Note

Discrepancies Between Recombination Frequencies and Physical Distances in *Aspergillus nidulans*: Implications for Gene Identification

Eduardo A. Espeso,* Laura Cobeño* and Herbert N. Arst, Jr.[†,1]

*Departamento de Microbiologia Molecular, Centro de Investigaciones Biologicas, CSIC, Madrid 28040, Spain and †Department of Infectious Diseases, Faculty of Medicine, Imperial College London, London W12 0NN, United Kingdom

ABSTRACT

A rapid route to gene molecular identification involves using recombination frequencies in locating mutational sequence changes. We describe a case where the recombination frequency is deceptively low, probably reflecting centromere proximity. Recombination frequencies are greatly reduced near the centromeres on the right arms of chromosomes III and IV of *Aspergillus nidulans*.

WITH the advent of whole-genome sequencing, an attractive route for identifying the sequence of a gene for which a mutant allele exists involves mapping the mutation and sequencing plausible candidate open reading frames (ORFs) to locate the mutational sequence change. Sometimes, however, recombination frequencies can be deceptive, as in the case described here.

As part of a project to investigate nuclear localization of the PacC transcription factor of the ascomycete *Aspergillus nidulans*, we sought spontaneous mutations that suppress the molybdate hypersensitivity phenotype of a *pacC* allele whose translation product has a K159M mutational change inactivating the PacC nuclear localization signal (described by FERNÁNDEZ-MARTINEZ *et al.* 2003). We selected for tolerance of 20 mM Na_2MoO_4 on supplemented minimal medium (COVE 1966). Among the mutations obtained, one, designated *halA24*, had no effect on PacC nuclear localization or on other aspects of the highly pleiotropic (see reviews by PEÑALVA and ARST 2002, 2004) mutant *pacC* allele phenotype apart from molybdate hypersensitivity. However, it has an otherwise interesting phenotype (to be described elsewhere). Using conventional parasexual and sexual mapping techniques (CLUTTERBUCK 1974) facilitated by the molybdate resistance of *halA24 pacC*+ strains, *halA24* was located between *phenA* and *cbxA/carC* on the right arm of chromosome III. Two preliminary crosses,

with 99 and 100 progeny analyzed, respectively, gave recombination frequencies of 1 cM to *phenA2* and 13 and 17 cM, respectively, to *cbxA17*. Using the *A. nidulans* genome sequence (http://www.broad.mit.edu/annotation/fungi/aspergillus) to identify likely ORFs, mutational sequence changes were determined for identifying definitively the *phenA* and *cbxA/carC* genes (Table 1). Following failure to find the *halA24* sequence change in plausible candidate ORFs near *phenA* on the *cbxA* side, it was eventually located at >35% of the physical distance between *phenA* and *cbxA* (Figure 1). The gene identity has been firmly established and its characterization will be described in detail elsewhere. Its predicted product is very likely to be a Ser/Thr protein kinase (http://www.sanger.ac.uk/Software/Pfam; *E*-value 5.7*e*-78) and *halA24* tandemly duplicates four residues within the predicted kinase domain (Table 1).

To better document the highly skewed relationship between recombination frequency and physical distance in the *phenA-cbxA* interval, 392 random progeny were analyzed from the cross *biA1*; *phenA2 cbxA17* × *halA24*; *inoB2*; *fwA1* (see CLUTTERBUCK 1993 for gene symbols), giving 1.8 ± 0.7 cM between *phenA2* and *halA24* and 13.5 ± 1.7 cM between *halA24* and *cbxA17*, a 7.5-fold difference. The physical distances are 67,765 and 124,414 bp, respectively (Table 1 and Figure 1), a 1.84-fold difference. Between *phenA2* and *halA24* 1 cM corresponds to 37.6 kb and between *halA24* and *cbxA17* 1 cM corresponds to 9.2 kb. To ensure that this difference is a genuine feature of relative recombination frequencies in these two intervals and not simply the consequence of a previously unrecognized chromosomal rearrangement(s), the two parental strains from this cross were first shown to be translocation-free using the

