JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

PERSPECTIVES

Natural Product Chemistry Meets Genetics: When Is a Genotype a Chemotype?

ANNE E. DESJARDINS*

Mycotoxin Research, National Center for Agricultural Utilization Research, Agricultural Research Service, U.S. Department of Agriculture, Peoria, Illinois 61604

The chemotype of a microbial or plant species has traditionally been defined as its profile of natural products, and the genotype has been defined as its genetic constitution or DNA sequence. The purpose of this perspective is to discuss applications of DNA genotyping, particularly by polymerase chain reaction (PCR)-amplification methods, to predicting natural product chemotypes of fungi and plants of importance in food and agriculture. Development of PCR genotyping for predicting chemotypes will require collaboration between molecular biologists and natural product chemists, as well as community standards for reporting data. PCR genotyping should be validated by chemical analysis of individuals that represent the allelic diversity of the target gene in the population. To avoid misinterpretation, it is critical to differentiate data obtained by genotyping from data obtained by chemical analysis. The obvious and appropriate solution is to retain the established meanings of genotype and chemotype, both of which have been in use for half a century in the fields of genetics and natural product chemistry.

KEYWORDS: Chemotype; natural product; trichothecene; genotype; PCR; *TRI* gene

INTRODUCTION

The chemotype of a microbial or plant species has traditionally been defined as its chemical phenotype, including its profile of natural products, and the genotype has been defined as its genetic constitution or DNA sequence. Decades before the discovery of DNA as the genetic code, pioneers of natural products chemistry recognized the interrelationship between the chemistry and the evolution of plants. In an era when many chemists did not believe in Darwin's theory of evolution and many botanists had no interest in chemistry, the brilliant young chemist Helen Abbott Michael emerged from her studies of plant glucosides and alkaloids to propose in 1886 that "the theory of evolution in plant life is best illustrated by the chemical constituents of vegetable form" (*1*).

Since the 19th century, determining the natural product profiles of plants, fungi, bacteria, and other organisms has been the focus of food, agricultural, and medicinal chemistry. Although most such investigations have had economical and practical objectives, some have focused on chemosystematics, the division of groups of organisms by chemotype, which is based on the principle that chemical profiles and other nonmorphological features are, like morphology, expressions of the genome. With the development of chromatographic and spectroscopic methods during the 1950s, large numbers of compounds could be identified quickly and unambiguously. Consequently, chemotyping became useful for distinguishing morphologically similar organisms, from identifying variants of the pathogenic bacterium *Salmonella* by their polysaccharides (*2*) to distinguishing hybrids of the plant genus *Baptisia* by their terpene profiles (*3*). Today, natural product chemistry and other classical morphological and nonmorphological methods retain a critical place in an integrated approach to biological systematics. However, with the development of DNA sequencing methods and the discovery of the polymerase chain reaction (PCR) for DNA amplification, biological systematics has increasingly been based on DNA sequence analysis.

The development of DNA-based molecular biology allowed the cloning of genes encoding enzymes for natural product biosynthesis, first from fungi in the late 1980s and then from plants in the early 1990s. At that time, scientists at the U.S. Department of Agriculture cloned trichothecene genes (*TRI*) from *Fusarium* and aflatoxin genes from *Aspergillus* and also discovered associated clusters of genes for the biosynthesis of these mycotoxins (*4, 5*). Since then, many fungal gene clusters for natural product biosynthetic pathways have been identified, including additional agriculturally important mycotoxins such as ergot alkaloids, fumonisins, ochratoxins, and zearalenones. In plants, clustering of genes for natural product biosynthetic pathways seems to be uncommon, although gene clusters have

* To whom correspondence should be addressed. Tel: 309-681-6378. Fax: 309-681-6665. E-mail: anne.desjardins@ars.usda.gov.

10.1021/jf801239j
Published on Web 08/09/2008

been found for some metabolites important for resistance to pests and diseases, including benzoxazinones in *Zea mays* (*6*), glucosinolates in *Arabidopsis thaliana* (*7*), and triterpenoids in *Avena* (*8*).

Natural products research has now entered the age of genomics with the sequencing and public release of complete or partial genomes of many agriculturally important microbes and plants. Access to fungal and plant genomes has revealed the presence of hundreds of candidate genes for polyketide synthases, nonribosomal peptide synthetases, terpene synthases, cytochrome P450 monooxygenases, acyltransferases, and other types of enzymes that synthesize natural products. Comparison of DNA sequences per se cannot prove gene function; however, most fungi and many plants are amenable to the techniques of biochemistry, classical genetics, and molecular genetics necessary to validate the function of candidate genes. With complete genomes and gene expression data on hand, researchers worldwide are working at a rapid pace to identify genes and gene clusters for the biosynthesis and regulation of natural products of importance in food and agriculture.

