Identification of a cosmid clone containing the *Neurospora crassa lys-5* and *un-4* genes, isolation of a partial *lys-5* cDNA and associated chromosome walking.

Thomas J. Schmidhauser, Dan Chen and Yihong Wan* - Department of Biology, The University of Southwestern Louisiana, Lafayette, LA and *University of Colorado Health Sciences Center, Denver, Colorado

The *un-4* gene of *Neurospora crassa* was cloned to determine the limits of a chromosome walk on linkage group VI (LGVI) and to allow analysis of *un* loci on LGVI. Subsequent analysis identified the *lys-5* locus on the same cosmid clone as *un-4*. We have isolated and sequenced a partial *lys-5* cDNA clone and initiated a chromosome walk from the *lys-5*, *un-4* cosmid clone.

A chromosome walk from the *cpc-1* locus has been extended 420 kb towards the left telomere of linkage group VI, (LGVIL, Wan *et al.* 1997 Fungal Genet. Biol. **21**:329-336). One of three heat-sensitive loci of unknown function on LGVI, *un-13*, was found in the *cpc-1* walk. The *un-4* locus maps to LGVIL. Three rounds of transformation using sib-selection with cosmid DNA pools from the Orbach/Sachs *Neurospora crassa* genomic library identified an *un-4*$^{+}$ cosmid, G13:8:G, by selection for transformants able to grow at the restrictive temperature of 34°C. A 1.2-kb cDNA isolate from a cDNA library (based on mRNA isolated from dormant conidia and kindly provided by M. Sachs), designated pYW19-2, was identified using a G13:8:G insert probe.

DNA sequence analysis of pYW19-2 identified an open reading frame encoding a deduced polypeptide with strong similarity to homocitrate synthases and isopropylmalate synthases from other organisms (Figure 1). Neurospora *lys-5* mutants lack homocitrate synthase activity. G13:8:G DNA complements *lys-5* spheroplasts allowing growth on minimal medium. *lys-5* maps 2% away from *un-4* and *un-4*, by definition, is irreparable by supplementation at the restrictive temperature. Thus, *un-4*$^{+}$ and *lys-5*$^{+}$ are separate loci and both are present in G13:8:G. pYW19-2 likely represents a partial *lys-5* cDNA clone. The partial deduced Lys-5 polypeptide has highest similarity to the homocitrate synthase of *Penicillium chrysogenum* with 80% identity in an optimized alignment (Figure 2).

```
   1 CGTTATTGAG TATGTCAAGT CCAAGGGACT TGAGGTTCGC TTCTCCTCCG AGGATTCCTT
     V  I  E  Y  V  K  S  K  G  L  E  V  R  P  S  S  E  D  S  F
  61 CCGCTCCGAT CTCGTCGATC TCCTTTCCCT TTACCGCGCT GTTGACAAGG TCGGCGTCCA
     R  S  D  L  V  D  L  L  S  L  Y  R  A  V  D  K  V  G  V  H
 121 CCGTGTCGGT ATCGCCGATA CCGTCGGCTG CGCTTCTCCC CGCCAGGTCT ATGACCTCGT
     R  V  G  I  A  D  T  V  G  C  A  S  P  R  Q  V  Y  D  L  V
 181 CCGTACCCTT CGCGGCGTCG TTTCGTGCGA TATCGAGACC CACTTCCACG ACGACACTGG
     R  T  L  R  G  V  V  S  C  D  I  E  T  H  F  H  D  D  T  G
 241 CTGCGCCATT GCCAACGCCT ACTGTGCTCT CGAGGCTGGT GCCACCCACA TCGACACCTC
     C  A  I  A  N  A  Y  C  A  L  E  A  G  A  T  H  I  D  T  S
 301 CGTTCTCGGT ATCGGCGAGC GTAACGGTAT CACCCCTCTC GGTGGCTTGA TGGCTCGCAT
     V  L  G  I  G  E  R  N  G  I  T  P  L  G  G  L  M  A  R  M
 361 GATCGTTACC AGCCCCGACT ACGTCAAGAG CAAGTACAAG CTCCACAAGC TCAAGGAGCT
     I  V  T  S  P  D  Y  V  K  S  K  Y  K  L  H  K  L  K  E  L
 421 CGAGGATTTG GTTGCCGAGG CTGTTGAGAT CAACACCCCC TTCAACAACC CCATCACTGG
     E  D  L  V  A  E  A  V  E  I  N  T  P  F  N  N  P  I  T  G
 481 TTTCTGCGCC TTCACCCACA AGGCTGGCAT CCACGCCAAG GCCATCCTCA ACAACCCCAG
     F  C  A  F  T  H  K  A  G  I  H  A  K  A  I  L  N  N  P  S
 541 CACCTATGAA ATTCTCAACC CTGCCGACTT CGGTCTCACC CGCTACGTCC ACTTCGCTTC
     T  Y  E  I  L  N  P  A  D  F  G  L  T  R  Y  V  H  F  A  S
 601 GCGCTTGACT GGCTGGAACG CCGTCAAGAC CCGTGTCGGC CAGCTTGGTC TCGAGATGAC
     R  L  T  G  W  N  A  V  K  T  R  V  G  Q  L  G  L  E  M  T
 661 CGACGACCAG GTCAAGGAAT GTACCGCCAA GATCAAGGCC CTTGCCGACG TGCGCCCAAT
     D  D  Q  V  K  E  C  T  A  K  I  K  A  L  A  D  V  R  P  I
 721 CGCCATCGAC GACGCCGATT CGATCATCCG TACTTTCCAC CTCGGTCTTC ACGAGCAGAA
     A  I  D  D  A  D  S  I  I  R  T  F  H  L  G  L  H  E  Q  N
 781 CAAGGTCCAG CCTCCCGCTG TTGTCGAGAA CTAAGCGGAA GCAGAGCGTT CGACCAACGG
     K  V  Q  P  P  A  V  V  E  N  *
 841 AGTTGTCCTT TAGCATGAAG GGGAATATAC CAGGATTTTT ACGAGGAGAG ATGCGGGCAT
 901 CATGACGATT TTCTTTTTAC TTGTGTTTGG GGTCATTTTT CACACATCCA CCGGAGTTCT
 961 TTGAGTACTA TAATCTCCCT GTTTGGGGAG CAAAAAGGGG GTTGATTGGG TTAACTGGGG
1021 ATGACTGAGC AGGCCAATAT TGCCGACTGT GTTCCTAATC AGGGGGAATG CTCGTCGAAA
1081 AATGAGCATG AGATAGACAA AATCAACGGG AGACGAAAGT AACAACGTCA CCTGATTGTC
1141 CTTCAAAAAA AAAAAAAAAA AA
```

