PROGRESS REPORT

Objectives for the past three years have been to: 1. elucidate the structure and genomic organization of genes involved in lignin degradation; and 2. investigate the expression of these genes in _Phanerochaete chrysosporium_ and in heterologous hosts.

Major accomplishments follow: 1. The _P. chrysosporium_ gene encoding glyoxal oxidase has been cloned, sequenced, and efficiently expressed in _Aspergillus_. 2. Mapping methods were developed allowing the integration of genetic and physical maps of _P. chrysosporium_. 3. Highly specific and sensitive PCR techniques were developed for discriminating closely related mRNAs. Applications of this technique will help to identify specific genes involved in degradation of lignin and organopollutants. 4. Investigations have revealed a novel insertional mutation in lignin peroxidase gene _lipI_.

I. Research summaries

Progress in 5 somewhat overlapping research areas is outlined below. As originally proposed, research focused on gene structure and genomic organization. However, unforeseen technical advances have shifted some emphasis toward studies on glyoxal oxidase and on an unusual _lip_ mutant. Efforts to further develop genetic transformation based on 5-fluoroorindole resistance have been sharply curtailed largely because an impressive number of systems are now available [1-7].

In the following section, brief summaries are provided for both published and unpublished studies. Full descriptions can be found in several attached reprints. In Section II, highly condensed versions of three completed but unpublished (submitted or in preparation) studies are presented. Unless otherwise specified, DOE funds supported all research and is acknowledged accordingly in the publications.

1. Glyoxal oxidase studies

Kersten, P. J. and D. Cullen. 1993. Cloning and characterization of a cDNA encoding a novel peroxide-producing enzyme from the lignin-degrading basidiomycete _Phanerochaete chrysosporium_. Proc. Nat. Acad. Sci. USA 90:7411-7413. (Reprint attached). These investigations were initiated by Cullen and Kersten with funding from DOE and USDA competitive grant 88-33521-4089 which expired in 1991. Since then, DOE funds have supported all GLOX work.

Kersten, P.J., C. Witek, and D. Cullen. 1994. Genomic organization and heterologous expression of the _Phanerochaete chrysosporium_ gene encoding glyoxal oxidase. (In preparation) Two genomic clones encoding allelic variants of glyoxal oxidase were isolated from a _P. chrysosporium_ BKM-F-1767 cosmid library. Sequence comparisons with cDNAs showed the presence of 4 introns ranging from 53 to 61 bp. Predicted amino acid sequences of allelic variants differed by a single residue; Lys vs Thr at position 308. Southern blot hybridization of pulse field gels localized _glx_ to a dimorphic chromosome pair separate from all known peroxidases. Expression vectors were constructed and fully active GLOX was obtained under the control of the glucoamylase promoter in _Aspergillus nidulans_. Expression was regulated by carbon source and extracellular yields were 50-fold higher than in _P. chrysosporium_ cultures. (More detailed results are presented in Section II.)
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
2. *Phanerochaete chrysosporium* expression studies.

Stewart, P., Kersten, P., Vanden Wymelenberg, A., Gaskell, J., and D. Cullen. 1992. Lignin peroxidase gene family of *Phanerochaete chrysosporium*: complex regulation by carbon and nitrogen limitation and identification of a second dimorphic chromosome. J. Bact. 174:5036-5042. (Reprint attached). This competitive RT-PCR method has been modified and expanded to include analysis of a wide range of transcripts (e.g. β-tubulin, cbh1s, mnp1s, and all lins) from a number of complex substrates. These improvements are described in the Proposed Research section.

3. Genomic organization of *Phanerochaete chrysosporium*.

Gaskell, J. and D. Cullen. 1993. Recent advances in the organization and regulation of lignin peroxidase genes of *Phanerochaete chrysosporium*. J. Biotechnol. 30:109-114. The chromosomal location of a number of closely related lins was determined by Southern blot analysis of CHEF gels.

Orth, A., M. Rzhetskaya, D. Cullen and M. Tien. 1994. Characterization of a cDNA encoding a manganese peroxidase from *Phanerochaete chrysosporium*: genomic organization of lignin and manganese peroxidase genes. Gene (In press). This was a collaborative study with Tien and co-workers. In this laboratory, chromosomal location of several mnp1s was shown by Southern blot analysis of CHEF gels. The bulk of the work was performed at Penn State and supported by Tien's grants.

