Comparative Sequence Analysis of the *qa-2* gene of *Neurospora crassa* and *Neurospora africana*

by

James Shevchuk

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In the

Department of Biological Sciences

Program

Youngstown State University

August 2004

Comparative Sequence Analysis of the *qa-2* gene of *Neurospora crassa* and *Neurospora africana*

James Shevchuk

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

James Shevchuk 8/3/04

James Shevchuk, Student Date

Approvals:

David Asch 8/3/04

Dr. David Asch, Thesis Advisor Date

Gary R. Walker 8/3/04

Dr. Gary Walker, Committee Member Date

Diana Fagan 8/5/04

Dr. Diana Fagan, Committee Member Date

Peter J Kasvinsky 8/4/04

Peter J Kasvinsky, Dean of Graduate Studies Date

Abstract

Gene systems like the quinic acid (*qa*) gene cluster have been studied in *Neurospora crassa* for many years. However, we know very little about gene systems in the homothallic species of *Neurospora*. Earlier it had been observed that *N. crassa* probes containing the *qa* gene cluster would hybridize to sequences in various homothallic species. To learn more about the *qa* systems in the homothallic species of *Neurospora* we have cloned the *qa* gene cluster from *Neurospora africana*. From these clones we have isolated and sequenced the *qa*-2 gene and compared it with the *qa-2* gene sequence of *N. crassa* and the sequence of the *qut-E* gene of *Aspergillus nidulans*.

TABLE OF CONTENTS

Introduction **1**

I. Kingdom Fungi 1

II. History of *Neurospora* 2

III. Gene regulation: Prokaryote versus Eukaryote 5

IV. Quinic Acid 6

V. The Galactose (*GAL*) pathway 6

VI. *GAL* Pathway Regulation 9

VII. *Qa* gene cluster organization and function 14

VIII. Comparison between the Quinic Acid (*qa*) Gene Cluster of *Neurospora crassa* and the Quinic Acid Utilization (*qut*) Gene Cluster of *Aspergillus nidulans* 24

IX. Homothallic versus Heterothallic 26

Materials and Methods **28**

I. Materials 28

II. Strains 29

III. Bluescript 30

IV. Phage Titration 30

V. Phage Recovery 31

VI. Lambda DNA Preparation 31

VII. Gel Electrophoresis 33

VIII. Southern Blot Digests 33

IX. Making Probe 33

X. Southern Blot 34

XI. Prehybridization and Hybridization 35

XII. Large Scale Isolation of Plasmid DNA 35

XIII. Isolation and Digestion of Vector 37

XIV. Isolation and Digest of Fragment 37

XV. Phenol Chloroform Extraction and DNA Precipitation 38

XVI. CIP and Ligation 38

XVII. Transformation 39

XVIII. Alkaline Plasmid Screen 39
XIX. Sequencing Reactions 40
XX. Comparative Analysis 41

Results **42**

I. Identification of DNA fragments containing the various *qa* genes from the homothallic speices of *Neurospora* 42
II. Subcloning of fragments containing *qa* genes from *N. africana* and *N. terricola* 51
III. Sequencing of the plasmid subclones pNAL, pNTL and pNTU 66
IV. Alignment of *Neurospora crassa qa-2* and *Neurospora africana qa-2* 66
V. Comparison of the *qa-2* protein of *N. africana*, the *qa-2* protein of *N. crassa*, and the *qutE* protein of *Aspergillus nidulans* 73
VI. Blast Search using *qa-2* protein sequences 78
VII. Sequence of entire Insert 78

Discussion **84**

Bibliography **89**

LIST OF FIGURES

1. Life cycle of *Neurospora crassa* 3

2. Representation of the galactose pathway 7

3. Illustration of possible mechanisms of catabolite repression 12

4. Quinic Acid gene organization in *Neurospora crassa* 14

5. Metabolic pathway of quinic acid and shikimic acid via the *qa* gene cluster 20

6. Sequences of the activator binding sites 22

7. Locations of the activator binding sites 26

8. Test gel giving insight to the quality and quantity of λNT41 DNA collected 43

9. Southern blot of lambda clone (λNT41) containing the *qa* gene cluster of *N. tericola* probed with the *qa-1S* gene of *N. crassa* 45

10. Southern blot of lambda clone (λNT41) containing the *qa* gene cluster of of *N. tericola* probed with the *qa-y* gene of *N. crassa* 47

11. Southern blot of lambda clone (λNA43) containing the *qa* gene cluster of *N. africana* probed with the *qa-2* gene of *N. crassa* 49

12. Test gel showing successful recovery of the 2400bp fragment of λNA41 52

13. Test gel showing that the 650bp and the 1900bp fragments of λNT41 were successfully recovered 54

14. Test gel showing that pBluescript KS^+ was successfully isolated 56

15. Construction of the pNAL plasmid 58

16. Construction of the pNTL plasmid 60

17. Construction of the pNTU plasmid 62

18. Test gel of the constructed clones 64

19. Sample of typical sequencing data 67

20. Results of NCBI BLAST search 69

21. Sequence alignment of the coding regions of the *qa-2* genes of *N. africana* and of *N. crassa*. 71

22. Comparison of the *qa-2* protein of *N. africana* and the *qa-2* protein of *N. crassa* 74

23. Comparison of the *qa-2* protein of *N. africana*, the *qa-2* protein of *N. crassa*, and the *qutE* protein of *Aspergillus nidulans* 76

24. Blast results of interest when using a high homology segment of the *qa-2* protein of *N. africana* 79

25. Entire NAL clone aligned to *N. crassa* 81

26. Map of the constructed clone in relation to the *qa* cluster 87

Introduction

I. Kingdom Fungi

Divided into three phyla, Zygomycota, Ascomycota, and Basidomycota, fungi are a highly abundant and adaptive group of eukaryotic organisms. Ascomycota are particularly noted by their filament shaped thalli. The hyphal growth consists of apical elongation; they form ascospores contained within an ascus, which is similar to budding in yeasts (Taylor 1993).