Figure 1. Nucleotide sequence of the cDNA insert of pYW19-2 and the deduced polypeptide product (GenBank AF142777). The stop codon is indicated by a *. Several isolates, including NC4A2-T7, from the Neurospora Genome Project, University of New Mexico, overlap pYW19-2 from position 549 to the polyadenylation site.

The pYW19-2 insert was used to probe a Southern blot of G13:8:G restriction digests. Results suggest that an approximately 6.3-kb *Eco*R1 fragment contains the *lys-5* gene. As cosmid G13:8:G was not identified in the *cpc-1* walk we initiated a chromosome walk from the *lys-5/un-4* region in an attempt to link up to our *cpc-1* walk. A G13:8:G based probe identified cosmid X6:6. A X6:6:F based probe identified cosmid X22:2:B. No new cosmid clones were identified with a X22:2:B based probe.

```
Penicillium c. HC Synthase    181  IEVIEFVKSK GIEIRFSSED
                                        |||+|||| |+|+||||||
Neurospora Lys-5                        VIEYVKSK GLEVRFSSED

Pc HC   SFRSDLVDLL SIYSAVDKVG VNRVGIADTV GCASPRQVYE LVRVLRGVVG
        |||||||||| |+| |||||| | |||||||| |||||||||+ |||+|||||+
Lys-5   SFRSDLVDLL SLYRAVDKVG VHRVGIADTV GCASPRQVYD LVRTLRGVVS

Pc HC   CDIETHFHND TGCAIANAFC ALEAGATHID TSVLGIGERN GITPLGGLMA
        ||||||||+| ||||||||+| |||||| ||| |||||||||| ||||||||||
Lys-5   CDIETHFHDD TGCAIANAYC ALEAGAYHID TSVLGIGERN GITPLGGLMA

Pc HC   RMMVADREYV KSKYKLEKLK EIEDLVAEAV EVNIPFNNYI TGFCAFTHKA
        ||+|   +|| |||||| ||| |+|||||||| |+| |||| | ||||||||||
Lys-5   RMIVTSPDYV KSKYKLHKLK ELEDLVAEAV EINTPFNNPI TGFCAFTHKA

Pc HC   GIHAKAILNN PSTYEIINPA DFGMSRYVHF ASRLTGWNAI KSRAQQLKLE
        |||||||||| ||||||+||| |||++|||| |||||||||+ |+|+ || ||
Lys-5   GIHAKAILNN PSTYEILNPA DFGLTRYVHF ASRLTGWNAV KTRVGQLGLE

Pc HC   MTDTQYKECT AKIKAMADIR PIAVDDADSI IRAYHRNLKS GENKPLLDLT
        ||| | |||| |||||+||+| |||+|||||| || +|  |+  +      ++
Lys-5   MTDDQVKECT AKIKALADVR PIAIDDADSI IRTFHLGLHE QNKVQPPAVV

Pc HC   AEEQAAFAAK EKELLEAQAA GLPV
         +
Lys-5   EN
```

Figure 2. Comparison of the amino acid sequence of the homocitrate synthase of *Penicillium chrysogenum* (Gene Bank AJ223630) and Lys-5. Identical residues are indicated by a vertical line. Residues of similar chemical properties are indicated by a +. A motif conserved in all known homocitrate synthases is underlined.