definition, two ordinal variables X and Y show strict covariation if one variable consistently increases or consistently decreases as the other variable increases. There thus exists a strictly monotonic relationship between the two variables. Product-moment correlations have to be considered as referring to the deviation from a special form of covariation as specified by linearity. Monotonous relationships other than linearity cannot be detected with this correlation. Since the mathematical

structures of δ and F_{ST} exclude a linear relationship, it is meaningful to look for methods of detecting types of covariation that are monotonous but not necessarily linear.

Generally, a strictly monotonous relationship between two variables X and Y is realized if for all pairs i and j ($i < j$) of objects the differences $X_i - X_j$ and $Y_i - Y_j$ consistently have the same sign or consistently have different signs. The variables X and Y correspond to any pair of our descriptors δ, F_{ST} and F'_{ST}, and the objects i and j correspond to gene loci. Thus, δ_i is the value of δ at the i-th locus, for example. This condition of monotonicity is equivalent to $|\Sigma_{i<j}(X_i - X_j) \cdot (Y_i - Y_j)| = \Sigma_{i<j}|(X_i - X_j) \cdot (Y_i - Y_j)|$ up to the existence of "ties", where one of $X_i - X_j$ or $Y_i - Y_j$ is zero and the other is not. Ignoring ties for a moment, a meaningful measure of covariation is suggested by

$$C := \frac{\sum_{i<j}(X_i - X_j) \cdot (Y_i - Y_j)}{\sum_{i<j}|(X_i - X_j) \cdot (Y_i - Y_j)|},$$

which varies between –1 and +1 such that $C = 1$ for strictly positive and $C = -1$ for strictly negative covariation.

There is a close relationship between C and the product-moment correlation C_p, which is determined by the fact that $\Sigma_{i<j}(X_i - X_j) \cdot (Y_i - Y_j)$ equals n^2 times the covariance of X and Y, where n denotes the number of objects (loci). The difference between C and C_p lies in their normalizations, which equals the product $\sigma(X) \cdot \sigma(Y)$ of the standard deviations in the case of C_p. In fact from the Cauchy-Schwarz inequality one obtains $\Sigma_{i<j}|X_i - X_j| \cdot |Y_i - Y_j| \leq \sqrt{\Sigma_{i<j}(X_i - X_j)^2} \cdot \sqrt{\Sigma_{i<j}(Y_i - Y_j)^2} = \sigma(X) \cdot \sigma(Y) \cdot n^2$. Hence, $|C| \geq |C_p|$ as is suggested by the fact that general covariation includes more than just linear relationships.

C is not defined if its denominator is zero. This situation arises only if there is no variation either in X- or in Y-values, which a priori indicates the complete absence of covariation and thus justifies to set $C = 0$, if at least one of X or Y varies. Moreover, "ties" may be considered to add significantly to the measurement of covariation only if $|C| = 1$, where monotonicity is still realized, but "steps" may occur in the graph of $X \times Y$. However, because of the bidirectional concept of measures of covariation, steps occur as the absence of change in one and steep change in the other direction. It may therefore be meaningful to consider ties as a separate feature of covariation if $|C| = 1$. Both C_p and C will be included in the data analysis. In the following we will be interested in degrees of covariation between $X = F_{ST}$ and $Y = \delta$ as well as between $X = F'_{ST}$ and $Y = \delta$.

Effects of allelic diversity on δ, F_{ST} and F'_{ST}

The above-mentioned concern about analyses of differentiation based on F_{ST} for highly variable genetic markers will be taken into account by determination of the effective number of alleles v_2 across all populations for each gene locus. The index v_2 is defined in CROW and KIMURA (1970, p.323f) as a measure of allelic diversity and is identical in concept to the effective numbers of alleles n applied in the Appendix. Since v_2 can be derived from δ as a special case (GREGORIUS, 1987), the basic condition that differentiation increases with increasing resolution of the genetic trait is equally fulfilled by the descriptor v_2 of genetic diversity. It is therefore meaningful to study effects of allelic diversity or polymorphism on descriptors of genetic variation among populations with the help of the measures C_p and C of covariation of v_2 with each of δ, F_{ST} and F'_{ST}. In this way, tendencies of populations to use higher genetic diversity for higher population differentiation can be observed.