Filling a niche as heterotrophic decomposers, fungi are an important part of the environment. During their metabolic activities they release carbon dioxide and nitrogenous materials into the environment. Fungi may live on living and non-living organic matter. In either case, they secrete special proteins known as enzymes to break down their carbon food source. They then adsorb the broken down food source for energy.

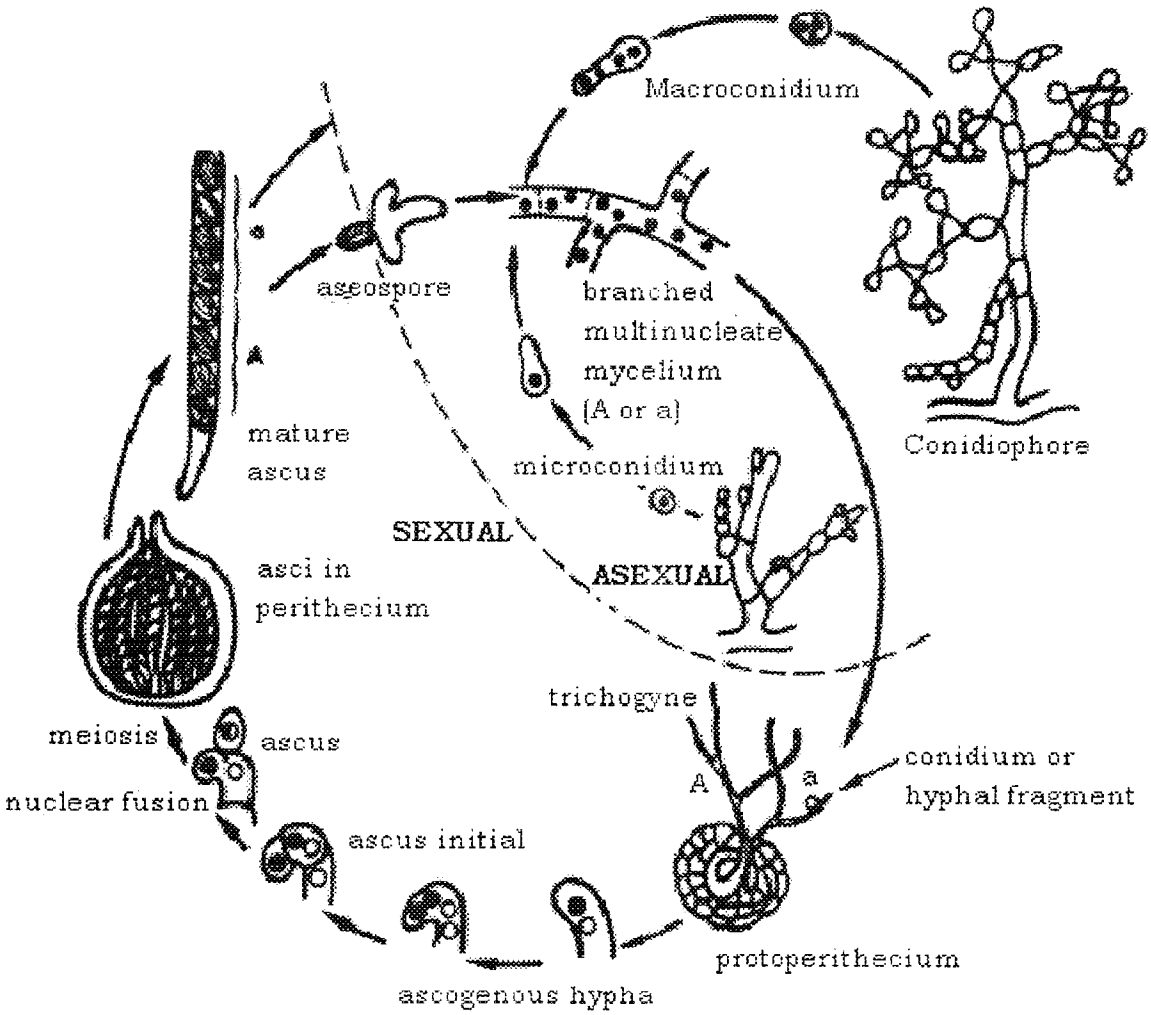
The *Neurospora* life cycle begins one of three ways. An asexual cycle originates either from a microconidia or a macroconidia. The environment appears to influence the type of reproduction that *Neurospora* uses. Macroconidia are multinucleated and form as hyphae mature. The circadian rhythm plays an intricate role in the timing involved during macroconidiation (Borkovich, Alex et al. 2004). *Neurospora* are found in vegetative cultures and as a male parent in fungal crossings. The microconidia are usually uninucleated and may also form male gametes. The conidia may then grow to form a branched network known as a conidiophore or may enter the sexual cycle.

Triggered by a limitation in nitrogen, the sexual life cycle causes a localized accumulation of hyphae and the formation of multicellular female gametes and organs (Borkovich, Alex et al. 2004). The sexual life cycle begins with the fusion of two gametes. The male gamete comes from the macro or microconidia while the female gamete is found in the protoperithecium. Upon fusion, a branched arrangement forms into the hypha. As the hypha matures an ascus may form at the tip. The ascus results from a meiotic division and contains ascospores. The ascospore may create another vegetative culture that will then repeat the cycle (Davis 1970). Figure 1 shows the life cycle of *Neurospora*.

II. History of *Neurospora*

Throughout scientific history, the name *Neurospora crassa* continually appears. It is the veritable white rat of geneticists. This eukaryotic fungus has a relatively small haploid genome composed of seven chromosomes containing all the features of normal eukaryotic DNA such as centromeres and nucleosomes. *Neurospora* transcripts are even subjected to intron removal (Mishra 1991). There are about 38 million base pairs within this genome (Borkovich, Alex et al. 2004); moreover, a shotgun sequencing method has made the genome sequence of *Neurospora crassa* available (Galagan 2003). In addition to the small genome, *Neurospora* can grow on many different carbon sources, which makes it quite inexpensive to maintain. For these reasons, molecular biologists frequently use *N. crassa* as a model system. Spending a large portion of its life in the haploid state makes Neurospora an excellent candidate for mutational studies, because

Figure 1: Life cycle of *Neurospora crassa* (Seale 1973).