Blaha, A., D. Cullen, and K. Messner 1994. Aspects of the life-cycle of *Phanerochaete chrysosporium* strain BKM-F-1767. Mycological Research (In press) This was a collaborative study with Kurt Messner and co-workers. Our laboratory determined the nuclear condition of a large number of single spore isolates by monitoring the segregation of specific lip and glx alleles. Cytological investigations of nuclear behavior during basidial development were conducted in Vienna. Both homokaryotic and heterokaryotic basidiospores were shown to be produced and the latter were attributed to postmeiotic mitotic divisions. The results may explain the aberrant ratios sometimes encountered among single basidiospore progeny and they clearly exclude apomixis and homokaryotic fruiting for *P. chrysosporium*.

Gaskell, J., Stewart, P., Kersten, P., Covert, S. F., Reiser, J. and D. Cullen. 1994. Genetic mapping in the lignin-degrading Basidiomycete *Phanerochaete chrysosporium* by allele-specific polymerase chain reaction analysis (Submitted). We describe an experimental approach whereby probes are designed and used to directly monitor the segregation of specific lip alleles during sexual fruiting. Five linkage groups were identified. The 10 lignin peroxidase genes were distributed in 3 linkage groups. One co-segregating group contained 8 closely-linked LiP genes. Another lip was linked to a cellobiohydrolase gene cluster. The method establishes linkage relationships among genes for which there are no mutations, and it is applicable to a wide range of genes and organisms. An integrated genetic and physical map of *P. chrysosporium* has been developed. (More detailed results are presented in Section II.)

4. Insertional mutant

Gaskell, J., P. Stewart, A. Vanden Wymelenberg, and D. Cullen. 1994. Insertional mutation of *Phanerochaete chrysosporium* BKM-F-1767 lignin peroxidase allele lip12: structure,
inheritance, and expression. (In preparation). A novel allelic variant of \textit{lip1} has been characterized. Specifically, a 1745 bp insert is located at the 3' end of the fourth intron of \textit{lip1}. The entire sequence was determined. RT-PCR experiments show the insert is not spliced-out and so \textit{lip1}2 appears to be transcriptionally inactive. Several short open reading frames (<125 codons) were identified within the insert, but their transcription remains to be studied. None of the open reading frames show significant homology to known proteins. Short inverted repeats were detected at insert termini. The insert is not repetitive; there appear to be only two copies. CHEF gel Southern blots show that both copies are present on the same 3.7 mb chromosome band of BKM-F-1767. The mutated allele is inherited in a strict 1:1 Mendelian ratio and therefore may play a significant role in generating genetic variability. (More detailed results are presented in Section II.)

5. Other ligninolytic species

Larrain, J., S. Lobos, L. Salas, D. Cullen, and R. Vicuna. 1994. Isozymes of manganese-dependent peroxidase and laccase produced by the lignin-degrading basidiomycete \textit{Ceriporiopsis subvermispora}. Microbiology (formerly J. Gen. Microbiol.) (In Press). In collaboration with Vicuna and coworkers, we have purified and characterized the manganese peroxidases of the white-rot fungus \textit{C. subvermispora}. Distinct N-terminal sequences were determined for several isoymes suggesting the presence of several related genes. The expression of the isoymes differed substantially between defined media and woody substrates. Sequence homology to \textit{P. chrysosporium} MnPs was evident. These studies represent the first step toward characterizing the ligninolytic system of the most efficient biopulpung fungal species known. As described in the Proposed Research section, characterization of \textit{C. subvermispora} peroxidase clones has begun.

Moen, M.A., K. Hammel, H. Burdsall, and D. Cullen. 1994. ATCC 26195 is a ligninolytic soil Basidiomycete. (In preparation). This collaborative study shows that a widely used "\textit{Geotrichum}" strain is closely related to white-rot basidiomycetes. DOE funds supported cloning, sequencing, and analysis of full length 18S rRNA genes from \textit{P. chrysosporium} and from ATCC 26195. The study firmly established the taxonomic position of ATCC 26195, a strain used in a number of organopollutant studies. As the first reported white-rot 18S sequence, results will be useful to many ongoing studies of fungal phylogeny.