Results

Genetic variation among populations

Tables 2 and *3* summarize the observed δ-, F_{ST}- and F'_{ST}-values as well as the outcomes from the permutation analyses all of which are based on 10000 permutations. For each descriptor δ, F_{ST} and F'_{ST}, $K^l_{0.025}$ and $K^u_{0.025}$ denote the critical values, below and above which, respectively, 2.5% of the descriptor values among all 10000 permutations were located. These values correspond to the lower and upper 0.025-quantile. $P(Z > \delta)$, $P(Z > F_{ST})$ and $P(Z > F'_{ST})$ refer to the proportion of δ-values, F_{ST}-values and F'_{ST}-values, respectively, which exceed the observed value of the descriptor among 10000 permutations. They will be referred to as significance probabilities.

Covariation between F_{ST}, F'_{ST}, δ and v^2

Figure 2 illustrates the covariation of F_{ST} and F'_{ST} with δ across loci for allozymes and microsatellites separately (the coefficients of determination R^2 are also provided in the figure). The C_p- and C-values of covaria-

Table 2. – The descriptors δ, F_{ST} and F'_{ST} of genetic variation among populations, and the effective number v_2 of alleles across all populations for allozymes at 7 loci; the variables are explained in the text. In the significance probability P (Z > ζ) the value of ζ corresponds to δ, F_{ST} and F'_{ST} in the second column, respectively.

δ, F_{ST} and F'_{ST} for **Allozymes**

Locus	$\delta/F_{ST}/F'_{ST}$	$K^l_{0.025}$	$K^u_{0.025}$	P(Z > ζ)	v_2
IDH	0.0337	0.0114	0.0634	0.465	
	0.0038	0.0003	0.0139	0.459	1.68
	0.0073	0.0007	0.0260	0.455	
6-PGDH	0.0105	0.0057	0.0207	0.762	
	0.0020	0.0005	0.0133	0.731	1.06
	0.0022	0.0006	0.0145	0.731	
PGM	0.0861	0.0191	0.0871	0.029	
	0.0212	0.0006	0.0178	0.012	2.11
	0.0499	0.0015	0.0427	0.013	
PGI	0.0379	0.0133	0.0488	0.146	
	0.0076	0.0007	0.0120	0.117	1.28
	0.0104	0.0001	0.0164	0.113	
AP	0.0538	0.0312	0.0933	0.608	
	0.0037	0.0012	0.0125	0.628	2.89
	0.0129	0.0041	0.0431	0.626	
SAP-C	0.0189	0.0101	0.0635	0.846	
	0.0014	0.0004	0.0164	0.840	1.56
	0.0025	0.0006	0.0282	0.839	
MNR	0.0246	0.0135	0.0422	0.548	
	0.0037	0.0009	0.0102	0.451	1.19
	0.0046	0.0012	0.0129	0.461	

Table 3. – The descriptors δ, F_{ST} and F'_{ST} of genetic variation among populations, and the effective number v_2 of alleles across all populations for microsatellites at 6 loci; the variables are explained in the text. In the significance probability P (Z > ζ) the value of ζ corresponds to δ, F_{ST} and F'_{ST} in the second column, respectively.

δ, F_{ST} and F'_{ST} for **Microsatellites**

Locus	$\delta/F_{ST}/F'_{ST}$	$K^l_{0.025}$	$K^u_{0.025}$	P(Z > ζ)	v_2
AG1_5	0.0957	0.0639	0.1086	0.178	
	0.0031	0.0013	0.0045	0.236	6.82
	0.0265	0.0111	0.0390	0.241	
AG104	0.1629	0.1131	0.1593	0.012	
	0.0034	0.0017	0.0040	0.100	14.75
	0.0635	0.0326	0.0746	0.111	
AG36	0.1555	0.0813	0.1271	0.000	
	0.0060	0.0014	0.0045	0.002	8.68
	0.0645	0.0159	0.0493	0.002	
AG9	0.1285	0.0629	0.1103	0.000	
	0.0076	0.0013	0.0049	0.000	7.46
	0.0697	0.0123	0.0451	0.000	
MSQ13	0.1336	0.0526	0.0975	0.000	
	0.0086	0.0011	0.0055	0.001	4.11
	0.0432	0.0055	0.0280	0.000	
MSQ4	0.1504	0.0835	0.1306	0.000	
	0.0059	0.0014	0.0050	0.006	7.38
	0.0541	0.0138	0.0468	0.007	

tion between F_{ST} and δ are C_p = 0.886 and C = 0.979 for allozymes, and they are C_p = 0.111 and C = 0.144 for microsatellites. As suggested by its particular normalization covariation of F'_{ST} with δ is stronger, where C_p = 0.938 and C = 1 for allozymes, and C_p = 0.748 and C = 0.841 for microsatellites.