Gene	Database notation	Allele	Nucleotide change	Amino acid change	Mutant protein
phenA	NA[a]	*phenA2*	A100G[d]	IVS1[g]	1–18 + 15
		phenA4	A100G	IVS1[g]	1–18 + 15
		phenA5	859insTT	F268fs[h]	1–267 + 19
		phenA7	G29A	G10D	G10D
		phenA8	G421A	W121[i]	1–120
halA	AN8830.2	*halA24*	1566insATCAACTCGTCC	insINSS	1–503 + INSS + 504–524
cbxA/carC	AN8793.2[b]	*cbxA17*	A684G	H156R	H156R
		carC9	T483C	L89P	L89P
suaB	Gln-tRNA[c]	*suaB111*	G145,904A[e]	—	—
hisA	AF159463 AN7430.2	*hisA10*	T973C	L296S	L296S
gdhB	AY223544 AN7451.2	*gdhB1*	ΔT1014-T1023	M318fs	1–317 + 12
meaA	AY049706 AN7463.2	*meaA8*	C855T[f]	Q231[i]	1–230

[a] Gene *not annotated* by the database. Predicted product is prephenate dehydratase; cDNA was sequenced.

[b] Predicted product is cytochrome B-560 subunit of succinate dehydrogenase; cDNA was sequenced.

[c] Predicted glutamine tRNA, anticodon CUG. Start is at 145,869, end at 145,954. Contig 1.161. Amber suppressor. G-to-A change in the anticodon sequence in *suaB111*.

[d] Nucleotide coordinates are based on the coding region in genomic DNA.

[e] Contig 1.161 coordinate; change in Figure 1 coordinate is G283,287A.

[f] Our nucleotide sequence change corrects that given by MONAHAN *et al.* (2002).

[g] Intron lariat sequence was altered; splicing defect is likely. The predicted protein would terminate after translation of the (mutant) intron and the first eight nucleotides of exon 2.

[h] fs, frameshift.

[i] Stop codon.

parasexual cycle. Second, a combination of Southern blotting and PCR fragment size determination demonstrated the absence of any significant rearrangement in either parental strain in a region of 96,747 bp extending from 17,847 bp centromere-proximal to *phenA2* to 11,134 bp centromere-distal to *halA24* (data not shown). It is possible that the recombination frequency continues to rise beyond *cbxA*. SEALY-LEWIS (1987) determined a map distance of 5.1 ± 2.2 cM between *cbxA17* and *suaB111*. Using the physical data in Figure 1 and Table 1 to calculate 38,038 bp between these mutations, this would equate to 7.5 kb/cM. Beyond *suaB*, the recombination frequency might be even higher in view of the 12.5-kb distance between *suaB* and *amdS* coding regions (Figure 1) and the *suaB111–amdS1350* 8.2 cM map distance (SEALY-LEWIS 1987), which would result in a recombination frequency <2 kb/cM.

The most obvious explanation for this markedly non-linear relationship between recombination frequency and physical distance is the proximity of the centromere. Reduced recombination in the vicinity of the centromere was reported first in *Drosophila melanogaster* (DOBZHANSKY 1930) and subsequently in other organisms, a particularly elegant demonstration being that of LAMBIE and ROEDER (1986) in *Saccharomyces cerevisiae*. In *A. nidulans* a translocation shifting a region near the chromosome II centromere to a much more

centromere-distal position on chromosome VIII increases the recombination frequency between two of the translocation-shifted genes by an order of magnitude (ARST *et al.* 1979). Arguably the most dramatic effect involves the silent mating-type loci of *Schizosaccharomyces pombe* where no recombination has been detected at a resolution of 0.001 cM in a 15-kb interval sharing homology with centromeric repeats (EGEL *et al.* 1989; GREWAL and KLAR 1997 and references therein). All eight *A. nidulans* centromeres lie within gaps in the genomic sequence, but from the contiguous right arm sequence of chromosome III the minimum distance between *phenA2* and the centromere is 53,065 bp.