II. More detailed descriptions of unpublished data

1. Glyoxal oxidase studies

A condensed version of a manuscript on genomic structure and organization of \textit{glx} alleles and expression of \textit{glx} in \textit{A. nidulans} follows. Full length sequences of two \textit{glx} alleles were determined (Fig. 2). Typical of allelic pairs in \textit{P. chrysosporium}, the sequences are highly conserved with most mismatches occurring in untranslated regions or in the third positions of codons. A single amino acid difference is located at position 308. In close agreement with isoelectric focusing gels [8], the predicted pls of \textit{glx1} and \textit{glx2} are 5.1 and 5.0, respectively. Conclusive proof for allelism was obtained by analyzing single basidiospore cultures (Fig. 3). A variety of low stringency Southern blots of genomic DNA showed no evidence for any other 'glx-like' sequences. The ME446 \textit{glx} alleles were PCR amplified, and partial sequence (>650 bp) showed them to be nearly identical to the corresponding BKM-1767 gene (data not shown).

Southern blots of pulse field gels localized \textit{glx} to a dimorphic pair of chromosomes separate from all known peroxidases and cellobiohydrolases (Figure 4). Non-linkage between
glx and other markers was confirmed by genetic segregation studies (See genomic mapping section below 1.2).

When fused to the A. niger glucoamylase promoter and terminator (Fig. 5), efficient glx expression was obtained in A. nidulans. Fully active extracellular protein was obtained under maltose induction (Fig. 6). Yields of 10-20 mg/l are suitable for various biochemical investigations and for site-specific mutagenesis. For example, plasmid pGLAGLXD308 is identical to pGLAGLX except that a single base was mutated to encode Thr at residue 308. Thus, the protein encoded by pGLAGLXD308 (Fig. 6, lane 1) corresponds to the allelic variant glx2. The physical characteristics and substrate specificity of both allelic forms are under investigation.

2. Genetic mapping

A novel strategy for genetic mapping by segregation analysis has been developed and submitted for publication. We systematically identified nucleotide sequence divergence among allelic pairs of all known LiP genes (lip) as well as GLOX and CBH genes. Allelic segregation was monitored by PCR amplification of genomic DNA from single basidiospore cultures and then probing with allele-specific oligonucleotides. Owing to the dikaryotic nuclear condition and homothallic fruiting of P. chrysosporium strain BKM-1767, it was not necessary to cross strains; one set of single basidiospore progeny was adequate for all analyses.

**Probe Design.** To prepare allele-specific oligonucleotide probes, allelic pairs were PCR amplified, subcloned, sequenced and compared. Initial emphasis was on the LiP gene family because, prior to this study, considerable confusion existed concerning the total number and allelic relationships among LiP genes. Based on published sequences or database depositions (Table 1), gene-specific primers were designed (Table 2) and LiP genes were PCR amplified from the genomic DNA of 5 separate homokaryotic cultures. The PCR products were subcloned, partially sequenced and compared. Assuming 1:1 Mendelian segregation, it was highly probable (~95%) that each allele would be represented at least once among the 5 cultures. This was confirmed; two highly similar, but distinct, sequences were consistently observed.

The nucleotide sequences of allelic homologues were more highly conserved in coding regions relative to untranslated regions. Sequence divergence within coding regions was typically in the third position of codons, so that the predicted amino acid sequence was rarely affected. Mismatches were most commonly T-C transitions. Only one major sequence difference was detected; a 1745 bp insertion in a lip allele. Aside from this discrete insertion, however, the lip alleles exhibited extreme sequence conservation similar to that observed in all other allelic pairs.

Two to three hundred base pairs of sequence were adequate to design allele-specific 17-mer probes (Table 2). In several instances, e.g. lipE, multiple mismatches were included in the differentiating probes. Not unexpectedly, the targets for such probes tended to be within introns.

**Segregation analysis.** To determine genetic linkage, the segregation of specific alleles was monitored within a population of homokaryotic segregants. The strategy involved isolation and identification of homokaryotic single basidiospore cultures, rapid extraction of genomic DNA, PCR amplification of genes using the primers listed in Table 2, and differentiation of allelic alternatives with 32P-labeled oligonucleotide probes. Linkage was then computed from allelic co-segregation frequencies.