With the discovery of genes for biosynthesis of natural products, chemistry and genetics have entered a new interrelationship. Whereas previously natural product profiles were used to predict genetic relationships between organisms, now gene sequences can potentially be used to predict natural product profiles of plants or cultured microbes or perhaps even in food or agricultural samples. Molecular biologists in this field are becoming interested in developing and promoting PCR genotyping for research and commercial applications; some view PCR genotyping methods as less expensive replacements for traditional chemical analysis (*9*). Others have reported PCR-derived genotypes as actual chemotypes (for examples, see refs *10*−*14*), which contradicts established definitions of a genotype as a DNA sequence and a chemotype as a chemical phenotype.

The purpose of this perspective is to discuss applications of DNA genotyping, particularly by PCR amplification methods, to the prediction of natural product chemotypes. Most examples will be from trichothecene (*TRI*) genotyping, because this is a field in which most of the relevant biosynthetic genes have been sequenced (as reviewed in ref *4*) and in which PCR amplification methods of predicting chemotype have been most widely applied to date. Consequently, in trichothecene research, sufficient data have accumulated to allow evaluation of some of the strengths and weaknesses of using genotypes to predict chemotypes. This writer's intent is to interest both chemists and molecular biologists in the evolving interrelationship between natural product chemistry and genetics and to encourage collaboration rather than a hostile takeover.

NATURAL PRODUCT CHEMOTYPES AND GENOTYPES

Genotyping by PCR. Development of chromatographic methods in the 1950s allowed sensitive, specific, and rapid chemotyping. Similarly, development of PCR methods in the 1980s allowed sensitive, specific, and rapid genotyping. The most critical aspect of PCR genotyping is the design of synthetic oligonucleotides to prime the synthesis of the target DNA sequence. If the goal is to amplify a gene between species, then sequence in a more-conserved, but not too conserved, region of the gene should be targeted. Conversely, if the goal is to compare DNA sequence differences (DNA polymorphisms) between individuals of the same species, then less-conserved regions of the gene should be targeted. Low specificity of primers may lead to false positives due to amplification of nontarget sequences, which may yield a product of similar or different size when compared to the target sequence. PCR products of similar size but different sequence can be distinguished by DNA sequencing, but for most routine genotyping applications, PCR products are identified by size alone. A size difference of at least 10% [optimally $\geq$100 base pairs (bp)] is required to distinguish bands easily under standard conditions of gel electrophoresis.

It is a truth universally acknowledged that false negatives are inconveniently common in PCR amplification, for many different reasons. Even with a high-quality DNA template preparation and well-designed primers, the reaction can fail due to unsuitable reaction conditions or to overly specific primers that are not amplified due to unanticipated DNA sequence polymorphisms in the target sequence. A single nucleotide polymorphism (SNP) mismatch in a 20-nucleotide primer, especially near the 3′-end, can be sufficient to prevent amplification of the target DNA. Detection of false negatives is a particular problem when genotypes are scored only by whether or not a PCR product is produced from a particular pair of forward and reverse primers. False negatives are easier to detect when genotypes are scored by different sizes of PCR products produced by the same pair of forward and reverse primers.

Genetic Homology and the Predictive Limits of PCR Genotyping. Natural product profiles can vary widely between species and between individuals of a species, and the rationale for genotyping is the existence of causative and detectable genetic differences. In the simplest system, a gene would be present when the corresponding metabolite is present and absent when the metabolite is absent. Systems are necessarily more complex for natural products, which are synthesized by pathways comprising multiple genes and gene products. However, genes encoding enzymes for reactions at branch points from primary metabolism are good candidates for some genotyping applications, since they are critical for function of the entire natural product biosynthetic pathway. In *Fusarium*, the *TRI5* gene encodes trichodiene synthase, which catalyzes the first dedicated step in trichothecene biosynthesis. In a survey of *Fusarium* species, primers specific for conserved regions of *TRI5* amplified a product from trichothecene-producing species but not from nonproducing species (*15*). In the fungus *Stachybotrys*, in contrast, *TRI5*-specific primers not only amplified a product from all of the trichothecene-producing strains but also from 40% of the strains that produced no trichothecenes detectable by high-performance liquid chromatography or gas chromatography−tandem mass spectrometry analysis (*16*). In the *Stachybotrys* study, the PCR products were not sequenced, and it remains unclear whether the discrepancy was a false positive due to amplification of a nontarget gene or resulted from a loss in strain toxigenicity due to a nonfunctional variant of *TRI5*, the absence of other genes required for trichothecene biosynthesis, environmental factors, or some other reason. The frequency of false positives due to amplification of nontarget genes should decrease as primer design becomes more informational due to additional DNA sequences, in particular, sequences of genes that are homologues (derived from a common ancestor).