Macroconidium
ascospore
branched
multinucleate
mycelium
(A or a)
Conidiophore
mature
ascus
microconidium
SEXUAL
ASEXUAL
asci in
perithecium
trichogyne
meiosis
ascus
conidium or
hyphal fragment
A
a
nuclear fusion
ascus initial
protoperithecium
ascogenous hypha

mutating one allele may result in an altered phenotype. Beadle and Tatum used *Neurospora* while performing their one gene - one protein studies (Perkins 1992). Because of these features, *Neurospora* is a good candidate for studying gene regulation in eukaryotic organisms.

III. Gene regulation: Prokaryote versus Eukaryote

Gene regulation is well understood in prokaryotic organisms. Prokaryotic organisms often cluster and regulate related genes as an operon. The purpose of an operon is to allow a coordinated transcription of related genes that are under a common promoter. This also ensures that there are equal amounts of gene products produced. Prokaryotes do not subject their transcripts to any postranscriptional modification. Eukaryotes however, often have related genes spread throughout the genome. In addition to control at the level of transcription, eukaryotes have many other mechanisms to control gene expression. Although there are more gene clusters in eukaryotes than originally thought, eukaryotic genes do not have a high frequency of gene clustering or operon arrangement (Keller and Hohn 1997). *N. crassa* does, however, contain a gene cluster coding for enzymes involved in the catabolism of quinic acid. The quinic acid gene cluster allows for a close inspection of gene regulation within a simple yet multicellular eukaryote.

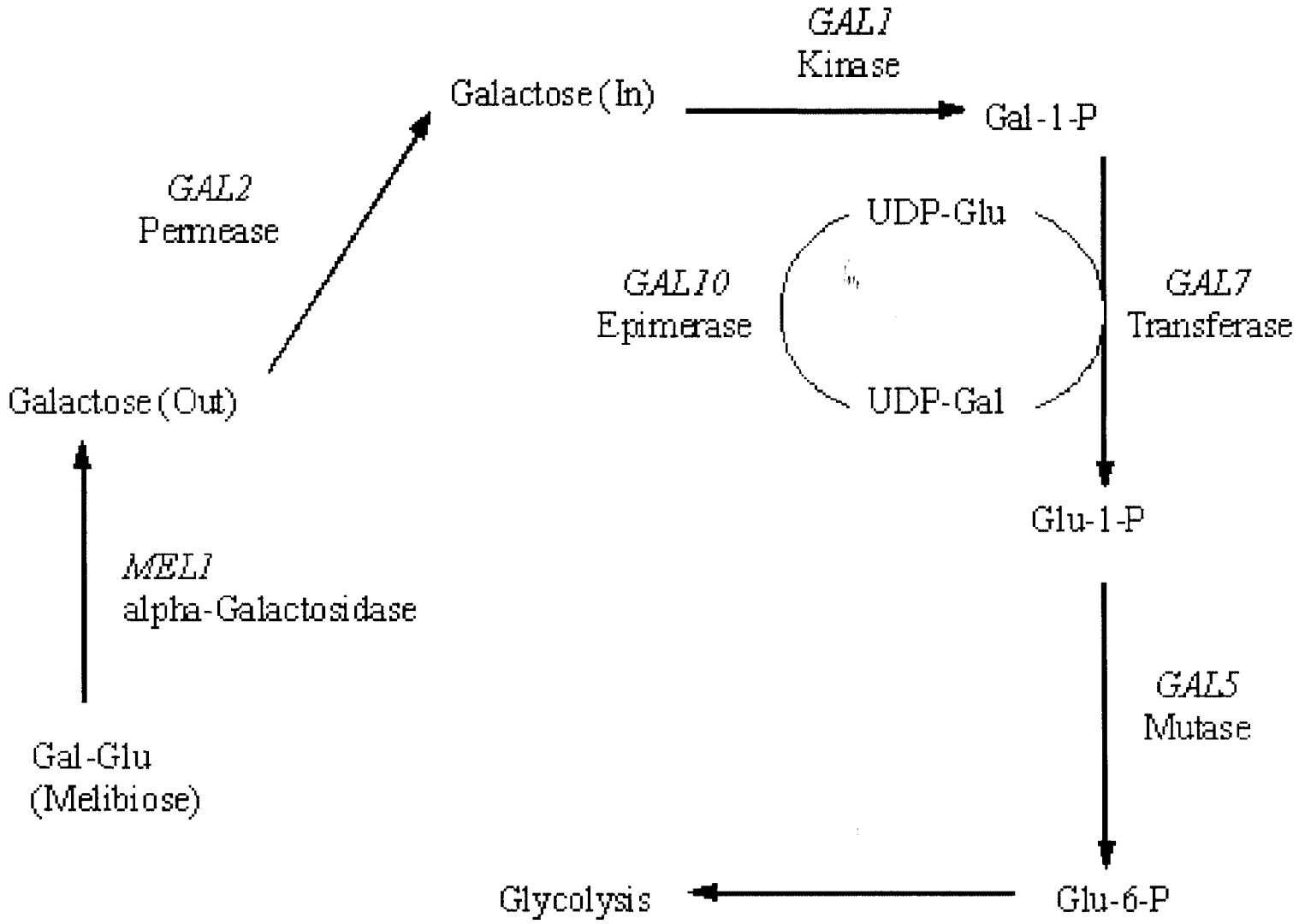
IV. Quinic Acid

Found in various plant tissues, quinic acid is soluble carboxylic acid. The quinic acid gene cluster imparts *Neurospora* with the ability to catabolize quinic acid as a sole carbon source. Since quinic acid is not a common or preferred carbon source, the gene products of the *qa* gene cluster are not always needed and, therefore, the genes encoding them are not always expressed. The *qa* gene cluster is only expressed as an alternate method of metabolism when under certain environmental pressures (Keller and Hohn 1997). This requires the same tight gene regulation one may find with the lac operon in prokaryotes or the galactose system of *Saccharomyces cerevisiae*.