To illustrate the process, analysis of 9 selected single basidiospore cultures is shown (Fig. 7). Three genes, lipA, lipH, and lipF, were PCR amplified and probed with their corresponding oligonucleotides (Table 2). In the case of two closely-linked genes, lipA and lipH, the pattern of hybridization is very similar except for 1 of the 9 basidiospore cultures shown (Fig. 7, lane 9). In contrast, the hybridization pattern for an unlinked gene (lipF) is
quite different, illustrating the random segregation of alleles. As expected, PCR products from
the parental dikaryotic culture (Fig. 7, lane 10) hybridized to all probes.

The 12 P. chrysosporium genes analyzed were assigned to 5 linkage groups (Table 3). Eight of the LiP genes were tightly linked. Five of these, lipA, lipB, lipC, lipI, and lipI, were previously shown to hybridize to two distinct bands on Southern blots of Clamped Homogeneous Electrical Field (CHEF) gels [9]. One interpretation, now confirmed, had been that the two bands represented chromosomal homologues corresponding to a single linkage group. Of the three new genes (lipE, G, H) mapped to this same linkage group, lipE may be most significant because of its high expression levels [10, 11]. Within this same linkage group, a single recombination event was detected between lipA and lipC. Cosmid and λ maps show lipA immediately adjacent to lipB and only 15 kb from lipC [9, 12].

lipD and lipF were unlinked to all other LiP genes. Previously, separate studies had localized lipD and cbh1-3 to a large dimorphic chromosome of approximately 4.4 and 4.8 mb[13, 14]. The segregation data is entirely consistent with these earlier CHEF gel blots; lipD is unlinked to all other known LiP genes but linked to cbh1-3. Recently, D'Souza et al [15] reported lipF hybridized to a relatively small doublet on Transverse Alternating Field Electrophoresis (TAFE) gels. We obtained a similar hybridization result on CHEF gel blots (data not shown). An integrated physical and genetic map is shown in Figure 8.

Overall, a relationship between genomic organization and transcriptional regulation is not obvious. Numerous studies have shown that LiP gene expression is derepressed under nutrient limitation and that certain genes respond differentially to carbon versus nitrogen starvation [10, 13, 16]. Of the 8 linked LiP genes, transcript levels of 5 have been studied in both C- and N-starved cultures. Transcript levels of 3 of these (lipA, B and I) differ substantially from each other but remain unchanged in C- versus N-starved cultures. In contrast, lipC and lipI mRNAs are at least 100-fold more abundant in N-starved media relative to C-starved media. Glyoxal oxidase, expressed in C- and N-starved cultures [13, 17], is unlinked (Table 3; Fig. 8).

Independent segregation by lipD and lipF may reflect physiological and evolutionary significance. Expression of lipD is uniquely and dramatically increased by carbon limitation [13, 16, 18], and it may be relevant that lipD is linked to cbh1-1 and cbh1-2, transcripts of which increase in response to carbon limitation [19]. Regulation of lipF expression has not been thoroughly investigated. Interestingly, when intron/exon criteria are used for delineating LiP genes into four groups, lipD and lipF are assigned separate subfamilies [11, 20]. Clustal analyses [21] of deduced LiP amino acid sequences clearly show lipD to be distinct from lipF and the other LiP genes.

The mapping strategy described here is generally applicable to cloned DNA from any organism from which haploid products of meiosis are available. For basidiomycetes such as P. chrysosporium, crosses and parental strains are unnecessary. Only homokaryotic basidiospores are required. Many fungi, although unable to fruit in culture, produce abundant viable basidiospores in nature. Using the approach described here, basidiocarps collected from the field or herbarium samples could provide the spores needed to construct genetic linkage maps. Other organisms in which haploid products of meiosis are readily available include a number of important conifer species. For example, Tulsieram et al [22] have exploited this characteristic to construct a RAPD map of white spruce.