For fungal natural products, the reliability of genotype in predicting chemotype is increased by the divergence within a species of the sequences of similar genes with different functions. In *Fusarium*, three cytochrome P450 monooxygenases (*TRI4*, *TRI11*, and *TRI13*) that encode enzymes specific for successive oxygenations in the trichothecene biosynthetic pathway are tightly linked. Within a *Fusarium* species, however, *TRI4*, *TRI11*, and *TRI13* share less than 40% identity at the amino acid level, suggesting that, despite their physical linkage,

they did not result from recent gene duplication events. In like manner, comparative analysis of cytochrome P450 genes involved in the biosynthesis of fumonisins (*FUM2* and *FUM15*) (*4*) and gibberellins (*P450-1* and *P450-4*) (*17*) in *Fusarium* and of aflatoxins (*aflG*, *aflL*, *aflN, aflQ*, *aflU*, and *aflV*) in *Aspergillus* (*5*) indicates less than 40% identity at the amino acid level for linked genes that encode similar enzymes with different functions.

In fungi, the reliability of genotypes in predicting natural product chemotypes also is increased by the conservation between species of the DNA sequences of homologues with the same function. Between *Fusarium* species, homologues of *TRI4*, *TRI11*, and *TRI13* are highly conserved, with 78−90% identity at the amino acid level. However, it cannot be assumed that all genes in a biosynthetic pathway will contain the same level of identity between species. Homologues of a fourth trichothecene cytochrome P450 gene, *TRI1*, can share as little as 66% identity at the amino acid level between species (*4, 18*). Consequently, probes to the *TRI1* gene from *Fusarium sporotrichioides* were unable to detect the *TRI1* homologue from *Fusarium graminearum*.

For plant natural products, comparative sequence analyses indicate that many genes in biosynthetic pathways are members of large gene families with complex relationships with each other and with genes of primary metabolism (*19*). Analyses of some plant gene clusters have revealed evidence for relatively recent gene duplication with significant conservation of sequence between linked genes that encode enzymes with different functions, such as terpene synthases in *Arabidopsis* and *Z. mays* and methylthioalkylmalate synthases in *Arabidopsis* (*7, 19*). In some cases, duplication of genes of primary metabolism has led to acquisition of new functions without significant sequence divergence, as in the case of indole-3-glycerol phosphate lyase (*BX1*) in benzoxazinoid synthesis and homospermidine synthase (*HSS*) in pyrrolizidine alkaloid synthesis (*6, 19*). In *Senecio vernalis*, the *HSS* gene shares 79% identity at the amino acid level with desoxyhypusine synthase, the primary metabolism gene from which *HSS* likely evolved. High levels of sequence conservation among homologues with different functions are likely to complicate the assignment of gene function based on sequence alone and the use of genotype to predict chemotype for some plant natural products.

PCR Genotyping with Allele-Nonspecific Primers. Most PCR-based genotyping applications to date, whether in fungi, plants, or humans, have focused on distinguishing alleles, gene variants with different DNA sequences due to insertions, deletions, inversions, and SNPs. Large insertions and deletions (indels) often produce nonfunctional genes or pseudogenes. Single nucleotide changes can have several outcomes, causing no change of amino acid sequence or a change of amino acid sequence with no effect on the phenotype or causing a change of amino acid sequence or creation of a stop codon with small to large effects on the phenotype. The human hemoglobin B gene, which is mutated in individuals with sickle cell anemia, has more than 400 known alleles, but the most common form of the disease is due to a GAG → GTG change that replaces the amino acid glutamine with valine and changes conformation of the protein (*20*). Thus, this A → T SNP is a causative polymorphism, whereas other polymorphisms in the hemoglobin B gene and nearby genes are not causative but rather are only linked to the phenotype of sickle cell anemia. For human genetic diseases and for fungal and plant natural products, the prediction of phenotype by genotype is most assured of success with single genes and causative mutations.

Alleles that differ by large indels often have dramatically different phenotypes and also are easily distinguishable by PCR. Using a pair of forward and reverse primers that are allele-nonspecific and that frame a region containing indels will produce PCR products that differ by the size of the indels, allowing simple identification by electrophoresis. This assay has an inherent PCR control, since a PCR product should be amplified from each allele, and if it is not, a false negative is suspected.