The degree to which δ, F_{ST} and F'_{ST} depend on the overall allelic diversity is characterized by the covariation between v_2 and each of δ, F_{ST} and F'_{ST} over all loci. For the combination $\delta \times v^2$: C_p = 0.858 and C = 0.990, for $F_{ST} \times v_2$: C_p = –0.106 and C = –0.238, and for $F'_{ST} \times v_2$: C_p = 0.786 and C = 0.962.

Discussion

Among the above results the following deserve further attention in view of the explanations in the Methods sections concerning the mathematical and conceptual properties of measures of differentiation and their interpretation with the help of permutation analyses.

– F_{ST}-values are about ten times smaller than δ-values for all gene markers. F'_{ST}-values are also smaller than δ-values, however to a lesser degree which varies among loci.

This cannot be viewed as a scaling effect, since F_{ST}, F'_{ST} and δ all vary between 0 and 1. It rather reflects the fact, that both F_{ST} and F'_{ST} fail to detect differences between populations beyond those inherent in the difference between the overall effective number of alleles and the effective number of alleles per population (see the pertaining definitions and demonstrations in the Appendix). The higher closeness of F_{ST} to δ is due to its special normalization, which reduces the dominance of the effective number of alleles per population in F_{ST} in favor of differences among populations.

– All three descriptors δ, F_{ST} and F'_{ST} yield consistently small values over all allozyme and microsatellite loci.

The gene markers can be assumed to be well distributed over the genome (at least for the allozyme and microsatellite loci), and there are no indications of recent joint history of all of the six oak populations (no artificial planting from a common source). Moreover, the fact that post glacial recolonization of Northern Germany by oak dates back about 5000–7000 years (HASEL and SCHWARTZ, 2002) leaves plenty of time for genetic differentiation among populations (at least 100 generations given a generation time of about 50 years). The small descriptor values could therefore be taken as an indication of sizable amounts of gene flow among all populations (as represented by the samples). However, the upper as well as the lower critical values (quantils) $K^u_{0.025}$ and $K^l_{0.025}$ are small for all loci and for both δ and F_{ST}. Hence, under the boundary conditions specified by the overall genetic variation across all population samples, by the sample sizes and by the number of populations, large degrees of differentiation are unlikely from the start. If gene flow is nevertheless suggested to have caused the low absolute amounts of differentiation, it must be held responsible for creating the overall genetic frequency distribution. This, however, is very difficult to argue for any type of gene flow mechanism. The low degrees of differentiation are therefore more likely to be a mathematical artifact of the boundary conditions as revealed by the permutation analysis. Consequently, gene flow can so far not be proposed as a major force producing the observed differentiation patterns.

To further analyze the differentiation pattern under its boundary conditions, the significance probabilities can be drawn upon. In this context it has to be recalled that small significance probabilities indicate relatively (given the boundary conditions) large differentiation as are expected from non-randomly acting diversifying forces. Conversely, large significance probabilities indicate relatively small differentiation as expected from non-randomly acting homogenizing forces. This gives prominence to the following observations:

– Among allozyme loci, significance probabilities vary from close to significantly large (at the SAP-C locus) to significantly small (at the PGM locus) values for all three descriptors δ, F_{ST} and F'_{ST}.

This suggests that non-randomly acting diversifying forces affect different allozymes differently including random effects at most loci and tendencies towards directed diversification and homogenization at other loci.

– Among microsatellite loci the significance probabilities for all three descriptors are consistently small, and they are highly significant with the exception of one locus. Moreover, the values of all three descriptors are smaller at allozyme than at microsatellite loci, although with distinctly larger differences between both marker types for δ.

