European Journal of Plant Pathology **102:** 385–395, 1996.

Analysis of diversity in populations of plant pathogens: the barley powdery mildew pathogen across Europe

Kaspar Müller[1], Joe M. McDermott[2], Martin S. Wolfe and Eckhard Limpert*
*Phytopathology, Institute of Plant Sciences, Swiss Federal Institute of Technology, Universitätsstr. 2, CH-8092 Zürich, Switzerland; [1]Present address: Ciba-Geigy Ltd., Crop Protection, Disease Control, R&D: Development; CH-4002 Basel, Switzerland; [2]Present address: West-East Centre for Microbial Diversity, 3650 Wesbrook Mall, Vancouver B.C. Canada V6S 2L2; *Author for correspondence (Fax: 1 632 1092)*

Accepted 18 December 1995

Key words: diversity indices, bootstrap method, genetic distance, Shannon, Simpson, dissimilarity

Abstract

Understanding population genetics and evolution within species requires recognition of variation within and between populations and the ability to distinguish between the potential causes of an observed distribution of variation. For this aim several established indices of diversity, and a novel one, were applied to population samples of the barley powdery mildew pathogen, *Erysiphe graminis* f. sp. *hordei*. Random spore samples were obtained from the air along transects through regions of interest across large parts of Europe in 1990. Significant geographical differences in diversity of virulence genotypes occurred among regional sub-samples. Diversity was highest in the samples from eastern Germany, Denmark and Austria, whereas the lowest values were found in the samples from Italy, southern France and parts of western Germany. Diversity in the pathogen population was generally related to the degree of diversification of host resistance in time and space, although there was considerable variation in ranking among different measures of diversity. Sensitivity to sample size proved to be the major problem with the use of several established indices of diversity. Working with very large sample sizes we used multiple random subsamples of various smaller sizes to determine how the mean index values changed with changing sample size. The Shannon index proved to be considerably affected by sample size, in contrast to the Simpson index that was therefore used as a global measure of diversity. Limits of confidence were estimated for the Simpson index using the bootstrap method of numerical resampling. The two aspects contributing to global diversity, richness and evenness, were considered separately to allow meaningful interpretation of the Simpson index. Random sub-sampling was used to reduce the influence of sample size differences for these measures. Dissimilarity, a novel measure of diversity for use in plant pathology, indicates the average number of major genes for host resistance against which pathotypes in a sample respond differently. It is thus able to account for the genetic relationship among pathotypes, which is not considered by any other index. The approaches developed in this study help to compare major forces driving evolution of large-scale populations of the barley mildew pathogen.

Introduction

There has been a general increase in interest in the geographical distribution and structure of plant pathogen populations. In particular this has focused on the processes involved including host selection, pathogen mutation, recombination and migration. Considerable progress has been made recently with analysis of variation of virulence genotypes in the cereal mildew pathogens in Europe [Wolfe and Limpert, 1987; Jørgensen, 1991; Zeller and Fischbeck, 1992]. However, in such studies on population genetics and dynamics, the diversity of populations has been considered little, so far, although this aspect seems to be particularly important for a population as diverse as that of the barley mildew pathogen across Europe.

Concepts have been proposed on how to use mathematical indices for a quantitative description of

diversity in samples of plant pathogens [e.g. Lebeda, 1982; Groth and Roelfs, 1987]. These indices were applied to several pathosystems for comparing variation among populations [Limpert et al., 1990; Leonard and Leath, 1990; Kolmer, 1991], for estimating the effect of recombination on population variability [Roelfs and Groth, 1980; Welz and Kranz, 1987] and for assessing the influence of migration or gene flow among populations [Andrivon and Limpert, 1992].

The theoretical basis and interpretation of diversity indices have been controversial topics in ecology, from where these measures were adopted [see Hurlbert, 1971; Ludwig and Reynolds, 1988; Magurran, 1988]. Three key problems have to be considered: (i) many indices may be strongly affected by the size of the samples analysed; (ii) the most popular indices (the Shannon- and the Simpson index) are 'global' measures (i.e. diversity is expressed as a single numerical value). Such global indices combine and, hence, confound the influence of their two distinct aspects: 'richness', the number of different types in a sample in relation to sample size and 'evenness', the dominance patterns among these types; (iii) diversity indices do not account for similarities or dissimilarities among the elements in a sample. Here we develop an approach for describing diversity of virulence genotypes (or haplotypes, multilocus phenotypes) in European pathogen samples of barley powdery mildew, based on a range of indices.

Materials and methods

Data collection

Random samples of barley powdery mildew isolates were obtained throughout Europe (Figure 1) in 1990 by trapping spores from the aerial population. The rapid procedure of Limpert et al. [1990] was used, with slight modifications [Müller, 1993], for sampling and analysing 3624 single colony progenies. Following the dotted lines in the figure, barley leaf segments were exposed in the jet spore trap on a car roof for distances of approximately 100 km. The test batches were incubated to provide the single colony population samples. Virulence reactions were assessed on a differential set containing twelve resistance genes or gene combinations (Table 1). Frequency data of pathotypes identified in this way were used for diversity investigations.