Genotyping by allele-nonspecific indel PCR has been used successfully for analysis of natural product biosynthetic genes in fungi and plants. In trichothecene biosynthesis, the gene *TRI13* encodes a cytochrome P450 responsible for 4-oxygenation of deoxynivalenol (DON) to produce nivalenol (NIV). Strains with a functional *TRI13* can produce both DON and NIV, whereas strains with a *TRI13* pseudogene (Ψ*TRI13*) can produce DON but not NIV. Production of NIV or DON has some mycotoxicological relevance since NIV is considerably more toxic than DON in most animal systems (*4*). The Ψ*TRI13* gene of DON producers of *F. graminearum* contains large deletions in the coding region, which render the gene nonfunctional. Using allele-nonspecific primers that frame the region containing the indels, *TRI13* and Ψ*TRI13* strains produce PCR products that differ by 276 bp, the size of the indels. With this strategy, Kim and associates (*21*) found a 100% correlation between *TRI13* genotyping and chemical analysis of DON and NIV production among 163 trichothecene-producing strains of *F. graminearum* from Korea and the United States.

Our research group also has used *TRI13* genotyping with allele-nonspecific primers (*21, 22*) to distinguish DON-producing and NIV-producing strains of *F. graminearum* from Nepal (*23*). For detection of chemically confirmed DON producers in this fungal population, however, we have found that primers that bind closer to the ends of the gene (*22*) are more reliable than primers that bind closer to the indels (*21*). One likely reason for these false negatives is primer mismatch due to unanticipated polymorphisms in the target sequence, since Ψ*TRI13* genes from Nepal strains are polymorphic, especially in the vicinity of the indels (*22*). Nonfunctional pseudogenes are known to be more varied than functional genes; thus, *TRI13* genotypes should be validated by chemical analysis of representative strains to confirm that all DON producers are being detected in the target fungal population.

Genotyping by allele-nonspecific PCR also has been used for analysis of carotenoid biosynthetic genes in maize. Different profiles of carotenoids, as determined by chemical analysis of maize grain, were associated with an indel in the promoter region of the gene lycopene epsilon cyclase (*lcyE*) (*9*). Using allele-nonspecific primers that frame the region containing the indel, maize lines with higher levels of β-carotenoids yielded an *lcyE* PCR product of 993 bp and lines with lower levels yielded a PCR product of 250 bp. The association of genotype with chemotype across more than 200 lines of maize, along with other evidence, supports the hypothesis that the *lcyE* 5′-indel is a causative mutation for carotenoid chemotype.

PCR Genotyping with Allele-Specific Primers. Whether they differ by large indels or only by SNPs, alleles can be distinguished by allele-specific PCR, in which primers are designed so that one allele produces a PCR product and all other alleles do not. Allele-specific PCR has numerous variations. In one variation, both forward and reverse primers are allele-specific, so that two unique primers are needed to amplify each allele of the target gene. In another variation, a forward primer common to all alleles is combined with two or more reverse

primers, each of which binds to only one allele. Allele-specific primers can bind to different locations or to the same location in the gene sequence. For the latter technique, the primer is designed to match one allele perfectly but to mismatch the other allele near the 3′-end such that amplification is inefficient. Alleles can be scored separately with each primer pair or scored simultaneously by combining multiple pairs of primers in multiplex PCR reactions. Detection of false negatives is a particular problem for allele-specific PCR because genotypes are scored only by whether or not a PCR product is produced from a particular pair of forward and reverse primers. For this reason, all allele-specific PCR reactions require a positive control that is not allele-specific.

Allele-specific PCR has been extensively developed for human genetic testing, where false positives and false negatives can have serious consequences. PCR-positive controls are incorporated in all assays, so that the number and size of PCR products will discriminate false negatives that are due to failure of the PCR amplification from true negatives that are due to absence of the allele. In some protocols, primers that amplify an external gene as a positive control are combined in the same reaction tube with allele-specific primers. This strategy is simple but may not be optimal if primers for the external control and target genes differ in efficiency, stability, or other factors. A more complicated but more rigorous strategy is to combine allele-nonspecific primers as a positive internal control in the same reaction tube with allele-specific primers for the same target gene.

Allele-specific PCR has been applied, to a limited extent, to genotyping of indels and SNPs of natural product biosynthetic genes in fungi and plants, including the *TRI13* gene from *Fusarium* and the *lcyE* gene from maize (*9, 22*). For indel PCR, both studies used similar strategies that combine an allele-nonspecific primer and two allele-specific primers: an insertion-specific primer and a deletion-specific primer targeted to the two sides of the insertion site. Minor variations in chemotype also have been associated with some SNPs in *icyE* and *TRI* genes, but causative roles have not been proven for any of these mutations. The allele-specific indel and SNP PCR assays of *lcyE* included external positive amplification controls, but allele-specific *TRI* gene assays appear to have included no positive controls (for examples, see refs *10*−*14*).