V. The Galactose (*GAL*) pathway

Insights to gene regulation in *Neurospora* often come by comparing it to other organisms. The galactose (*GAL*) genes of *Saccharomyces cerevisiae* are under tight regulation. Composed of six structural genes [Figure 2 (Leloir 1951)], *Gal1*, *Gal7*, *Gal10*, *Gal5*, *Gal2* and *Mel1*, the galactose system is repressed by glucose and induced by galactose. The expression of these genes is tightly regulated with exception of *Gal5*, as it is constitutively expressed (Bevan and Douglas 1969). *Mel1* encodes an alpha-Galactosidase that will split melibiose into galactose and glucose. *Gal2* encodes a permease that allows for the uptake of galactose. *Gal1* encodes a kinase that is responsible for converting galactose into galactose-1-phosphate (Gal-1-P). *Gal7* is next

Figure 2: Representation of the galactose pathway.
MEL1 encodes an alpha-galactosidase, *GAL2* encodes a permease, GAL1 encodes a galactokinase, GAL7 encodes a galactose-1-phosphate uridylytransferase, GAL10 encodes a uridine diphosphoglucose-4-epimerase, and GAL5 encodes a phosphoglucomutase. Together they convert Galactose into Glucose-6-phosphate in preparation for glycolysis (Leloir 1951).


GAL1
Kinase
Galactose (In)
Gal-1-P
GAL2
Permease
UDP-Glu
GAL10
Epimerase
GAL7
Transferase
UDP-Gal
Galactose (Out)
Glu-1-P
MEL1
alpha-Galactosidase
GAL5
Mutase
Gal-Glu
(Melibiose)
Glycolysis
Glu-6-P

in line and encodes a transferase that converts the Gal-1-P product into galactose uridylyphosphate (UDP-Gal). *Gal10* encodes an epimerase that converts the UDP-Gal to glucose uridylyphosphate (UDP-Glu). The combination of Gal-1-P and UDP-Glu allows for the conversion to Glucose-1-Phosphate (Glu-1-P). Gal5 is then responsible for encoding a mutase that converts Glu-1-P into Glucose-6-Phosphate, which is glycolysis ready(Douglas and Hawthorne 1964; Douglas and Hawthorne 1972).

VI. *GAL* Pathway Regulation

The regulation of this pathway relies on three genes. *GAL4* and *GAL80* regulate the transcription of the other *GAL* genes. *Gal4* is produced in the cytoplasm and transported into the nucleus where it acts as an activator that binds to a site upstream of the *GAL* genes (Lue, Chasman et al. 1987). Through a portion of the GAL4 protein distinct from the DNA binding region, transcription is activated. *GAL80* inhibits transcription through an interaction with *GAL4*. *GAL3* is the key to activating this whole system. *GAL3* responds to galactose and causes some type of signal. *GAL3* does not directly encode the inducer, but does encode an enzyme that catalyses the synthesis of the inducer (Johnston 1987). The current theory is that galactose somehow interacts with GAL3 which in turn interacts with the GAL4-GAL80 complex (Blank, Woods et al. 1997). The interaction activates the GAL4 transcriptional domain and increases the level of transcription.

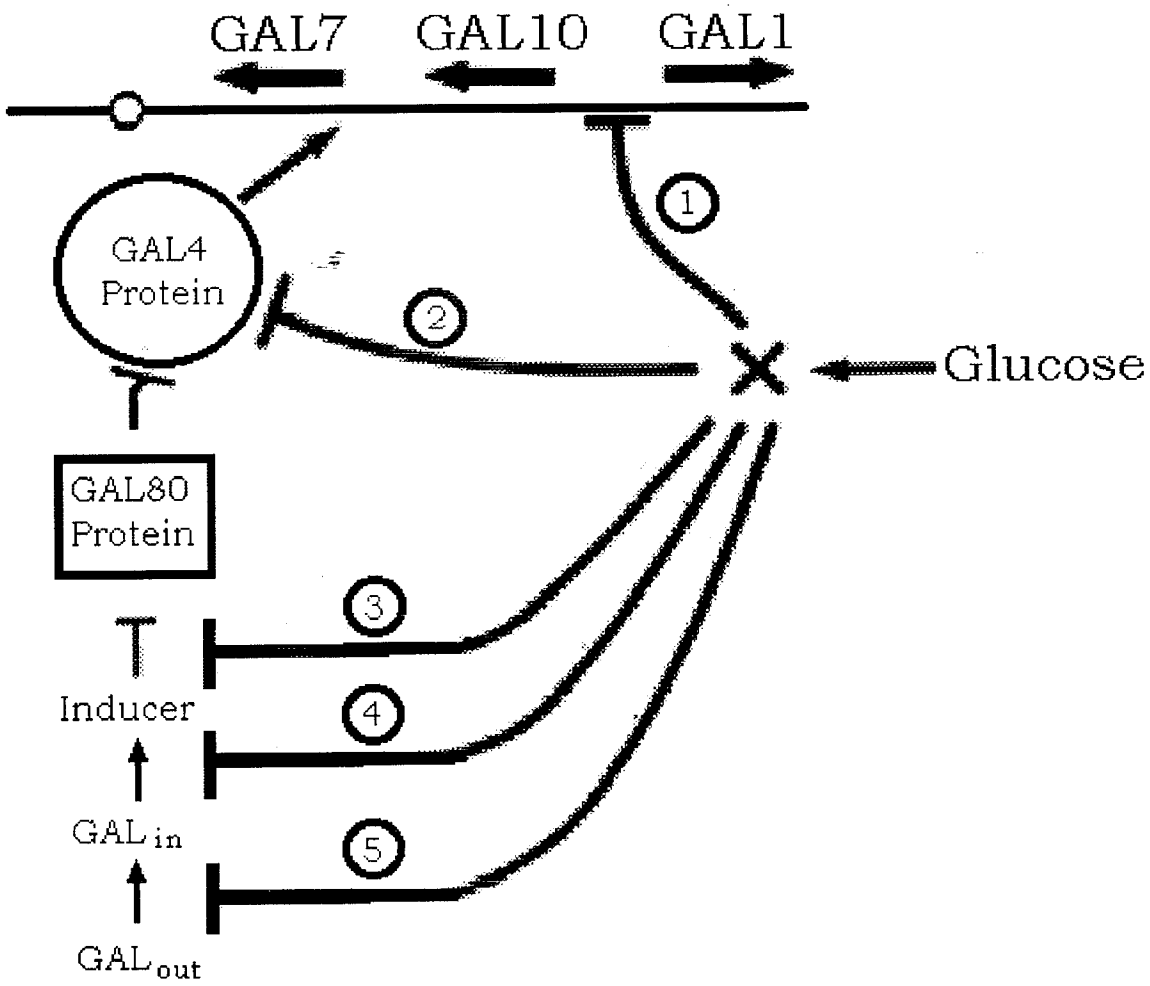
GAL4 has many functions that are dependant on the different regions of the protein. DNA-binding, transcriptional activation and transportation into the nucleus are