A striking centromere effect on recombination frequency can also be seen on the right arm of chromosome IV, as previously noted by ALEKSENKO *et al.* (2001). With a gene order of centromere, *hisA/hisHF*, *gdhB*, and *meaA* and using mapping data from ARST *et al.* (1989) in conjunction with the genome sequence (see Figure 1) and mutational sequence changes (Table 1), the distances between *hisA10* and *gdhB1* are 1.2 ± 0.3 cM and 62,103 bp and those between *gdhB1* and *meaA8* are 5.9 ± 1.3 cM and 33,641 bp. Between *hisA10* and *gdhB1* there are 51.8 kb/cM, in good agreement with 49.3 kb/cM between *hisA134* and *gdhB1* reported by ALEKSENKO *et al.* (2001), and between *gdhB1* and *meaA8* 1 cM corresponds to 5.7 kb. The higher rate of recombination
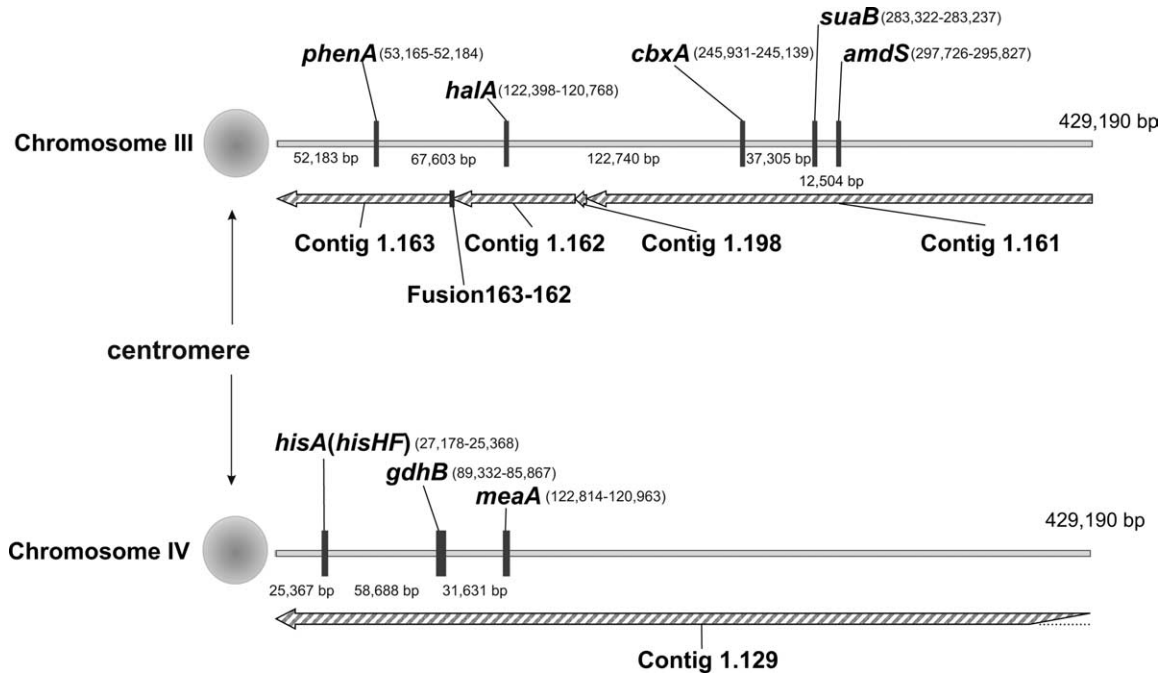
FIGURE 1.—Physical maps of the centromere-proximal regions of the right arms of chromosomes III and IV. Contig numbering follows that of the Broad Institute (http://www.broad.mit.edu/annotation/fungi/aspergillus/). Distances between genes or to the end of the available genomic sequence are given. Coordinates following the gene name indicate the coding region with that of the first base of the initiation codon appearing first. Contigs 1.161 and 1.162 were joined by alignment with contig 1.198. To join contigs 1.162 and 1.163, genomic DNA was PCR amplified using oligonucleotides annealing at a distance of 170 nt from the end of 1.162 and the beginning of 1.163. Sequencing of the PCR fragment revealed database errors at the ends of both contigs (a consequence of shotgun cloning) and, after deletion of a total of 36 incorrect nt, a 34-nt addition was necessary to join the two contigs. The chromosome III contigs thus joined in total 429,190 bp, but only 429,190 bp of the 627,835 bp of contig 1.129 from chromosome IV are indicated.