The method can be modified to suit a variety of experimental situations. The targeted markers need not be specific structural genes. Degenerate primers amplifying highly conserved genes, or random PCR products could also be used. Allele-specific probes could be tagged with fluorescent dyes, or perhaps eliminated altogether by choosing suitably selective PCR primers. Finally, the number of PCR reactions might be substantially reduced by simultaneously amplifying several genes of varying lengths.

LiP gene nomenclature. Like other complex gene families, considerable uncertainty exists as to total number and structure of LiP genes. For example, identical upstream regions were recently reported in six LiP "genes" [23], but our analysis shows that these six clones are actually the same gene or allelic variants thereof. The number of lignin
peroxidase genes in *P. chrysosporium* had been variously set from 5 [15] to 15 [23]. Contributing to this confusion are the widespread use of two distinct laboratory strains (BKM-F-1767 & ME446), the presence of allelic variants within these dikaryotic strains, the publication of partial sequences with or without database deposition, inconsistent nomenclature for clones/genes, and separate publication of highly similar or identical sequences. In the process of mapping LiP genes allelic relationships were established. On that basis, we include a uniform nomenclature which may reduce confusion in the future (Table 1). The system arbitrarily employs letters to designate genes and numbers to assign alleles (Table 1; column 1). This system makes no attempt to assign specific isozymes to these genes. A working assumption is that strains feature two alleles of each gene; these are labeled 1 and 2 for BKM-1767 alleles, 3 and 4 for ME446 alleles. New genes and alleles can be designated with alphabetical and numerical suffixes, respectively.

3. lip1*2* insertional mutant

In the course of genetic mapping studies, an unusual mutant allele of *lip1* was observed. Specifically, a 1745 bp insertion was located adjacent to the 3' end of the fourth intron of *lip1* (Fig. 9, top). Southern blots of genomic DNA show that the mutated allele segregates from the wild type *lip1l* in homokaryotic cultures (Fig. 9, bottom). No sequences with strong homology to the insert were detected in ME446 by cross hybridization (Fig. 9). A second band in restricted BKM-F-1767 DNA suggested the presence of one structurally similar sequence in the genome (Fig. 9; 4.6 kb band). This additional 'copy' must be linked and cis to *lip1l* because it cosegregates on Southern blots of single basidiospore genomic DNA (Fig. 9, right panel; lanes 2 and 4) and hybridizes to a single band on CHEF gel blots (Fig. 10).

Full length sequence of the insert and most of the surrounding LiP-like region have been determined (Fig. 12). Open reading frames within the insert and through adjacent regions are relatively small, i.e. <120 codons. Predicted amino acid sequences from these ORFs showed no significant homology to known proteins (NBRF-PIR release 81; Swiss-Prot release 27). The region surrounding the insert is much like a typical lip allele, exhibiting extreme sequence conservation with *lip1l*. No extended repeats (direct or inverted) were observed, but a short inverted terminal repeat was detected (Fig. 12):

1 ATCAGCAATTTCCCTTAG-CGGATTCTTTACAGGCCTTTT 35
ATCAGCAAT- TCC- AG- C- G- TCTT- CAGGC- TTT
2150 ATCAGCAATATCCCAAGACC-GTCTGCAGGCC-TTT 2116

RT-PCR amplification with *lip1*-specific primers suggests that *lip1l* is not transcribed (Fig 11). In contrast, *lip1l* transcripts are detected in both N-limited and C-limited cultures. Thus, if *lip1l* is transcribed, it is not likely to involve splicing of the insert. Transcription of potential insert ORFs under a variety of cultural conditions, and *lip1l* transcription in single basidiospore cultures are being evaluated (see proposed research section). The distribution of the insert among *P. chrysosporium* strains is also being assessed.

The mutated allele is inherited in a simple Mendelian fashion. Sixty nine single basidiospore cultures were examined, and *lip1l* and *lip1l2* segregated at 1:1 (P=0.01 by χ2). Cosegregation data (Table 3) and CHEF gel blots (Fig. 10) show the *lip1l* locus to be very closely linked to three other LiP genes, *lipG*, *lipH*, and *lipI*. The insertion appears to be mitotically stable, in that both *lip1l* and *lip1l2* are consistently maintained in asexual conidia (data not shown). The role of *lip1l2* in ligninolytic variability remains to be established.