Complex Genetics and the Predictive Limits of PCR Genotyping. Natural products are synthesized by long and complex biosynthesis pathways that are controlled by a network of genes. The translation of genotype into chemotype is complicated by gene interactions and environmental factors (*24*). In studies of fungal natural product pathways, gene-disruption mutants and precursor feeding experiments have played critical roles in elucidating gene functions. In trichothecene biosynthesis, chemotypes of multiple mutants in an otherwise isogenic background have shown that the acetylations and deacetylations at carbons 3 and 15 are controlled by at least three genes: *TRI3*, *TRI8*, and *TRI101*. In *F. graminearum*, *TRI3* and *TRI8* are tightly linked, but *TRI101* is on a different chromosome. Interactions among these genes and the effect of environmental factors on patterns of acetylation are poorly understood, and the amounts and ratios of acetylated compounds can vary widely for an individual strain grown in different environments (*25*−*27*). *TRI3*-PCR has found associations between SNPs and different ratios of the minor acetylated metabolites 3-acetyldeoxynivalenol and 15-acetyldeoxynivalenol in some *Fusarium* populations (*10*−*14*). None of these mutations, however, has yet been proven to alter the function of the *TRI3* enzyme or the trichothecene chemotype. Such noncausative mutations may be associated with a chemotype in one genetic background but not in another. Therefore, the predictive value of noncausative SNPs of *TRI3*, and of other genes of complex natural product biosynthetic pathways, must be determined by chemical analysis of individuals that represent the allelic diversity of each population under study. Some studies have validated *TRI3* genotyping by chemical analysis of at least some representative strains (*12*−*14*), but other studies appear not to have done so (*10, 11*).

Linkage Analysis and the Predictive Limits of PCR Genotyping. For many traits, causative genes and mutations are not yet known; thus, genotyping by PCR is based on genetic differences that are only associated or linked with the trait of interest. Such combinations of linked alleles or DNA polymorphisms that tend to be inherited together are called haplotypes. Haplotypes can be broken by mutation, which is rare, and by recombination through sexual reproduction, which is a frequent phase of the life cycle of many plants and of some mycotoxigenic fungi. DNA sequence analyses provide evidence for a history of recombination in the trichothecene and aflatoxin gene clusters (*28, 29*). Furthermore, experimental crosses of *F. verticillioides* suggest a recombination frequency of about 6% between *FUM1* and *FUM2* within a 1780 bp region of the fumonisin gene cluster (*30, 31*). It also cannot be assumed that biosynthetic genes that are clustered in one species will be similarly linked in another species. Among NIV-producing *Fusarium* species, the cytochromes P450 *TRI1* and *TRI13* are clustered in *F. equiseti* but are on different chromosomes in *F. graminearum* (*18*). Recombination between haplotypes or between unlinked genes could lead to loss of associations in some populations and the occurrence of both false positives and false negatives in prediction of chemotypes by PCR genotypes.

Despite the problems inherent in association studies, a number of linked *TRI* genes and intergenic regions have been proposed as markers for predicting trichothecene chemotypes, especially to distinguish DON producers and NIV producers. Proposed DON−NIV diagnostic markers include *TRI3*, *TRI5*, the *TRI5-6* intergenic region, *TRI12* (a transporter), and the 4-acetyltransferase *TRI7*, which modifies NIV (*10*−*14, 21, 22*). In *F. graminearum* and related species, all of these markers are in the trichothecene gene cluster upstream (5′) of *TRI13*, which encodes the cytochrome P450 actually required for the 4-oxygenation of DON to NIV. Among *F. graminearum* strains from Korea, *TRI5* assays (95% correlation) and *TRI7* assays (98% correlation) were almost as reliable as *TRI13* assays (100% correlation) in predicting actual chemotypes of DON producers and NIV producers (*21*), likely because sequences from Korean strains were included in primer design.

Recent *TRI* genotype surveys of *F. graminearum* from China that have targeted different *TRI* genes appear to have obtained contradictory results. Using primers for *TRI13*, Ji and associates (*10*) detected NIV genotypes at a frequency of only 6% among strains from Hubei Province. In contrast, Zhang and associates (*11*) detected NIV genotypes in Hubei at much higher frequencies of 22%, based on primers for the *TRI5-6* intergenic region, or 35%, based on primers for *TRI7*. The apparent *TRI5*-DON/*TRI7*-NIV haplotype of several of these strains may have resulted from recombination between these two loci. In addition, DNA sequence and chemical analysis from an earlier study (*22*) indicate that the reported *TRI7* NIV genotypes of at least two of the Hubei strains are false positives. These two strains produced only DON by chemical analysis, had a Ψ*TRI13* gene with large deletions, and had a *TRI7* PCR product that is the size expected for NIV producers but with a predicted translation stop codon in the middle of the gene (*22*). In addition, previous

chemical analyses of cereal grains from eastern China have consistently reported DON as the major trichothecene present, with NIV usually present at 10-fold lower levels than DON (*4*). Because neither of the cited PCR-based surveys (*10, 11*) of *F. graminearum* from China included a thorough validation of the NIV genotype by chemical analysis, the actual frequency of NIV producers in this population remains uncertain.