between *phenA2* and *halA24* as compared to that between *hisA10* and *gdhB1* might be explained by the minimum distance of *phenA2* to the centromere being approximately twice that of *hisA10* to the centromere (53,065 bp *vs.* 26,205 bp). Caution is needed, however, because the actual distances to the centromere have not been determined. Figure 2 is a summary showing relationships between recombination frequencies and physical distances on the right arms of chromosomes III and IV.

In the above cases it is very likely that centromere proximity is responsible for greatly reduced recombination frequencies. However, when recombination frequencies are used in conjunction with genome sequences for gene identification, it should be remembered that other factors can also affect the relationship between recombination frequency and physical distance. For brevity we give a few examples from fungi but the factors involved affect recombination frequencies across a wide range of organisms. Such factors include environmental influences such as nutritional state (*e.g.,* ABDULLAH and BORTS 2001), genetic factors including both *trans*-acting genes (*e.g.,* CATCHESIDE 1981; YEADON *et al.* 2004) and *cis*-acting recombination hotspots (*e.g.,* PRYCE *et al.* 2005; STEINER and SMITH 2005), and the long-known effects of chromosomal

rearrangements (reviewed by KÄFER 1977 and PERKINS 1997). These effects can be considerable. For example, effects of more than an order of magnitude have been observed for both *trans*- and *cis*-acting genetic factors (CATCHESIDE 1981; STEINER and SMITH 2005), and
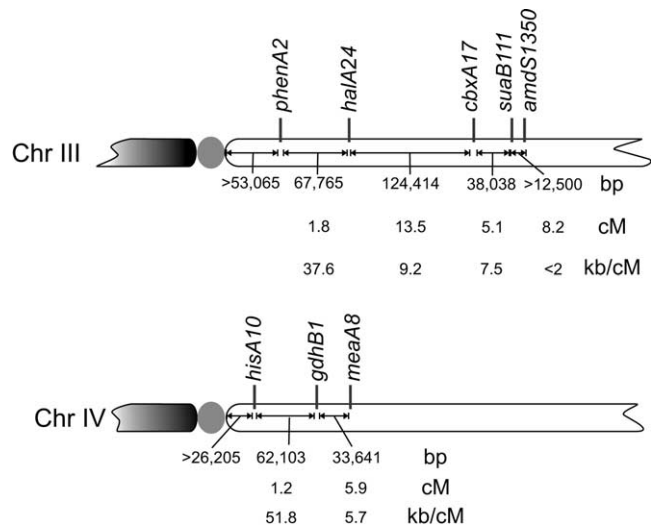

FIGURE 2.—Summary indicating meiotic map distances and relationships between recombination frequencies and physical distances.

there is evidence that, in the wild, selection operates on recombination frequencies (*e.g.,* Saleem *et al.* 2001). Judicious choice of strains and crossing conditions and, *in extremis,* molecular screening for chromosomal rearrangements can facilitate the use of recombination data for gene identification.

LITERATURE CITED

Abdullah, M. F. F., and R. H. Borts, 2001   Meiotic recombination frequencies are affected by nutritional states in *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA **98:** 14524–14529.

Aleksenko, A., M. L. Nielsen and A. J. Clutterbuck, 2001   Genetic and physical mapping of two centromere-proximal regions of chromosome IV in *Aspergillus nidulans.* Fungal Genet. Biol. **32:** 45–54.

Arst, H. N., Jr., K. N. Rand and C. R. Bailey, 1979   Do the tightly linked structural genes for nitrate and nitrite reductases in *Aspergillus nidulans* form an operon? Evidence from an insertional translocation which separates them. Mol. Gen. Genet. **174:** 89–100.