Apparently, diversifying forces act more strongly and consistently on microsatellites than on allozymes, where δ displays this effect most clearly. Microsatellites are known for their "slippage" effects that constitute a form
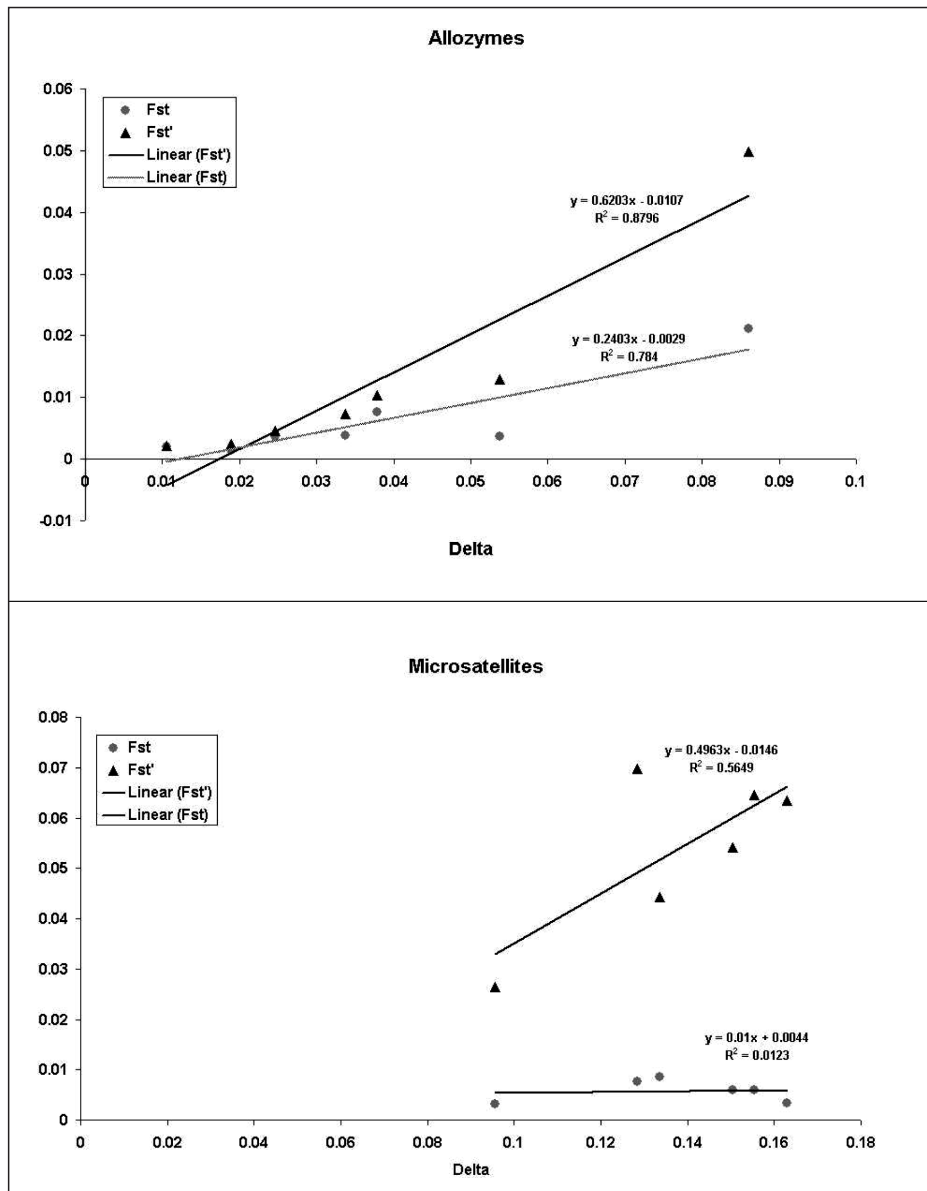
Figure 2. – Regression of F_{ST} and F'_{ST} on δ for allozymes and microsatellites.

of non-recurrent mutation. Moreover, evidence seems to accumulate that "SSRs, by virtue of their special mutational and functional qualities, have a major role in generating the genetic variation underlying adaptive evolution" (KASHI and KING, 2006). While the mutational system alone constitutes a major diversifying force that explains the consistently large differentiation across all loci, this effect would even be enhanced by the possibility for considerable adaptational differentiation inherent in regulatory genetic variation of SSRs (see the review of KASHI and KING, 2006). This could also be a consequence of the larger distances among the stands scored for microsatellites (see *Table 1* and *Figure 1*).

This observation contrasts clearly with that for allozyme loci, where two loci show high though not significant, two loci show low (but only one significant), and the remaining loci show intermediate significance probabilities for all three descriptors. Hence, at two loci uniformly acting selection may have had a sizable share,

at one locus selective diversification explains the differentiation pattern, and at two loci the differentiation patterns are more likely to be dominated by random events in the sense of a balanced mixture of diversifying and homogenizing forces. Under the boundary conditions gene flow can thus again not be argued to have played a significant role in the creation of the observed differentiation patterns. Instead, mutation type and variable selection regimes appear to be the dominant forces.

– The (sample) correlation between δ and F_{ST} across loci is strong (C_p = 0.886) for allozymes and weak (C_p = 0.111) for microsatellites. The correlation between δ and F'_{ST} is strong for both trait categories (up to complete covariation C = 1 for allozymes) but again distinctly weaker for the highly polymorphic microsatellites.