CONCLUSIONS AND TERMINOLOGY

Fifty years ago, Holger Erdtman cautioned his colleagues in natural products chemistry: "Chemical arguments should be taken only as contributions to a taxonomic discussion which are valuable because they represent a completely different approach. Nothing can discredit the chemical approach to taxonomy more than the uncritical over-estimation of the chemical method" (*32*). The argument of this perspective is that similar caution is necessary now in applying PCR genotyping and other DNA sequence-based methods to the prediction of natural product chemotypes.

It is clear that PCR-based genotyping is a rapidly growing field of diagnostic testing that will contribute to, but not replace, the analysis of natural products of agriculturally important fungi, plants, and plant products. Development and validation of PCR-based technology for predicting chemotypes will require collaboration between molecular biologists and natural product chemists, as well as community standards for reporting and interpreting data. Validation of the predictive relationship requires fulfilling two important criteria. First, PCR genotyping assays should contain the appropriate controls to identify false positives and false negatives. Second, PCR genotyping data should be validated by chemical analysis of individuals that represent the allelic diversity of the target gene in the population. The assumptions and limitations of genetic testing, and especially of linkage analysis, should be discussed, and confidence intervals of the data should be provided, if possible. An accuracy of 95% may be sufficient for most applications of genetic testing in natural product chemistry but may not be sufficient for applications related to food safety, international trade, or other politically and economically sensitive areas of food and agriculture.

The complex relationship between genotype and chemotype ensures that the presence of a gene for a natural product biosynthetic pathway neither guarantees the presence of a metabolite produced by that pathway nor predicts the levels produced of that metabolite. Therefore, it is critical and important that there remain a differentiation between genotype data obtained by DNA analysis and chemotype data obtained by chemical analysis. The obvious and appropriate solution is to retain the established meanings of genotype and chemotype, both of which have been in use for half a century in the fields of genetics and natural product chemistry. The terms "genotype" and "phenotype" have a long history; both were introduced in 1909 by the Danish botanist Wilhelm Johannsen, who defined the genotype as the genetic constitution of the organism and the phenotype as characteristics that result from the interaction of genotype and environment (*33*). The term "chemotype" appears to have come into use much later but had appeared by the early 1960s in the chemical, genetic, and microbiological literatures (*34, 35*), with a clear and precise definition as a chemical phenotype, the "chemical expression of a genotype" (*35*). Although it is true that the meanings of words, even scientific words, can change over time, such a change requires a strong rationale and community consensus. A survey of standard journals in the fields of genetics and natural product chemistry since 1960 indicates that, to geneticists and natural product chemists alike, a genotype remains a genetic constitution or DNA sequence and a chemotype remains a chemical phenotype such as a natural product profile. In a text search of recent publications, including 500 papers in the journal *Genetics*, which is published by the American Genetics Society, and more than 250 papers in journals published by the American Chemical Society, this writer found no evidence of the use of genotype to indicate a natural product chemical phenotype. To date, the misusage of the term chemotype to refer to a natural product genotype appears to be restricted to the use of the term trichothecene chemotype to refer to *TRI*-gene alleles (for examples, see refs *10*−*14*), and let us hope that this error is not further propagated. Rather, I propose that strains characterized only by *TRI* genotyping be referred to as trichothecene-genotypic strains or trichothecene genotypes. The misusage of the term chemotype blurs an important and long-established distinction between genotype and phenotype and should not be accepted as a community standard for reporting alleles of genes for natural product biosynthetic pathways.

ABBREVIATIONS USED

Allele, an alternative form of a gene; haplotype, linked alleles or polymorphisms; homologues, genes derived from a common ancestor; indel, insertion or deletion; bp, base pair; PCR, polymerase chain reaction; polymorphism, DNA sequence difference; SNP, single nucleotide polymorphism; *TRI*, trichothecene gene; DON, deoxynivalenol; NIV, nivalenol.

ACKNOWLEDGMENT

For their comments on the manuscript, I thank my colleagues Nancy Alexander, Michael Appell, Mark Busman, Chris Maragos, Susan McCormick, and Robert Proctor. I thank Robert Proctor for DNA sequence comparisons.