Table 1. Differential set of barley lines used for the determination of virulence pathotypes in the European barley mildew pathosystem

Differential no.	Name	Resistance genes
1	P-21[a]	Mlg + (CP)
2	P-03[a]	Mla6 + Mla14
3	P-23[a]	Ml(La)
4	P-4B[a]	Mla7 + (No3)
5	P-16[a]	Mlk
6	P-10[a]	Mla12 + (Em2)
7	P-01[a]	Mla1
8	P-8B[a]	Mla9
9	P-02[a]	Mla3
10	P-11[a]	Mla13 + (Ru3)
11	Triumph	Mla7 + (Ab)
12	Alexis[b]	mlo

[a] Near-isogenic lines (Kølster et al., 1986). Host – pathogen interactions: scoring of infection types 0–IV and colony density.
[b] No differentiation; only avirulent interactions observed.

Data analysis

The logarithmic series model [Fisher et al., 1943] was used to describe mathematically the relationship between the number of different pathotypes and their frequency in the overall sample. The logseries model takes the form:

$$\alpha x, \alpha x^2/2, \alpha x^3/3, \ldots \alpha x^n/n$$

where αx = the number of pathotypes predicted by the model to occur with one individual isolate in the sample, $\alpha x^2/2$ = the number with two isolates and so on. The detailed procedure for the calculation of α and x is given by Poole [1974] or Magurran [1988].

Two measures of global diversity have been widely used in ecology and were also applied to describe variability in samples of plant pathogens: one is the Shannon index [Shannon and Weaver, 1949]:

$$H = -\sum_{i=1}^{s}(n_i/n) \cdot \ln(n_i/n)$$

The second is Simpson's [1949] estimate for the 'measure of concentration':

$$\lambda = \sum_{i=1}^{s} n_i(n_i - 1)/n(n-1)$$

where n_i = the number of isolates of the i-th pathotype, n = the sample size and s = the number of different

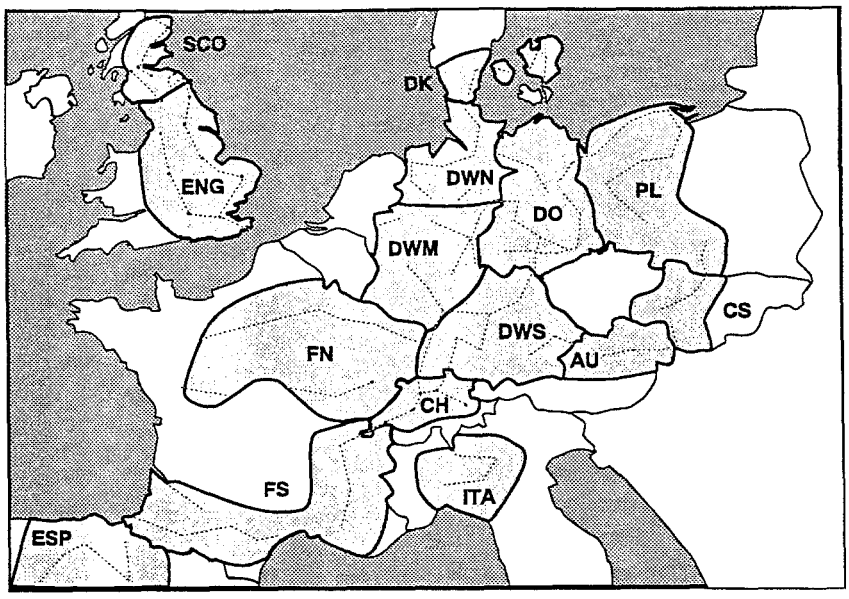

Fig. 1. Spore sampling areas in 1990. SCO = Scotland, ENG = England, FN = northern France, FS = southern France, ESP = northern Spain, DK = Denmark, DWN = northwestern Germany, DWM = western Germany, DWS = southern Germany, DO = eastern Germany, PL = western Poland, CH = Switzerland, CS = Czechoslovakia, AU = lower Austria, ITA = northern Italy. Dotted lines represent sampling routes.

pathotypes in the sample. The complement $1 - \lambda$ was used in the present study as the Simpson index of diversity. The effect of sample size on different measures of diversity was determined for the present sample by taking the means of one hundred random sub-samples with replacement from the total set of isolates where the actual sample size varied (Figures 3 and 4).

For analysing regional differentiation, the overall area was divided into 15 sub-areas considering within-region homogeneity and comparable geographic size and continuity (Figure 1). Homogeneity within regions was determined in a separate study [Müller, 1993] based on Rogers distance [Rogers, 1972] and Nei's standard genetic distance [Nei, 1987; Leonard and Leath, 1990] between sampling units. Diversity in the regional samples was characterised by a global measure (the Simpson index) and three features of diversity including richness, evenness and a novel measure, dissimilarity). A numerical estimate of the standard error of the Simpson index was calculated from 600 bootstrap samples [Efron and Tibshirani, 1991].

Genetic richness d (the number of different types divided by the total number of individuals in a sample) and genetic evenness E_5 (the relative abundance of different types in the sample [Ludwig and Reynolds, 1988]) were calculated by random sub-sampling. Thus, they represent theoretical values for sub-samples of a standard size (n_{sub}), here: n_{sub} equals the size of the smallest sample.

$$d = \frac{s_{sub}}{n_{sub}} \qquad E_5 = \frac{(1/\lambda_{sub}) - 1}{e^{H_{sub}} - 1}$$

where s_{sub} = the number of different pathotypes, λ_{sub} = the complement of the Simpson index and H_{sub} = the Shannon index in random sub-samples of n_{sub} isolates, here: $n_{sub} = 85$.