Even though these correlations are based on generally small values for both F_{ST} and δ, they reflect true associations, since the ranges of variation for both descriptors

are strongly limited by the boundary conditions. The low correlation for microsatellites is not just an artifact of non-linearity as is confirmed by the degrees C of covariation ($C = 0.979$ for allozymes, $C = 0.144$ for microsatellites). Hence, many of the numerous F_{ST} studies based on allozyme or related marker types may be acceptable as studies of population differentiation. Microsatellites are different, however. The poor covariation found for this marker type is in accordance with the fact that for higher degrees of polymorphism, F_{ST} is dominated by the effective genetic variation within populations and thus loses sensitivity towards differences among populations. Since this particular effect is diminished by the special normalization of F'_{ST}, its covariation with δ is distinctly stronger for this highly polymorhic marker type (compare the demonstrations in the Appendix). Yet, in accordance with F_{ST}, the covariation increases with the lower degrees of polymorphism at the allozyme loci. This discrepancy did not lead to contradictory conclusions in the above causal analysis because of the comparatively small differences among population samples implied by their boundary conditions.

– The correlation of the overall effective number of alleles is strongest with δ, still strong but weaker with F'_{ST}, and even negative with F_{ST}.

This is the probably most impressive demonstration of disobeyance of the basic resolution criterion of differentiation and diversity (see Appendix). Particularly the near strict covariation ($C = 0.99$) of δ with the effective number of alleles demonstrates the high sensitivity of this descriptor towards the differentiating capacity of genetic polymorphism. At the other extreme we have the particularly counter-intuitive result for F_{ST} that suggests no relationship between differentiation and polymorphism with a slight tendency towards decreasing differentiation with increasing polymorphism.

Some of the above less familiar aspects were observed in earlier experimental work on genetic differentiation. In a study using several nuclear RAPD and allozyme loci (besides mitochondrial and chloroplast DNA-markers) among pine populations, LATTA and MITTON (1997) also found consistently small values of F_{ST} for all allozyme loci. For the RAPD markers, however, F_{ST} values varied from very small to very large among loci, where at two loci all populations with the excpetion of one were close to fixation for the same allele. The authors discuss these values in connection with historical division and diversifying selection. Yet, at least the latter suggestion is difficult to maintain, when considering the fact that the large F_{ST} values were due to almost fixation of the same allele and thus to effectively small genetic differentation among populations. As shown in the Appendix, F'_{ST} values may also be high in this case and would thus, in contrast with δ, not have revealed the genetic similarity either. Unfortunately, further analysis of the small F_{ST} values at the allozyme loci was not performed in the work of LATTA and MITTON.

In summary, the present observations strongly suggest that in studies of population differentiation the boundary conditions specified by the overall genetic variation across all population samples, by the population (sample) sizes, and by the number of populations should be taken into consideration. These boundary conditions may set bounds to the degree of differentiation, which can be estimated by permutation analysis, and which may essentially alter conclusions about the significance of the basic factors that prevent or promote population divergence. In particular, the widely favored reasoning that the small degrees of differentiation observed at all loci should indicate sizable amounts of gene flow among our oak stands is not supported by the present analysis. Since, for central European tree species, our choice of populations is not special in any sense, the above results may give rise to reconsideration of current opinions about the prevalence of gene flow, random drift, selection and mutation in creating patterns of genetic differentiation. Such reconsideration should take advantage of δ as a basic descriptor of genetic differentiation among populations. Moreover, the comparatively high covaration between δ and F'_{ST} observed in our study is unlikely to extend to genetic traits typically showing higher degrees of genetic fixation in local populations (such as extranuclear but also particular nuclear markers). Conclusions as to the forces of genetic differentiation may in such cases diverge more significantly between δ and the F_{ST} descriptors.

References

BARRENECHE, T., M. CASASOLI, K. RUSSELL, A. AKKAK, H. MEDDOUR, C. PLOMION, F. VILLANI and A. KREMER (2004): Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). Theoretical and Applied Genetics **108**: 558–566.

CHARLESWORTH, B. (1998): Measures of Divergence Between Populations and the Effect of Forces that Reduce Variability. Mol. Biol. Evol. **15**: 538–543.

CROW, J. F. and M. KIMURA (1970): An Introduction to Population Genetics Theory. Harper and Row, New York.

DEGEN, B., R. STREIFF and B. ZIEGENHAGEN (1999): Comparative study of genetic variation and differentiation of two pedunculate oak (*Quercus robur*) stands using microsatellite and allozyme loci. Heredity **83**: 597–603.