some of the domain dependant functions of GAL4. Being produced in the cytoplasm but functioning in the nucleus presents a problem for GAL4. It must be transported across the nuclear envelope. A region on the N-Terminal of GAL4 is believed to be responsible for transporting this molecule into the nucleus (Silver, Keegan et al. 1984). This region also appears to contain a zinc finger similar to that found on other DNA binding proteins that allows for GAL4-DNA binding. Evidence comes from *gal4* mutants that alter this structure, variations in this area terminates all GAL4-DNA interaction (Johnston and Dover 1987). The central region of GAL4 appears to be similar to an on/off switch. It is within this region that repression and activation appear to be controlled (Kang, Martins et al. 1993). When glucose is present, the GAL pathway will be quiescent. Catabolite repression may result from a few mechanisms that function at the level of transcription. The first mechanism occurs when GAL4 is inhibited from binding DNA when glucose is present; however, transcription is also repressed when GAL80 is present. The GAL80 does not interfere with GAL4-DNA binding, but interacts with GAL4 in such a way that transcription is blocked.

GAL80 has the properties of a repressor gene. Mutants of this gene have been shown to express *GAL* genes constitutively (Torchia, Hamilton et al. 1984). GAL80 works to repress *GAL* transcription by directly inhibiting GAL4 protein. This has been supported by various chromatography and gel mobility shift assays (Johnston 1987). Since GAL80 functions by directly interacting with GAL4, protein concentrations must be very important in controlling *GAL* transcription. One study shows this is the case as an increase in *GAL4* gene copy results in transcription of *GAL* genes independently of an inducer (Hashimoto, Kikuchi et al. 1983).

There are five proposed mechanisms of catoblite repression (Johnston 1987). The easiest way to regulate the pathway is at the level of transcription. Simply put, do not transcribe genes that are not going to be used. The first mechanism suggested is to repress the promoter. Evidence for this comes from deleting certain regulatory sequences that lie upstream of in a normal *GAL* promoter (Johnston 1987). The next opportunity for catabolite repression occurs with the sequence activator. Experiments where different promoter regions were used showed that there may be another protein present that prevents GAL4 from binding DNA (Johnston 1987). If GAL4 cannot bind, it cannot activate. The next opportunity seen is in the GAL80-inducer interaction. The presence of glucose may inhibit the inducer from inactivating GAL80; however, this cannot be the sole source of repression as GAL transcription is still inhibited in strains grown on glucose that contain no GAL80 (Torchia, Hamilton et al. 1984; Johnston 1987). Another way for repression would be to inhibit the inducer. Glucose may inhibit the transcription of *GAL3*, which indirectly inhibits synthesis of the inducer (Johnston 1987). The last proposed mechanism involves the permease. Not only is *GAL2* expression repressed, but glucose appears to inactivate preexisting permease proteins. Figure 3 illustrates the possible actions of catabolite repression (Johnston 1987).

Figure 3: Illustration of possible mechanisms of catabolite repression. Label 1: promoter repression; Label 2: GAL4 activator repression; Label 3: Inhibition of GAL80 inactivation; Label 4: Inhibition of *GAL3* transcription; Label 5: Galactose permease inhibition (Johnston 1987).


GAL7
GAL10
GAL1
GAL4
Protein
1
2
Glucose
GAL80
Protein
3
Inducer
4
GAL in
5
GAL out

VII. *Qa* gene cluster organization and function

The catoblite repression model demonstrated by the *GAL* pathway may also serve as a model for *Neurospora*. Like the workers along an assembly line, each gene within the *qa* gene cluster codes for a product that performs a specific function in quinic acid metabolism. If one worker is lost, a stop in the assembly line will result. Gene cloning technology has allowed a look at the molecular level of regulation and organization of these genes (Huiet and Case 1985). Through mutational studies and various other DNA and RNA analysis, there is an abundance of information characterizing the *qa* cluster. Composed of 5 structural and 2 regulatory genes (Geever 1989; Battogtokh, Asch et al. 2002) the *qa* gene cluster occupies 17.3kb on *Neurospora* linkage group VII. The organization of the quinic acid gene cluster is pictured in Figure 4 (Geever, Baum et al. 1987).

Recombinant technology has greatly facilitated the analysis of the *qa* gene cluster. The *qa-2* gene was the first gene of the cluster that was cloned. Clones containing the *qa-2* gene could be selected for through complementation in *Escherichia coli aroD* mutants (Vapnek, Hautala et al. 1977) These clones were then used to clone the entire *qa* cluster through a *Neurospora* transformation protocol (Huiet and Case 1985). Earlier studies, in which the *qa-2* gene was mutated, have shown that the *qa-2* gene encodes a catabolic dehydroquinase, which aids in the breakdown of 5-dehydroquniate to 5-dehydroshikimate (Huiet and Case 1985). If compared to the galactose pathway, *GAL1* would share a similar niche except *GAL1* phosphorylates the galactose while *qa-2* removes a hydroxyl group from the dehydroquinate. In both cases

the substrates are advanced in their metabolic pathways, allowing them to be further processed. The next steps rely on other proteins produced in the pathways to obtain the end products of Glucose-6-Phosphate and protocatechuic acid respectively. These products are ready for entry into glycolysis.