1000 random sub-samples each containing n_{sub} isolates were drawn with replacement from every test sample to obtain the mean sub-sample estimate of d and E_5. By this procedure, numerical estimates were computed exclusively from the sample data without any assumptions about the unknown underlying distribution in the entire population.

The mean dissimilarity (D) of a sample is determined as follows. The reaction of a pathotype on a differential host is regarded as an alternative response with V indicating virulence and A avirulence. For example, from sample a the four pathotypes a_1 through a_4 (Table 2) are characterized on a set of four differentials representing different genes or alleles or, in general, characters. The number of pairwise comparisons among n genotypes is $n(n - 1)/2$. In the above case there are six pairs of genotypes possible. The number of different characters $d(a_i, a_j)$ for each pair a_i, a_j

Table 2. Dissimilarity analysis. The example is based on four pathotypes (1). The six pairwise comparisons possible within sample a and their dissimilarity (2) lead to a mean dissimilarity D_a of 2.33

1. pathotypes		2. dissimilarity				
a_1	V A A A	a_1	-			
a_2	A V A A	a_2	2	-		
a_3	V V A A	a_3	1	1	-	
a_4	A A V V	a_4	3	3	4	-
			a_1	a_2	a_3	a_4

is given in Table 2, as well as the mean dissimilarity D_a, according to the formula:

$$D = \frac{\Sigma_{i<j} d(a_i, a_j)}{n(n-1)/2}$$

Means of the samples and their standard deviations are determined. The definition of complexity is given in Table 3.

Results

Pathotype abundance

With the exception of the differential variety 'Alexis' (resistance mlo) where only avirulence was found, both virulent and avirulent isolates occurred for all differentials tested. Thus the test-set could detect a maximum of 2048 (2^{11}) pathotypes; of these 525 were observed within the total sample of 3624 isolates.

The ranked abundance list (Figure 2) was used to summarise the abundance data of the observed pathotypes in the total European sample. The frequency of each pathotype in the sample is plotted in rank order from the most abundant (4.7%) to the least abundant (0.03%, equivalent to a single occurrence in the sample). The ten most abundant pathotypes comprised 23.7% of the total sample, with 71% of the isolates represented among the 100 most common pathotypes. This distribution corresponds to a frequently observed phenomenon [Pielou, 1977]: types represented by one individual are numerous, while those with successively more representatives are progressively less numerous. To quantify this abundance pattern, the logarithmic series model was fitted to the observed frequency distribution (Figure 2). The model constants were determined as α = 168.65 and x = 0.956, where α is a measure of diversity and x is related to the sample size [Fisher et al., 1943; Poole, 1974; Magurran, 1988].

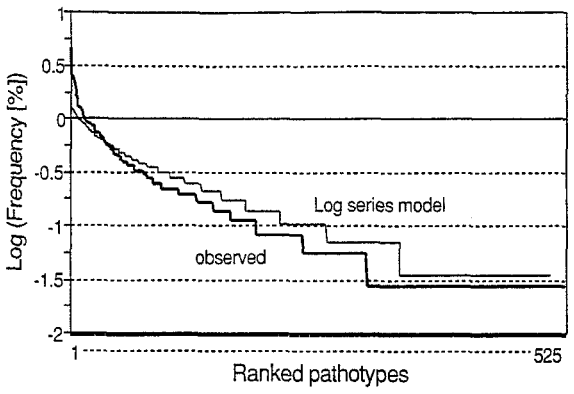

Fig. 2. Observed frequency distribution of virulence phenotypes of *E. graminis* f. sp. *hordei*, collected in western and central Europe in spring 1990 (bold line). The expected abundance distribution (logarithmic series, treated as a continuous function) is shown as a normal line. (Spearman's rank correlation coefficient r = 0.859, significant at p = 0.001). Model constants: x = 0.956, α = 168.65.

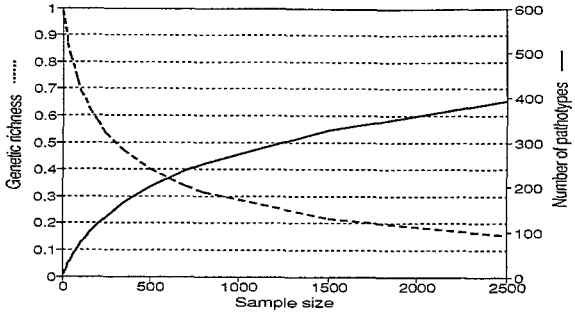

Fig. 3. Relationship between sample size and genetic richness in the sample of *E. graminis* f. sp. *hordei* isolates from Europe in 1990. The effect of simple size was simulated by random sub-sampling, a variant of numerical resampling.

Influence of sample size on diversity

An estimate for the relation between the sample size and the genetic richness in the European collection of barley mildew isolates was calculated by random sub-sampling (Figure 3). The number of pathotypes detected is strongly dependent on sample size. By testing a European sample of 500 isolates the expected number of different pathotypes is about 200, whereas about 2500 isolates have to be tested to identify 400 different pathotypes.