DOW, B. D., M. V. ASHLEY and H. F. HOWE (1995): Characterization of Highly Variable (Ga/Ct) (N) Microsatellites in the Bur Oak, Quercus inacrocarpa. Theoretical and Applied Genetics **91**: 137–141.

DUMOLIN, S., B. DEMESURE and R. J. PETIT (1995): Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated by an efficient PCR method. Theoretical and Applied Genetics **91**: 1253–1256.

GREGORIUS, H.-R. (1987): The relationship between the concepts of genetic diversity and differentiation. Theor. Appl. Genetics **74**: 397–401.

GREGORIUS, H.-R. (1998): The system analytical approach to the study of hypotheses. URL http://www.uni-forst.gwdg.de/forst/fg/index.htm.

GREGORIUS, H.-R. (2002): An integrative approach to modeling mating systems of tree populations. *In:* B. DEGEN, M. D. LOVELESS, A. KREMER 2002. Modelling and Experimental Research on Genetic Processes in Tropical and Temperate Forests. Belém, PA: Embrapa Amazônia Oriental, pp. 42–68.

GREGORIUS, H.-R. and F. BERGMANN (1995): Analysis of isoenzyme genetic profiles observed in forest tree populations. Pp. 79–96. *In:* PH. BARADAT, W. T. ADAMS, G. MÜLLER-STARCK (eds.): Population Genetics and Genetic Conservation of Forest Trees. SPB Academic Publishing, Amsterdam.

GREGORIUS, H.-R. and J. H. ROBERDS (1986): Measurement of genetical differentiation among subpopulations. Theoretical and Applied Genetics **71**: 826–834.

HASEL, K. and E. SCHWARTZ (2002): Forstgeschichte. Ein Grundriss für Studium und Praxis. Kessel, Remagen.

HEDRICK, P. W. (1999): Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution **53**: 313–318.

HEDRICK, P. W. (2005): A standardized genetic differentiation measure. Evolution **59**: 1633–1638.

HERTEL, H. and I. ZASPEL (1996): Investigations on vitality and genetic structure in oak stands. Annales Des Sciences Forestieres **53**: 761–773.

HEUERTZ, M., J. F. HAUSMAN, O. J. HARDY, G. G. VENDRAMIN, N. FRASCARIA-LACOSTE and X. VEKEMANS (2004): Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). Evolution **58**: 976–988.

KASHI, Y. and D. G. KING (2006): Simple sequence repeats as advantageous mutators in evolution. TRENDS in Genetics **22**: 253–259.

KREMER, A., J. L. DUPOUEY, J. D. DEANS, J. COTTRELL, U. CSAIKL, R. FINKELDEY, S. ESPINEL, J. JENSEN, J. KLEINSCHMIT, B. VAN DAM, A. DUCOUSSO, I. FORREST, U. L. DE HEREDIA, A. J. LOWE, M. TUTKOVA, R. C. MUNRO, C. STEINHOFF and V. BADEAU (2002): Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak stands. Annals of Forest Science **59**: 777–787.

LATTA, R. G. and J. B. MITTON (1997): A comparison of population differentiation across four classes of gene marker in Limber Pine (*Pinus flexilis* James). Genetics **146**: 1153–1163.

MANLY, B. F. J. (1997): Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman & Hall, London, 399 pp.

NEI, M. (1973): Analysis of gene diversity in subdivided populations. Pro. Nat. Acad. Sci. USA **70**(12): 3321–3323.

NAGYLAKI, T. (1998): Fixation Indices in Subdivided Populations. Genetics **148**: 1325–1332.

NEIGEL, J. E. (1997): A comparison of alternative strategies for estimating gene flow from genetic markers. Annu. Rev. Ecol. Syst. **28**: 105–128.

NEIGEL, J. E. (2002): Is FST obsolete? Conservation Genetics **3**: 167–173.

NIELSEN, R. (2001): Statistical tests of selective neutrality in the age of genomics. Heredity **86**: 641–647.

SLATKIN, M. (1995): A Measure of Population Subdivision Based on Microsatellite Allele Frequencies. Genetics **139**: 457–462.

STEINKELLNER, H., S. FLUSH, C. E. TURETSHEK, C. LEXER, R. STREIFF, A. KREMER, K. BURG and J. GLÖSSL (1997): Identification and characterization of (GA/TC)n-microsatellite loci from *Quercus petraea*. Pl. Mol. Bio. **33**: 1093–1096.