LITERATURE CITED

(1) Michael, H. A. *Studies in Plant and Organic Chemistry and Literary Papers*; Riverside Press: Cambridge, MA, 1907.
(2) Beale, G. H.; Wilkinson, J. F. Antigenic variation in unicellular organisms. *Annu. Rev. Microbiol.* **1961**, *15*, 263–296.
(3) Alston, R. E.; Turner, B. L. *Biochemical Systematics*; Prentice-Hall: Englewood Cliffs, NJ, 1963.
(4) Desjardins, A. E. *Fusarium Mycotoxins: Chemistry, Genetics, and Biology*; American Phytopathology Society Press: St. Paul, MN, 2006.
(5) Yu, J.; Chang, P.-K.; Ehrlich, K. C.; Cary, J. W.; Bhatnagar, D.; Cleveland, T. E.; Payne, G. A.; Linz, J. E.; Woloshuk, C. P.; Bennett, J. W. Clustered pathway genes in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2004**, *70*, 1253–1262.
(6) Gierl, A.; Frey, M. Evolution of benzoxazinone biosynthesis and indole production in maize. *Planta* **2001**, *213*, 493–498.
(7) Benderoth, M.; Textor, S.; Windsor, A. J.; Mitchell-Olds, T.; Gershenzon, J.; Kroymann, J. Positive selection driving diversification in plant secondary metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9118–9123.
(8) Qi, X.; Bakht, S.; Leggett, M.; Maxwell, C.; Melton, R.; Osbourn, A. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8233–8238.
(9) Harjes, C. E.; Rocheford, T. R.; Bai, L.; Brutnell, T. P.; Kandianis, C. B.; Sowinski, S. G.; Stapleton, A. E.; Vallabhaneni, R.; Williams, M.; Wurtzel, E. T.; Yan, J.; Buckler, E. S. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **2008**, *319*, 330–333.
(10) Ji, L.; Cao, T.; Wang, S. Determination of deoxynivalenol and nivalenol chemotypes of *Fusarium graminearum* isolates from China by PCR assay. *J. Phytopathol.* **2007**, *155*, 505–512.

(11) Zhang, J.-B.; Li, H.-P.; Dang, F.-J.; Qu, B.; Xu, Y.-B.; Zhao, C.-S.; Liao, Y.-C. Determination of the trichothecene mycotoxin chemotypes and associated geographical distribution and phylogenetic species of the *Fusarium graminearum* clade from China. *Mycol. Res.* **2007**, *111*, 967–975.

(12) Jennings, P.; Coates, M. E.; Walsh, K.; Turner, J. A.; Nicholson, P. Determination of deoxynivalenol and nivalenol chemotypes of *Fusarium graminearum* isolated from wheat crops in England and Wales. *Plant Pathol.* **2004**, *53*, 643–652.

(13) Quarta, A.; Mita, G.; Haidukowski, M.; Santino, A.; Mulè, G.; Visconti, A. Assessment of trichothecene chemotypes of *Fusarium culmorum* occurring in Europe. *Food Addit. Contam.* **2005**, *22*, 309–315.

(14) Starkey, D. E.; Ward, T. J.; Aoki, T.; Gale, L. R.; Kistler, H. C.; Geiser, D. M.; Suga, H.; Tóth, B.; Varga, J.; O'Donnell, K. Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity. *Fungal Genet. Biol.* **2007**, *44*, 1191–1204.

(15) Edwards, S. G.; Pirgozliev, S. R.; Hare, M. C.; Jenkinson, P. Quantification of trichothecene-producing *Fusarium* species in harvested grain by competitive PCR to determine efficacies of fungicides against *Fusarium* head blight of winter wheat. *Appl. Environ. Microbiol.* **2001**, *657*, 1575–1580.

(16) Peltolta, J.; Niessen, L.; Nielsen, K. F.; Jarvis, B. B.; Andersen, B.; Salkinoja-Salonen, M.; Moller, E. M. Toxigenic diversity of two different RAPD groups of *Stachybotrys chartarum* isolates analyzed for potential for trichothecene production and for boar sperm cell motility inhibition. *Can. J. Microbiol.* **2002**, *48*, 1017–1029.

(17) Tudzynski, B. Gibberellin biosynthesis in fungi: genes, enzymes, evolution, and impact on biotechnology. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 597–611.

(18) Proctor, R.; McCormick, S.; Alexander, N.; Desjardins, A. Variation in *TRI1* in thirteen trichothecene-producing species of *Fusarium*: Evidence for a complex evolutionary history of a mycotoxin biosynthetic locus. In *9th European Conference on Fungal Genetics, Meeting Abstracts*; Archer, D. , Eds.; University of Edinburgh, UK, Fungal Genetics Ltd.: Nottingham, United Kingdom, 2008; p 29.