The Shannon index (H) also changes with sample size. With small samples such changes in H may be large (Figure 4). As sample sizes increase, H approaches asymptotically some maximum value. This is because this index is related logarithmically to the number of pathotypes in the collection, and partly because the effect of detecting new pathotypes is off-

Table 3. Diversity of the barley mildew pathogen across Europe in 1990. Regions are ordered according to the ranking of values of the Simpson index

Region[a]	n	Simpson index	Richness[d]	Evenness[d]	Dissimilarity[e]	Complexity[g]
DO	443	0.990[b]±0.002[c]	0.69	1.71	4.68±0.04[f]	4.66
DK	185	0.987 ±0.006	0.59	1.33	4.43±0.07	5.29
A	122	0.986 ±0.010	0.54	1.25	4.31±0.08	4.24
FN	183	0.984 ±0.006	0.53	1.27	3.75±0.07	3.58
ESP	139	0.982 ±0.007	0.48	1.23	3.56±0.08	3.27
CH	218	0.976 ±0.009	0.52	1.06	3.99±0.07	4.00
DWN	355	0.975 ±0.006	0.53	1.10	4.40±0.05	5.32
SCO	124	0.975 ±0.010	0.43	1.09	3.80±0.11	3.73
ENG	310	0.975 ±0.006	0.52	1.08	4.20±0.06	4.37
PL	314	0.972 ±0.008	0.53	0.98	4.27±0.05	4.70
DWS	341	0.972 ±0.008	0.52	1.00	4.11±0.06	4.02
CS	85	0.966 ±0.021	0.66	0.73	4.26±0.13	4.96
DWM	262	0.956 ±0.016	0.49	0.77	4.24±0.07	3.86
FS	303	0.956 ±0.010	0.42	0.87	3.89±0.08	2.97
ITA	99	0.946 ±0.028	0.36	0.78	3.54±0.14	1.90
EURO	3624	0.990 ±0.0003	0.70	1.57	4.93±0.01	4.16

[a] Sampling regions: see Figure 1. EURO = total European sample, SCO = Scotland, ENG = England, FN = northern France, FS = southern France, ESP = northern Spain, DK = Denmark, DWN = north-western Germany, DWM = western Germany, DWS = southern Germany, DO = eastern Germany, PL = western Poland, CH = Switzerland, CS = Czechoslovakia, AU = lower Austria, ITA = northern Italy.
[b] $1 - \lambda$ (values in the original sample).
[c] Estimate for the 95%-confidence intervals, calculated from 600 bootstrap samples.
[d] Numerical estimate for subsamples of 85 isolates, obtained by random sub-sampling.
[e] Mean number of virulence loci differences between the different pathotypes in the sample.
[f] $3\ S_x$ (S_x = standard error of the mean dissimilarity).
[g] Mean number of virulence loci per isolate.
n = number of isolates.

set by the addition of more individuals of the common types. In the present sample, the value of H changes little only for sample sizes larger than about 800 isolates. In contrast to genetic richness and the Shannon index, the Simpson index seems to be virtually independent of sample size (Figure 4).

Regional differentiation

Diversity was greatest in the sample from eastern Germany. The Simpson index for this sample was significantly different from all other samples except for those from Denmark and Austria (Table 3). The high global value in the sample from eastern Germany is caused by both the greatest richness and most evenness compared with the other regional samples. The combination of the least genetic richness and evenness of pathotypes in the Italian sample led to the lowest global diversity among the regions.

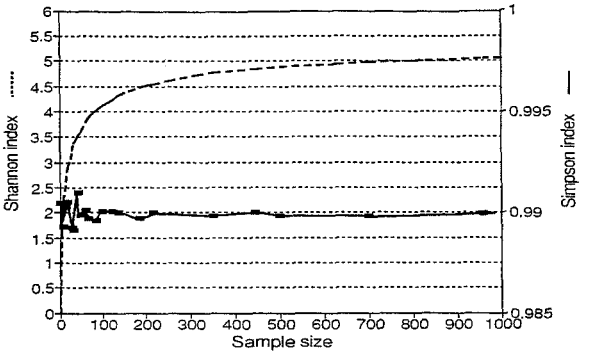

Fig. 4. Relationship between sample size and two measures of global diversity in the sample of *E. graminis* f. sp. *hordei* virulence pathotypes from Europe in 1990. The effect of sample sizes on the Shannon and Simpson index ($1 - \lambda$) was simulated by random sub-sampling. Note that the scale for the Simpson index is very fine.

Similar values for the Simpson index did not indicate similar population structures. For example, the

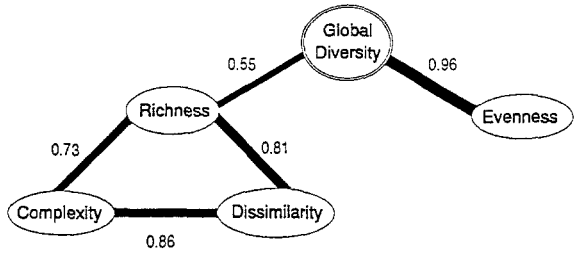

Fig. 5. Correlation between global diversity (Simpson index) and different features of diversity (evenness, genetic richness and dissimilarity of pathotypes) in samples of *E. graminis* f. sp. *hordei* virulence pathotypes from Europe in 1990. Spearman's rank correlation, calculated from data in Tables 3, $p < 0.01$.