STREIFF, R., T. LABBE, R. BACILIERI, H. STEINKELLNER, J. GLÖSSL and A. KREMER (1998): Withinpopulation genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. Molecular Ecology **7**: 317–328.

WEIR, B. S. and C. C. COCKERHAM (1984): Estimating F-Statistics for the analysis of population structure. Evolution **38**: 1358–1370.

WHITLOCK, M. C. and D. E. MCCAULEY (1999): Indirect measures of gene flow and migration: F_{ST} not equal $1/(4Nm + 1)$. Heredity **82**: 117–125.

WRIGHT, S. (1978): Evolution and the Genetics of Populations, Vol. 4: Variability within and among Natural Populations. The University of Chicago Press, Chicago.

ZANETTO, A., A. KREMER, G. MÜLLER-STARCK and H. H. HATTEMER (1996): Inheritance of isozymes in pedunculate oak (*Quercus robur* L.). Journal of Heredity **87**: 364–370.

Appendix: Characteristic properties of F_{ST}

In the following an attempt is made to achieve a characterization of F_{ST} that is analogous to that presented by HEDRICK (1999) but uses effective numbers of alleles in place of "heterozygosities". By this it is intended to demonstrate more explicitly the effects of allelic polymorphism on F_{ST}.

With the help of the above notation one obtains the representation

$$F_{ST} := \frac{\sum_j \sum_i p_{i,j}^2 \cdot c_j - \sum_i p_i^2}{1 - \sum_i p_i^2}$$

in its common version that considers variable population sizes (see e.g. GREGORIUS and ROBERDS, 1986). Basically, F_{ST} is not defined for the situation where all populations are fixed for the same allele (monomorhpy of the collection of populations). Otherwise $0 \leq F_{ST} \leq 1$ with $F_{ST} = 0$ if and only if all populations are genetically identical, and $F_{ST} = 1$ if and only if all populations are monomorphic (with at least one population differing from the others). Hence, F_{ST} could be considered as a measure of genetic fixation of the collection of populations. It cannot be considered as a measure of genetic differentiation among populations, since for monomorphy of the collection of populations, in which there is no differentiation among populations, it is not defined.

Notations

$p_{i,j}$: = relative frequency of the *i*-th allele in the *j*-th population.

c_j : = proportion of individuals belonging to the *j*-th population.

p_i : = $\sum_j p_{i,j} \cdot c_j$ or the relative frequency of the *i*-th allele in the total collection of populations.

n_e : = $1/\sum_i p_i^2$, or the overall *effective number of alleles* in the totality of populations.

$n_e(j)$: $= 1/\Sigma_i p_{i,j}^2$, or the *effective number of alleles* in the j-th population.

$\bar{n}_e$: $= 1/\Sigma_j c_j/n_e(j)$, or the *effective number of alleles* per population defined by the harmonic mean.

Concerning distribution of allelic variation across populations, the significance of F_{ST} can be more directly demonstrated by expressing F_{ST} in terms of the effective numbers of alleles specified in the above table of notations. Thus

$$F_{ST} = \frac{\sum_j c_j/n_e(j) - 1/n_e}{1 - 1/n_e} = \frac{1/\bar{n}_e - 1/n_e}{1 - 1/n_e} = \frac{n_e - \bar{n}_e}{\bar{n}_e(n_e - 1)},$$

which shows that F_{ST} depends solely on the overall effective number of alleles and the effective number of alleles per population. Since $(n_e - \bar{n}_e) < (n_e - 1)$, one obtains

$$F_{ST} < \frac{1}{\bar{n}_e},$$

which shows that F_{ST} may become very small for large effective numbers of alleles per population (the inequality is equivalent to (lb) in HEDRICK 1999). This includes the situation where all populations are genetically disjoint.

On the other hand, F_{ST} also approaches zero as all populations become genetically identical, irrespective of the number of genetic types involved. F_{ST} therefore may realize small values both in the presence of large similarity and large dissimilarity among the populations (c.f. HEDRICK, 1999). At the other extreme, large values of F_{ST} can be realized only for high degrees of fixation of all populations, however, with the possibility of small average differentiation as a consequence of fixation to the same genetic variant in almost all populations. Large degrees of differentiation can therefore be realized for small as well as for large values of F_{ST}.