The *qa-3* gene encodes an enzyme that has two functions. It is responsible for the dehydrogenation of both quinate and shikimate acids (Giles, Case et al. 1985). This bifunctional enzyme contains one binding site for both substrates. This is similar to *GAL7* transferase which also has two functions.

The *qa-4* gene encodes an enzyme, 3-dehydroshikimate dehydratase, that is responsible for catalyzing the conversion of dehydroshikimate to protocatechuic acid (Stroman, Reinert et al. 1978). Like the *GAL5* mutase, *qa-4* is the last step in preparation of entering another pathway.

Transcriptional studies revealed the presence of two additional genes, *qa-x* and *qa-y* (Patel, Schweizer et al. 1981). The *qa-y* gene encodes a quinic acid permease that allows quinic acid uptake. Evidence suggests that *qa-y* is similar to other quinate permeases as well as other permeases found in other organisms. One study comparing *Neurospora crassa* and *Aspergillus nidulans*, suggests that the *qutD* gene of *A. nidulans* is responsible for quinic acid uptake and is homologous to the *qa-y* gene of *Neurospora crassa* (Whittington, Grant et al. 1987; Geever 1989). The function of the *qa-y* gene is comparable to that of the *GAL2* gene in *S. cereviseae*. Mutational deletions of the *qa-y* gene resulted in low levels of transcription for the entire cluster. This suggests that the *qa* inducer, quinic acid, was unable to enter the cell (Case, Geever et al. 1992).

Figure 4: Quinic Acid gene organization in *Neurospora crassa* (Geever, Baum et al. 1987).

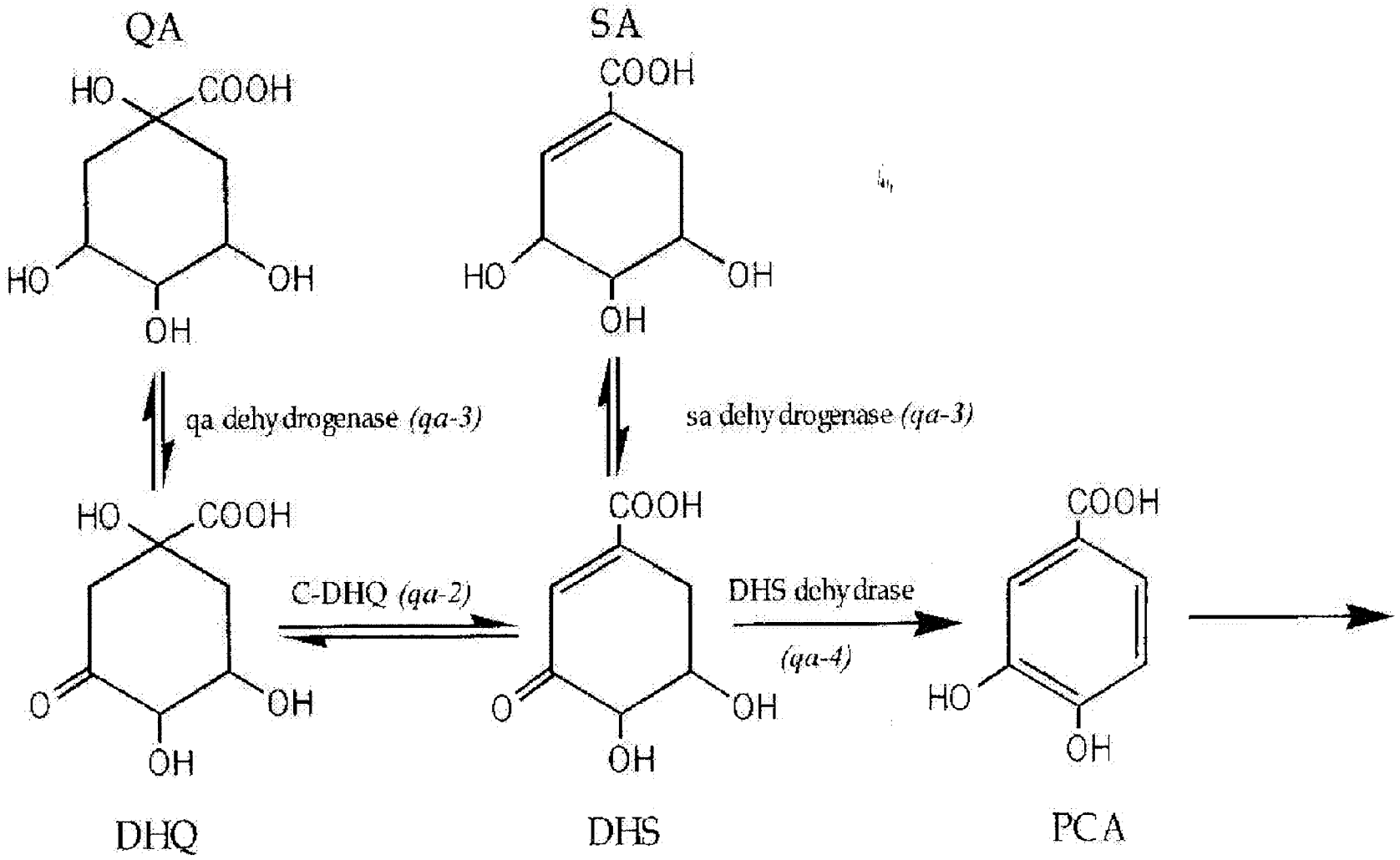
qa-1F
qa-1S
qa-3
qa-y
qa-4
qa-2
qa-x

The function of the *qa-x* gene product is currently unknown. A deletion study attempted to illuminate the role of *qa-x*. When *qa-x* was deleted, *N. crassa* continued to grow on quinic acid as a lone carbon source; however a brown pigment formed (Case, Geever et al. 1992). A transcription study shows that *qa-x* appears to be controlled by both catabolite repression and the activity of *qa-1S* (Huiet and Case 1985). One theory suggests that *qa-x* may be re responsible for the conversion of chlorogenic acid into quinic acid (Giles, Case et al. 1985). However, the current theory is that *qa-x* plays a role in catabolite repression. There is a noted increase in *qa-x* expression when growing on a limited carbon source; however, its exact functions remains yet to be elucidated (Giles, Case et al. 1985). Figure 5 illustrates the role of each gene product in the quinate and shikimate pathways (Arnett MS Thesis 2000).