Simpson index for the Spanish and Czechoslovakian samples did not differ significantly but their structures were quite different (Table 3). The lowest value for evenness occurred in the Czechoslovakian sample, although the values for pathotype richness and dissimilarity were high. This pattern contrasts with that for the Spanish where evenness was high but richness and dissimilarity were low. It is worth noting that the difference in dissimilarity among the two samples was highly significant.

Overall, both richness and evenness were positively correlated with the Simpson index, although the correlation was closer for evenness. A close correlation was observed particularly for dissimilarity and complexity (Figure 5). A large portion of the total variation was detected at the regional level. The diversity in the regional sample from eastern Germany, for example, was close to that for the whole European sample (Table 3) and, for evenness, the sample from eastern Germany even had the highest value.

Discussion

The large data set from a monitoring programme for the barley mildew pathogen throughout Europe in 1990 [Müller, 1993] was used to analyse the specific virulence characteristics of this pathosystem. Novel methods of computer intensive statistics were applied to calculate statistical significance of the values. Diversity indices obtained in this way allow comparisons among the regions investigated with regard to variability in their pathogen populations and contribute to the development of hypotheses on the large scale population biology of the barley mildew pathogen.

Diversity values and sample size

Comparisons of diversity in space and time should be based on representative samples of the populations investigated. Sampling the air-spora along transects through regions of interest allows random spore samples of *E. g.* f. sp. *hordei* populations [Limpert, 1987; Müller, 1993] to be obtained. Depending on local conditions, differences in sample size can occur that may cause problems because increasing sample size leads to the detection of rarer types in the population (Figure 3). Diversity values therefore may depend on sample size. This makes analyses of the powdery mildew pathosystem more difficult than those of less complex systems, such as stem or leaf rust of wheat, where sample size appears to have less effect on diversity indices [Groth and Roelfs, 1987a; Kolmer, 1991].

For summarising global diversity, the Simpson index $(1-\lambda)$ proved to be more useful than the Shannon index because it was virtually independent of sample size. In agreement with Simpson [1949]; Hurlbert [1971]; Roelfs and Groth [1980] and Groth and Roelfs [1987], $1-\lambda$ can therefore be considered as an unbiased estimate of diversity in a population sample.

Confidence intervals for the Simpson index may be directly estimated for samples of different size using the bootstrap procedure. However, this procedure should be used with care because, after resampling, the mean diversity tends to be below the value of the original sample. In our data the difference was mostly one to two standard deviations. This is related to the shape of the frequency distribution of pathotypes generally resembling a logseries model as in Figure 2. That means that the number of pathotypes occurring only once in a sample is the largest category. This group, together with others of rare pathotypes, will be most reduced after resampling, thus changing the value of $1-\lambda$.

The Shannon index was extremely sensitive to sample size, especially for samples of less than about 500 isolates. Insensitivity to sample size and the calculation of standard errors for the Shannon index as described by Poole [1974], requires that all types in the population considered are represented in the sample, which is rarely possible for most biological systems. This contrasts with the judgement of the Shannon index given by Lebeda [1982] or Groth and Roelfs [1987], to which other workers referred [Welz and Kranz, 1987; Limpert et al., 1990; Andrivon and Limpert, 1992]. It is therefore probable that some of the variation in pathotype diversity in *E. g.* f. sp. *hordei* reported

by these authors resulted from differences in sample size, although the main conclusions drawn from these studies seem still to be valid. Effects of sample size on estimations of diversity seem also to be evident in samples of *Cochliobolus carbonum* [Leonard and Leath, 1990]. Similarly, the mere difference in the size of samples of *Phytophthora infestans* [Drenth, pers.comm.] may partly account for the increase of diversity described in populations of this pathogen during the recent past [Fry et al., 1992].

Interpretation of diversity values

Since a given value of some global diversity index may result from very different combinations of richness and evenness, interpretation and correct usage of global indices have led to many critical and cautionary remarks [Groth and Roelfs, 1987a; Poole, 1974; Hurlbert, 1971; Magurran, 1988]. For interpreting global diversity it is essential to assess the relative importance of its main aspects. However, the two properties, pathotype richness and evenness are again highly sensitive to sample size [Hurlbert, 1971]. One way to avoid this problem is to have constant sample size which may, however, be difficult to realize. The expected numbers and frequencies of pathotypes in sub-samples of equal size can be computed from different samples using an underlying model that applies to all populations studied [e.g. 'rarefraction-method', Ludwig and Reynolds, 1988]. However, the manner in which diversity changes with sample size varies according to the abundance distribution of specific pathotypes in different populations. When using random sub-sampling, as proposed in this study, no assumptions need to be made about the unknown underlying probability distributions in the entire populations.

It should be noticed that diversity values of the total European sample did not correspond to the means of the regional sub-samples (Table 3). This shows that pooling of sub-samples has an additive effect on diversity of the combined sample, because combined samples represent the results of a variable number of sub-populations which have, in general, interacted little with one another. Thus, pathotype diversity in a sample changes not only with sample size but also when the sampling area varies (or when cultivars in different areas are differently susceptible). Samples compared should therefore ideally represent populations with similar geographical dimensions or dimensions of corresponding host areas. Such conditions can only be partly fulfilled in practice.

Dissimilarity – a novel index and its value

The dissimilarity parameter allows determination of an additional dimension of diversity, quantifying the genetic relationship of the genotypes occurring in a sample (Table 2). Its significance becomes obvious with an example, considering two samples with three genotypes, identified by virulence (V) or avirulence (A) at four loci. Both samples contain pathotypes AAAA and VVAA; sample 1, in addition VAAA, sample 2 VVVV instead. The dissimilarity of sample 1 is 1.33, that of sample 2 is 2.66, pointing to the obvious genetic differences present in both samples. For any established index, in contrast, the diversity of both samples would be the same.