While these deliberations show clearly that F_{ST} cannot be accepted unambiguously as a measure of differentiation, it remains to demonstrate its suitability as a measure of fixation. Since fixation refers to monomorphy of individual populations, it is meaningful to require for a measure of fixation that it decreases with increasing polymorphism within populations. By the above results, F_{ST} seems to fulfill this requirement. However, the previously indicated discontinuity of F_{ST} for overall monomorphy corrupts even this requirement. To see this, consider any set of genetically identical polymorphic populations in which one allele predominates. $F_{ST} = 0$ for these populations, and this continues as the dominant allele approaches fixation. Moreover, for genetically identical populations their degree of polymorphism does not affect F_{ST}. Hence, F_{ST} is not a consistent measure of fixation either. Such a measure could be rather of the form $1/\bar{n}_e = \Sigma_{i,j} c_j \cdot p_{i,j}^2$ since it is defined for a monomorphic set of populations and approaches a value of 0 asymptotically with increasing numbers of alleles per population and thus with increasing deviation from fixation.

Finally, it should be mentioned that any change in the distribution of the overall allelic variation (permutation of individuals) over populations does not affect F_{ST} as long as it does not change the effective number of alleles $\bar{n}_e$ per population. Hence, for given overall allelic variation, F_{ST} is insensitive towards changes in differences among populations that do not affect the average effective number of alleles per population, if there are any. Obviously, for small overall effective numbers of alleles there are fewer opportunities to distribute genetic variation over populations without changing $\bar{n}_e$, and F_{ST} may therefore be more sensitive to differences among populations for lower degrees of polymorphism.

Hedricks normalization of F_{ST}

In a recent paper, HEDRICK (2005) suggests a normalization of F_{ST} with the intention to take account of the problem of genetic disjointness among polymorphic populations. In its version for arbitrary population sizes, the normalization is determined by writing $\Sigma_i p_i^2 = \Sigma_i \Sigma_{j,k} p_{i,j} c_j \cdot p_{i,k} c_k$, so that obviously $\Sigma_i p_i^2 \geq \Sigma_{i,j}(p_{i,j} c_j)^2$ with equality between the two sides if the "off-diagonal" elements $p_{i,j} c_j \cdot p_{i,k} c_k$ equal zero for each i and $j \neq k$. This is equivalent to the statement that no individual genetic type appears in two different populations, i.e. all populations are genetically disjoint. Consequently, setting $F_{ST(max)} := (\Sigma_{i,j} p_{i,j}^2 \cdot c_j - \Sigma_{i,j}(p_{i,j} c_j)^2)/(1 - \Sigma_{i,j}(p_{i,j} c_j)^2)$, so that $F_{ST(max)} \geq F_{ST}$, Hedrick obtains for the normalized version $F'_{ST} := F_{ST}/F_{ST(max)}$ that $F'_{ST} = 1$ if all populations are genetically disjoint.

The latter does, however, not extend to monomorphic populations with polymorphism among populations. Even if all populations with the exception of one are fixed for the same allele, so that there is considerable overlap among populations, both $F_{ST} = 1$ and $F'_{ST} = 1$. This observation is due to the intended property of $F_{ST(max)}$ to be affected only by the effective number of alleles within populations. Writing $F_{ST(max)} = (1/\bar{n}_e - \Sigma_j c_j^2/n_e(j))/(1 - \Sigma_j c_j^2/n_e(j))$ this becomes obvious. For k equally sized populations one obtains $F_{ST(max)} = (k - 1)/(k \cdot \bar{n}_e - 1)$. Therefore, F'_{ST} equals F_{ST} in that differences between populations are considered only via the overall effective number of alleles and the effective number of alleles per population. As an additional effect, increasing the number of populations tends to decrease F'_{ST} even if n_e and $\bar{n}_e$ do not change.

Hence, the ambiguity of F_{ST} as a measure of differentiation or fixation is not solved by this normalization, and it cannot be solved by any normalization based on a frequency dependent factor N with $N \geq F_{ST}$. It is therefore not surprising that neither F_{ST} nor its normalized version F'_{ST}, consistently fullfil the basic criterion that with increasing resolution of a genetic marker the differentiation should (not necessarily strictly) increase. This was demonstrated above for genetically fixed populations. It is, however, also relevant if all populations are polymorphic as can be seen from the simple example of two populations with two alleles and frequencies 0.04 and 0.96 in one population and frequencies 0.67 and 0.33 in the other population. The resulting F_{ST}, F'_{ST} and δ values are 0.4333, 0.7371 and 0.63, respectively. Suppose that an increase in resolution of the first allele yields two alleles with frequencies 0.02 and 0.02 in the first population and frequencies 0.34 and 0.33 in the second population. In contradiction with the resolution criterion, the F_{ST} and F'_{ST} values now decrease to 0.2857 and 0.6244, while δ remains the same.