The remaining two genes of the *qa* gene cluster are regulatory genes. Mutations in these genes resulted in two different phenotypes. The first mutants were uninducable; there were no *qa* gene products in the presence of quinic acid. The second mutant was a constitutive mutation. Constitutive mutants always produced *qa* gene products even in the absence of quinic acid. The entire cluster appears to be regulated at the transcriptional level, as no mutational studies have implicated a translational or post-translational level of control (Geever, Baum et al. 1987).

The *qa-1F* gene is a regulatory gene that encodes an activator for the entire *qa* gene cluster and increases transcription when quinic acid is present. This particular gene product appears to encode a protein that binds to the inducer quinic acid (Huiet and Case 1985). *Qa-1F* mutants are pleiotropic negative which means that none of the *qa* enzymes are active (Geever, Case et al. 1983). This activator protein is a DNA-binding protein. It

Figure 5: Metabolic pathway of quinic acid and shikimic acid via the *qa* gene cluster (Arnett MS Thesis 2000).


QA
SA
HO
COOH
OH
qa dehydrogenase (qa-3)
sa dehydrogenase (qa-3)
C-DHQ (qa-2)
DHS dehydrase
(qa-4)
O
DHQ
DHS
PCA

binds to consensus sequences before all of the *qa* genes. Figure 6 shows the consensus sequences to which *qa-1F* binds (Baum, Geever et al. 1987). Figure 7 shows where the activator protein binds in relation to the *qa* genes (Baum, Geever et al. 1987; Arnett MS Thesis 2000). The presence of an activator binding site in front of *qa-1F* supports that *qa-1F* is subject to autoregulation (Patel and Giles 1985; Baum, Geever et al. 1987). This activator protein appears to be similar to the N-terminal of GAL4 (Baum and Giles 1985). When the DNA-binding domains of *qa-1F* and *GAL4* are compared, there is similarity in both the localization and amino acid content (Baum, Geever et al. 1987).

The *qa-1S* gene encodes a repressor protein. The repressor target is not yet determined, but the favorite theory involves an interaction with the activator protein (Giles, Case et al. 1985). Quinic acid inactivates the *qa-1S* gene product (Baum and Giles 1985). Induction of the *qa-1S* requires both the *qa-1S* and *qa-1F* products to be active (Huiet 1984). *Qa-1S* mutants are recessive constitutive or semi-dominant non-inducible (Baum and Giles 1985). The entire *qa* gene cluster in *Neurospora crassa* has been mapped and sequenced (Geever 1989).

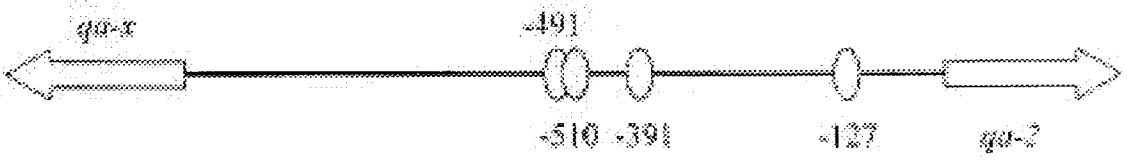
Figure 6: Sequences of the activator binding sites (Baum, Geever et al. 1987).

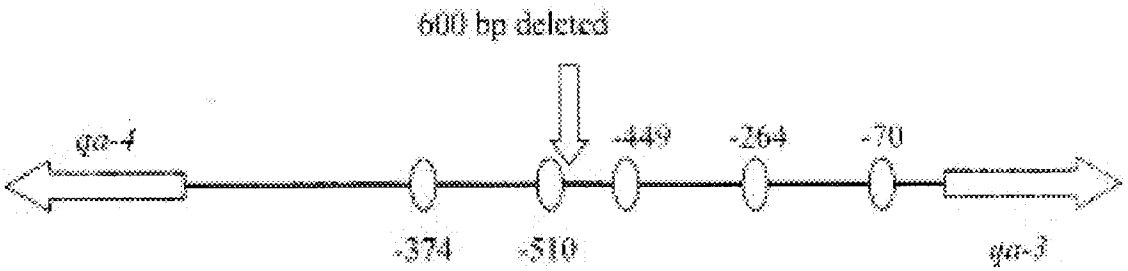
Gene	Position	Conserved Sequence		
qa-2	-510	GGATGA	GTTT	TTAACC
qa-2	-391	GGATAA	ATCC	TAACCC
qa-2	-127	GGGTAA	TCGC	TTATCC
qa-4	-510	GGCTAA	GTGT	TTAACA
qa-4	-347	CGTTAA	TGCT	TTATTC
qa-3	-499	GGGTAA	GTAT	TTAAGC
qa-3	-264	GGCAAA	ACGA	TCATCC
qa-3	-70	GGGGAA	TAAC	TTATAG
qa-y	-681	GGGTAA	TGGC	TTTTCC
qa-y	-512	GGTTAT	ACAT	TCATCC
qa-y	-412	GGCTCA	ACAC	TCATCA
qa-1S	+236	GGATGA	GTGA	TTCTCC
qa-1F	-144	GGATAA	ACAA	TTATCC
qa-x	-487	GGCGAA	CGTT	TTACCC
Consensus sequence		GGRTAA	RYRY	TTATCC