A comparison of the Simpson index with dissimilarity reveals some interesting points. With the established index, the ranking of the sample from Spain is much in contrast to that of Italy. However, the populations in these areas have two aspects in common: they are particularly homogeneous and differ considerably from populations in the rest of Europe, which can be explained by the geographic barriers of mountain chains [Limpert et al., 1990, 1991; Müller et al., 1992]. Thus, a low dissimilarity would be expected and, indeed, the lowest values observed were from these areas (Table 3). In contrast, dissimilarity was highest in Eastern Germany, where due to immigration from neighbouring countries and selection by new resistance genes and variety mixtures the population was described to be extremely variable and heterogeneous [Limpert et al., 1991]. Concluding from both the underlying concept and the results obtained, dissimilarity seems to be well suited to measure the genetic distance among the pathotypes in a sample.

Regional differentiation and population biology

The observed frequency distribution of pathotypes of the entire European sample accords well with the logseries model (Figure 2). In general, however, models should be interpreted with caution because different models derived from sometimes conflicting hypotheses can be fitted to the same observed data set [Hurlbert, 1971; Ludwig and Reynolds, 1988; Poole, 1974]. We used the logarithmic series model only for summarising the observed distribution pattern, but it seems worth to be considered also for regional samples.

The European population of *E. g.* f. sp. *hordei* is characterised by an extraordinarily large genetic variability, even though only a rather limited number of resistance genes have been used in the European barley cultivars [Jørgensen, 1991]. The following factors may be crucial for the regional differences in diversity observed.

Selection. The virulence structure of the pathogen population is principally a function of the structure of the host population in terms of the number of resistance genes and their dynamics. These factors will tend to affect the range of pathotypes, their complexity and their (dis)similarity in the same direction, thus explaining the correlations (Figure 5) among these aspects.

An example illustrates the potential increase of parameter values when progressively more resistance genes are introduced into a model-system (Table 4). The general difference between populations from the north and the south of the area investigated can be explained in this way. For example, the high values of complexity and dissimilarity in the Danish sample seem to be due to the extensive use and rapid turnover of resistance genes [Munk et al., 1991]. The opposite holds true for Italy, southern France and Spain where very few resistance genes have been used [Wolfe and Limpert, 1987].

We may expect an unusual distribution of pathotypes in a region in which a resistance gene has been recently overcome. In Czechoslovakia, following the demise of *Mla13*, the population was characterised by extremely low evenness and high richness, indicating the varied success of a newly selected range of pathotypes. The high complexity in this sample may be, in addition to selection, due to the positive gametic disequilibria between *Va13*, *Va9* and *Va7* [Østergård and Hovmøller, 1991; Limpert et al., 1991; Müller, 1993] that was also expressed in the populations from Denmark, Poland and northern Germany.

Response to the use of variety mixtures. Selection exerted by variety mixtures differs from that in pure stands, and it is important to try to assess the impact of mixtures on the pathogen population [Huang et al., 1995; Wolfe et al., 1992]. The sample from eastern Germany came from an area with intensive use of mixtures in the spring crop for several years and was characterized by the highest values for, dissimilarity, evenness, richness and the Simpson index (Table 3). Most unusual is the big and highly significant difference in evenness and dissimilarity compared with other regions. The only other region in which mixtures have played a significant role, though considerably less so, is Denmark, which was also characterized by a range of high diversity values, in particular, again, evenness and dissimilarity. Variety mixtures may maintain a more even distribution of a greater, and more dissimilar, range of pathotypes than do uniform host populations.

Migration and gene flow. Another crucial factor is migration across Europe, making large parts of the continent into one unit of epidemiology and population genetics of the pathogen [Limpert, 1987; Limpert et al., 1990; Andrivon and Limpert, 1992; Wolfe et al., 1992]. From Table 3 the centrally-located regions are likely to receive immigrant spores from numerous neighbouring regions. In contrast, the lowest import of spores from distant populations can be expected for peripherally located areas, thus explaining the low richness and dissimilarity found in the samples from Scotland, southern France, Italy and Spain.

Migration and increase of complexity. As a consequence of migration and its effect combined with selection we would also expect an increase of complexity and diversity following the prevailing westerly winds in Europe, as proposed by Limpert [1985]. The present data tend to confirm this hypothesis: among the samples with greatest pathotype complexity and richness we find those from central and eastern Europe like Czechoslovakia, eastern Germany and Poland. The only more easterly part investigated is Byelorussia where complexity was still higher [Müller et al., 1992; Müller, 1993], which fits the hypothesis.

Population size. Diversity is correlated positively with the size of the population from which the samples are taken [Hurlbert, 1971; Poole, 1974]. Populations of the barley mildew pathogen are extraordinarily large and there are considerable differences among regions. Moreover, the pathogen is endemic in all areas considered and, in contrast to potato blight [Fry et al., 1992] there is no indication of bottle necks that would reduce diversity. The high pathotype richness in the Danish, east German and Czech samples could be influenced by the high density of host cultivation in these regions. Northern Italy and southern France, on the other hand, have the lowest density of barley in the area investigated [Müller, 1993] which leads to smaller pathogen populations. This may be reflected in the diversity values.