(19) Hartmann, T. From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochemistry* **2007**, *68*, 2831–2846.

(20) Waterfall, C. M.; Cobb, B. D. Single tube genotyping of sickle cell anaemia using PCR-based SNP analysis. *Nucleic Acids Res.* **2001**, *29* (23), e119.

(21) Kim, H.-S.; Lee, T.; Dawlatana, M.; Yun, S.-H.; Lee, Y.-W. Polymorphisms of trichothecene biosynthesis genes in deoxynivalenol- and nivalenol-producing *Fusarium graminearum* isolates. *Mycol. Res.* **2003**, *107*, 190–197.

(22) Chandler, E. A.; Simpson, D. R.; Thomsett, M. A.; Nicholson, P. Development of PCR assays to *Tri7* and *Tri13* trichothecene biosynthetic genes, and characterization of chemotypes of *Fusarium graminearum*, *Fusarium culmorum* and *Fusarium cerealis*. *Physiol. Mol. Plant Pathol.* **2003**, *62*, 355–367.

(23) Desjardins, A. E.; Busman, M.; Manandhar, G.; Jarosz, A. M.; Manandhar, H. K.; Proctor, R. H. Gibberella ear rot of maize (*Zea mays*) in Nepal: Distribution of the mycotoxins nivalenol and deoxynivalenol in naturally and experimentally infected maize. *J. Agric. Food Chem.* **2008**, *56*, 5428–5436.

(24) Dixon, R. A.; Gang, D. R.; Charlton, A. J.; Fiehn, O.; Kuiper, H. A.; Reynolds, T. L.; Tjeerdema, R. S.; Jeffrey, E. H.; German, J. B.; Ridley, W. P.; Seiber, J. N. Applications of metabolomics in agriculture. *J. Agric. Food Chem.* **2006**, *54*, 8984–8994.

(25) Desjardins, A. E.; Plattner, R. D. Diverse traits for pathogen fitness in *Gibberella zeae*. *Can. J. Plant Pathol.* **2003**, *25*, 21–27.

(26) Bily, A. C.; Reid, L. M.; Savard, M. E.; Reddy, R.; Blackwell, B. A.; Campbell, C. M.; Krantis, A.; Durst, T.; Philogène, B. J. R.; Arnason, J. T.; Regnault-Roger, C. Analysis of *Fusarium graminearum* mycotoxins in different biological matrices by LC/MS. *Mycopathologia* **2004**, *1457*, 117–126.

(27) McMullen, M.; Jordahl, J.; Meyer, S. Differential effects of infection timing on *Fusarium* head blight and on DON and DON derivatives in three spring grains. In Proceedings of the 2007 National Fusarium Head Blight Forum, 2−4 December, 2007, Kansas City, MI; Canty, S. , Eds.; Michigan State University: East Lansing, MI, 2007; p 100.

(28) Carbone, I.; Jakobek, J. L.; Ramirez-Prado, J. H.; Horn, B. W. Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*. *Mol. Ecol.* **2007**, *16*, 4401–4417.

(29) Ward, T. J.; Bielawski, J. P.; Kistler, H. C.; Sullivan, E.; O'Donnell, K. Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9278–9283.

(30) Desjardins, A. E.; Plattner, R. D.; Proctor, R. H. Linkage among genes responsible for fumonisin biosynthesis in *Gibberella fujikuroi* mating population A. *Appl. Environ. Microbiol.* **1996**, *62*, 2571–2576.

(31) Proctor, R. H.; Plattner, R. D.; Desjardins, A. E.; Busman, M.; Butchko, R. A. E. Fumonisin production in the maize pathogen *Fusarium verticillioides*: genetic basis of naturally occurring chemical variation. *J. Agric. Food Chem.* **2007**, *54*, 2424–2430.

(32) Erdtman, H. Organic chemistry and conifer taxonomy. In Perspectives in Organic Chemistry; Todd, A., Ed.; Interscience: New York, 1956. Quoted in the following: Swain, T., Ed. Chemical Plant Taxonomy; Academic Press: London, 1963.

(33) Mahner, M.; Kary, M. What exactly are genomes, genotypes and phenotypes? And what about phenomes. *J. Theor. Biol.* **1997**, *186*, 55–63.

(34) Hubby, J. L.; Forrest, H. S. Studies on the mutant maroon-like in *Drosophila melanogaster*. *Genetics* **1960**, *45*, 211–224.

(35) Westphal, O.; Lüderitz, O.; Staub, A. M. Bacterial endotoxins. *J. Med. Pharm. Chem.* **1961**, *4*, 497–504.

Received for review April 18, 2008. Revised manuscript received June 6, 2008. Accepted July 10, 2008. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

JF801239J