Arst, H. N., Jr., D. Tollervey and M. X. Caddick, 1989   A translocation associated loss-of-function mutation in the nitrogen metabolite repression regulatory gene of *Aspergillus nidulans* can revert intracistronically. Mol. Gen. Genet. **215:** 364–367.

Catcheside, D. E. A., 1981   Genes in Neurospora that suppress recombination when they are heterozygous. Genetics **98:** 55–76.

Clutterbuck, A. J., 1974   *Aspergillus nidulans,* pp. 447–510 in *Handbook of Genetics, Vol. 1: Bacteria, Bacteriophage and Fungi,* edited by R. C. King. Plenum Press, New York.

Clutterbuck, A. J., 1993   *Aspergillus nidulans,* pp. 3.71–3.84 in *Genetic Maps, Vol. 3: Locus Maps of Complex Genomes,* Ed. 6, edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Cove, D. J., 1966   The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans.* Biochim. Biophys. Acta **113:** 51–56.

Dobzhansky, T., 1930   Translocations involving the third and fourth chromosomes of *Drosophila melanogaster.* Genetics **15:** 347–399.

Egel, R., M. Willer and O. Neilsen, 1989   Unblocking of meiotic crossing-over between the silent mating-type cassettes of fission yeast, conditioned by the recessive, pleiotropic mutant *rik1.* Curr. Genet. **15:** 407–410.

Fernández-Martinez, J., C. V. Brown, E. Díez, J. Tilburn, H. N. Arst, Jr. *et al.,* 2003   Overlap of nuclear localisation signal and specific DNA-binding residues within the zinc-finger domain of PacC. J. Mol. Biol. **334:** 667–684.

Grewal, S. I. S., and A. J. S. Klar, 1997   A recombinationally repressed region between *mat2* and *mat3* loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. Genetics **146:** 1221–1238.

Käfer, E., 1977   Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. Adv. Genet. **19:** 33–131.

Lambie, E. J., and G. S. Roeder, 1986   Repression of meiotic crossing over by a centromere (CEN3) in *Saccharomyces cerevisiae.* Genetics **114:** 769–789.

Monahan, B. J., S. E. Unkles, I. T. Tsing, J. R. Kinghorn, M. J. Hynes *et al.,* 2002   Mutation and functional analysis of the *Aspergillus nidulans* ammonium permease MeaA and evidence for interaction with itself and MepA. Fungal Genet. Biol. **36:** 35–46.

Peñalva, M. A., and H. N. Arst, Jr., 2002   Regulation of gene expression by ambient pH in filamentous fungi and yeasts. Microbiol. Mol. Biol. Rev. **66:** 426–446.

Peñalva, M. A., and H. N. Arst, Jr., 2004   Recent advances in the characterization of ambient pH regulation of gene expression in filamentous fungi and yeasts. Annu. Rev. Microbiol. **58:** 425–451.

Perkins, D. D., 1997   Chromosome rearrangements in *Neurospora* and other filamentous fungi. Adv. Genet. **36:** 239–398.

Pryce, D. W., A. Lorenz, J. B. Smirnova, J. Loidl and R. J. McFarlane, 2005   Differential activation of *M26*-containing meiotic recombination hot spots in *Schizosaccharomyces pombe.* Genetics **170:** 95–106.

Saleem, M., B. C. Lamb and E. Nevo, 2001   Inherited differences in crossing over and gene conversion frequencies between wild strains of *Sordaria fimicola* from "Evolution Canyon." Genetics **159:** 1573–1593.

Sealy-Lewis, H. M., 1987   Suppressor specificity in *Aspergillus nidulans.* Curr. Genet. **12:** 141–148.

Steiner, W. W., and G. R. Smith, 2005   Optimizing the nucleotide sequence of a meiotic recombination hotspot in *Schizosaccharomyces pombe.* Genetics **169:** 1973–1983.

Yeadon, P. J., F. J. Bowring and D. E. A. Catcheside, 2004   Alleles of the hotspot *cog* are codominant in effect on recombination in the *his-3* region of Neurospora. Genetics **167:** 1143–1153.

Communicating editor: J. J. Loros