Table 4. Relationship of genetic parameters in a host-pathogen system. Indicated is the effect of the number of resistance genes in the population of the host on selected parameters of the pathogen population

Number of resistance genes used:	0	1	2	3
Maximum number of virulence genes selected:	0	1	2	3
Maximum number of virulence pathotypes possible:	1	2	4	8
Average complexity of pathotypes*:	0	0,5	1	1,5
Average dissimilarity of pathotypes*:	0	1	1,3	1,7
Maximum Simpson index $(1 - \lambda)$**:	0	0,51	0,76	0,89
Maximal richness**:	0,012	0,024	0,047	0,094

* If virulence factors randomly distributed.
** In a sample of 85 isolates.

Sexual recombination. High levels of sexual recombination should extend variation and may promote high levels of richness, evenness and dissimilarity, depending on the gene frequencies. Unfortunately, we do not have reliable estimates for the extent of sexual recombination, but it may not be a dominant factor given the wide ranges of variation evident in Table 3.

Conclusions and outlook

Diversity indices will be useful only if their principal properties and problems are recognised. Generally, the composition of pathotype samples cannot be described and interpreted using only one index. The diversity measures proposed in this study offer an approach for describing the most important aspects of the population structure based on pathotype frequencies.

Dissimilarity, the novel index of diversity, appears useful for population genetical studies, as it indicates the genetic distance among pathotypes in a sample. This index may therefore fill an important need, not only in plant pathology.

However, instead of describing populations independently from each other, the same concept can be applied to estimate the genetic distance between populations. It can thus be used as an alternative to Nei's standard genetic distance [Nei, 1987] or Rogers distance index [Rogers, 1972]. These established distance indices share the drawbacks of the established indices of diversity in that they do not account for the degree of genetic distance of pathotypes. Recently a similar concept of pairwise comparisons of isolates was developed for the latter purpose [Kosman pers. comm.], however using a different further approach; each isolate sampled from one population is matched with an isolate sampled from the second population so as to minimize the sum of distances between corresponding pairs of isolates.

Extending the scope of previous work [Lebeda, 1982; Groth and Roelfs, 1987a; Welz and Kranz, 1987; Leonard and Leath, 1990; Kolmer, 1991; Fry et al., 1992; Andrivon and Limpert, 1992] we also considered Fisher's logarithmic series model [Fisher et al., 1943]. The observed frequency distribution of pathotypes across Europe accorded well with the model (Figure 2). The logseries index α is widely applied in ecology and combines the advantage of being easy to interpret and statistically sound [Magurran, 1988]. Therefore, this index would merit further consideration in future studies of populations of plant pathogens.

The detection of specific diversity patterns can help in deriving and testing hypotheses about the underlying structure, causal factors and processes of population dynamics. For this aim, separate studies have to be carried out at different levels. Further insight into population structure and dynamics can also be obtained by the use of characters other than virulence, such as fungicide resistance or molecular markers. The latter are, in contrast to the former, under no, more limited, or different selection.

Existing evidence [Limpert et al., 1990; Munk et al., 1991; Wolfe, 1984] as well as models for the evolution of populations on a local scale [Brändle et al., 1992; Hovmøller et al., 1993] demonstrate the primary significance of local selection for the composition of a population. However we see much evidence showing that the exchange of genotypes on a large scale may also influence significantly local populations [Limpert, 1987; Andrivon and Limpert, 1992; Wolfe et al., 1992; Müller, 1993]. On the other hand our knowledge about quantitative effects of recombination and mutation is

still insufficient to estimate reliably the impact of these processes on population development.

A comprehensive model for the population biology of *E. g.* f. sp. *hordei* on a European scale should include a quantitative consideration of the aspects mentioned above. Such an extended model would be a highly valuable tool for the forecast of disease occurrence and for the development of improved disease control strategies.

Acknowledgements

We gratefully acknowledge financial support from the Swiss Federal Institute of Technology (ETH) Zürich, and appreciated the comments of the referees.

References

Andrivon D and Limpert E (1992) Origin and proportions of the components of composite populations of *Erysiphe graminis* f. sp. *hordei*. J Phytopathology 135: 6–19

Brändle U, Schaffner D, Wolfe MS and McDermott JM (1992) DNA and virulence variation in a field population of *Erysiphe graminis* f. sp. *hordei*. Vortr Pflanzenzüchtg 24: 37–38

Efron B and Tibshirani R (1991) Statistical data analysis in the computer age. Science 253: 390–395

Fisher RA, Corbet AS and Williams CB (1943) The relation between the number of individuals and the number of species in a random sample of an animal population. J Anim Ecol 12: 42–58

Fry WE, Goodwin SB, Matuszak JM, Spielman LJ, Milgroom MG and Drenth A (1992) Population genetics and intercontinental migrations of *Phythophthora infestans*. Annual Review of Phytopathology 30: 107–129

Groth JV and Roelfs AP (1987a) The concept and measurement of phenotypic diversity in *Puccinia graminis* on wheat. Phytopathology 77: 1395–1399

Groth JV and Roelfs AP (1987b) Analysis of virulence diversity in populations of plant pathogens. In: Wolfe MS and Caten CE (eds) Populations of plant pathogens (pp. 63–74). Blackwell Scientific Publications

Groth JV and Roelfs AP (1989) The analysis of genetic variation in populations of rust fungi. In: Leonard KJ and Fry WE (eds) Plant disease epidemiology. Vol II. Genetics, resistance and management (pp. 318–339). McGraw-Hill, New York

Hovmøller MS, Munk L and Østergård H (1993) Observed and predicted changes in virulence gene frequencies at 11 loci in a local barley powdery mildew population. Phytopathology 83: 253–260

Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. Ecology 52: 577–586

Huang R, Kranz J and HG Welz (1995) Increase of complex pathotypes of *Erysiphe graminis* f. sp. *hordei* in two-component mixtures of spring barley cultivars. J Phytopathology 143: 281–286

Jørgensen JH (ed) (1991) Integrated control of cereal mildews: virulence patterns and their change. Risø National Laboratory, Roskilde, Denmark

Kolmer JA (1991) Phenotypic diversity in two populations of *Puccinia recondita* f. sp. *tritici* in Canada during 1931–1987. Phytopathology 81: 311–315

Kolster P, Munk L, Stolen O and Lohde J (1986) Near-isogenic barley lines with genes for resistance to powdery mildew. Crop Science 26: 903–907

Lebeda A (1982) Measurement of genetic diversity of virulence in populations of phytopathogenic fungi. Journal of Plant Diseases and Protection 89: 88–95

Leonard KJ and Leath S (1990) Genetic diversity in field populations of *Cochliobolus carbonum* on corn in North Carolina. Phytopathology 80: 1154–1159

Limpert E (1985) Ursachen unterschiedlicher Zusammensetzung des Gerstenmehltaus, *Erysiphe graminis* f. sp. *hordei* Marchal, und deren Bedeutung für Züchtung und Anbau von Gerste in Europa. Thesis. TU München

Limpert E (1987) Frequencies of virulence and fungicide resistance in the European barley mildew population in 1985. J Phytopathology 119: 298–311

Limpert E, Andrivon D and Fischbeck G (1990) Virulence patterns in populations of *Erysiphe graminis* f. sp. *hordei* in Europe in 1986. Plant Pathology 39: 402–415

Limpert E, Andrivon D, Knittel R and Fischbeck G (1991) Barley mildew in Europe: patterns of composition of the pathogen population during the period 1985–1988. In: Jørgensen JH (ed) Integrated control of cereal mildews: virulence patterns and their change (pp. 87–103). Risø National Laboratory, Roskilde, Denmark Ludwig JA and Reynolds JF (1988) Statistical Ecology. A Primer on Methods and Computing. John Wiley and Sons, New York

Magurran AE (1988) Ecological Diversity and Its Measurement. Princeton, New Jersey

Müller K (1993) Virulenzstruktur und -dynamik grossräumiger Populationen des Gerstenmehltaus (*Erysiphe graminis* DC f. sp. *hordei* Marchal) in Europa. Thesis No. 10008, ETH Zürich

Müller K, Limpert E and Wolfe MS (1992) Patterns and dynamics of populations of *E. graminis* f. sp. *hordei*: virulence analysis. Vortr Pflanzenzüchtg 24: 150–152

Munk L, Jensen HP and Jørgensen JH (1991) Virulence frequencies in barley powdery mildew in eastern Denmark 1974–1986. In: Jørgensen JH (ed) Integrated control of cereal mildews: Virulence patterns and their change (pp. 55–65). Risø National Laboratory, Roskilde, Denmark

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, 512 pp

Østergård H and Hovmøller MS (1991) Gametic disequilibria between virulence genes in barley powdery mildew populations in relation to selection and recombinations. I. Models. Plant Pathology 40: 166–177

Pielou EC (1975) Ecological Diversity. John Wiley and Sons, New York

Pielou EC (1977) Mathematical Ecology. John Wiley and Sons, New York

Poole RW (1974) An Introduction to Quantitative Ecology. McGraw-Hill series in Population Biology, New York

Roelfs AP and Groth JW (1980) A comparison of virulence phenotypes in wheat stem rust populations reproducing sexually and asexually. Phytopathology 70: 855–862

Rogers JS (1972) Measure of genetic similarity and genetic distance. In: Studies in genetics (pp. 145–153). University of Texas Publ 7213

Shannon CE and Weaver W (1949) The Mathematical Theory of Communication. Univ of Illinois Press, Urbana

Simpson EH (1949) Measurement of diversity. Nature 163: 688

Welz G and Kranz J (1987) Effects of recombination on races of a barley powdery mildew populations. Plant Pathology 36: 107–113

Wolfe MS (1984) Trying to understand and control powdery mildew. Plant Pathology 33: 451–466

Wolfe MS and Limpert E (eds) (1987) Integrated control of cereal mildews: monitoring the pathogen. Martinus Nijhoff Publishers, Dordrecht

Wolfe MS, Brändle U, Koller B, Limpert E, McDermott JM, Müller K and Schaffner D (1992) Barley mildew in Europe: population biology and host resistance. In: Johnson R and Jellis GJ (eds) Breeding for disease resistance. Kluwer Acad, Dordrecht

Zeller FJ and Fischbeck G (1992) Cereal Rusts and Mildews. Proc. 8th European and Mediterranean Cereal Rusts and Mildews Conference, Weihenstephan, Germany. Vortr Pflanzenzüchtung 24, 344 pp