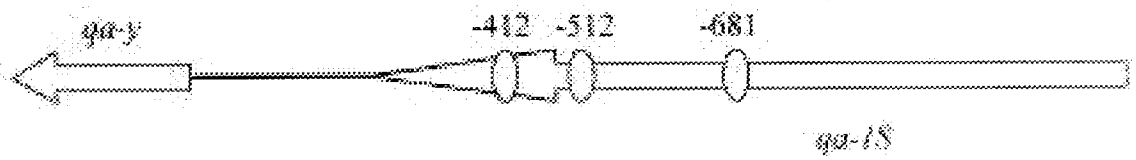
Consensus Sequence: $G^{93}G^{100}R^{57}T^{76}A^{76}A^{93}$ $R^{64}Y^{71}R^{71}Y^{86}$ $T^{100}T^{71}A^{86}T^{64}C^{76}C^{76}$

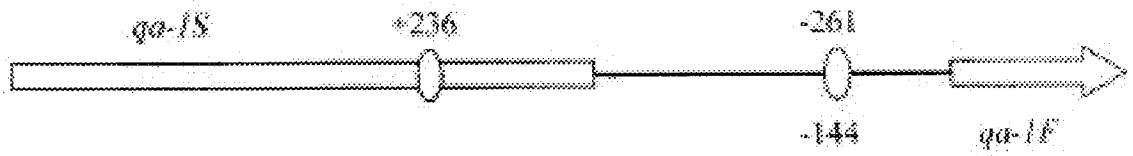
Figure 7: Locations of the activator binding sites (Baum, Geever et al. 1987; Arnett MS Thesis 2000).


-1
-500
-1


qa-x
-491
-510
-391
-127
qa-2


600 bp deleted
qa-4
-449
-264
-70
-374
-510
qa-3


qa-y
-412
-512
-681
qa-1S


qa-1S
+236
-261
-144
qa-1F

VIII. Comparison between the Quinic Acid (*qa*) Gene Cluster of *Neurospora crassa* and the Quinic Acid Utilization (*qut*) Gene Cluster of *Aspergillus nidulans*

Although it is from a different family and DNA sequence comparisons of the genomes have little in common, the *qut* gene cluster of *A. nidulans* has a high functional similarity to that of the *qa* gene cluster of *N. crassa*. The *qa* cluster has two regulatory proteins, as does the *qut* cluster. *QutR* is similar to *qa-1S* as they both encode repressor proteins, while *qutA* encodes an activator like the *qa-1F* gene. *Aspergillus nidulans* also contains a permease similar to that of *N. crassa*. The *qutD* gene encodes this permease. The *qutG* gene shares the same transcriptional characteristics and same mystery function as the *qa-x* gene. Even though similar functions are observed, amino acid similarity varies except at the functional domains. At a slightly higher level, there are two other features that are similar. First the genes are divergently transcribed and the structural genes are separated from the regulatory genes (Grant, Roberts et al. 1988; Hawkins, Lamb et al. 1988).

IX. Homothallic versus Heterothallic

There are several homothallic species of *Neurospora*. Homothallic species are capable of self-fertilization. All the homothallic *Neurospora* species tested are capable of using quinic acid as a sole carbon source. A previous study compared homothallic *N. africana qa* genes to the heterothallic *N. crassa qa* genes (Asch, Orejas et al. 1991). This study showed that there was a high amount of conservation in the heterothallic gene

cluster; moreover, it showed that the gene organization remained highly conserved. However there were different restriction fragment length polymorphisms within the two clusters.

This study is going to further this research by isolating the *qa-2* gene of *N. africana* and comparing its DNA to sequence to the DNA sequences of the *qa-2* gene of *N. crassa*. It is also going compare the amino acid sequences of *N. africana qa-2*, *N. crassa qa-2* and *A. nidulans qutE.*

Materials and Methods

I. Materials

Bacto-tryptone bacto agar was purchased from Difco Laboratories (Detroit, MI). Restriction endonucleases (*EcoRI, BamHI* and *HindIII*), RnaseA, CIP, 10XCIP buffer, T4 DNA ligase, were purchased from Boehringer Mannheim (Indianapolis, IN). Ethanol was purchased from Aaper Alcohol and Chemical Company (Shelbyville, KY). Agarose was purchased from EM Science (Cherry Hill, NJ). Ethidium Bromide, 85% phosphoric acid, sodium citrate and acetic acid were purchased from Fisher Scientific (Fairlawn, NJ). Isopropanol was purchased from Pharmco Products (Brookfield, CT). Concentrated HCl was purchased from VWR Scientific (San Francisco, CA). QIAGEN columns were purchased from QIAGEN, Inc. (Chatsworth, CA). pBluescript was purchased from Stratagene (LaJolla, CA). Polaroid film, yeast extract, sodium chloride, ampicillin, isopropyl-β-D-thiogalactoside (ITPG), 5'-bromo-4-chloro-3-indoyl- β -D-galactopyronoside (X-gal), dextrose, Trizma base, ethylenediaminetetraacetic acid – disodium salt (EDTA), sodium hydroxide, sodium lauryl sulfate (SLS), potassium acetate, phenol, cresol, chloroform, sodium acetate, magnesium chloride, cesium chloride, glycerol, polyethylene glycol (PEG) and adenosine triphosphate (ATP) were purchased from Sigma Chemical Company (St. Louis, MO). dNTPs were purchased from PE Biosystems (Norwalk, CT). Sequencing supplies (contains primer pUC: 5'-CGC CAG GGT TTT CCC AGT CAC GAC-3') were purchased from Beckman Coulter, Inc., (Fullerton, CA). Primers (M13R: 5'-GGA AAC AGC TAT GAC CAT G-3', NAL-F-1:

5'-GGT AGT AGC TGG CAT AGG TGG –3', NAL-F-2: 5'-GGT TGA TGA TGA TGG CCG ACA CTT-3', NAL-F-3: 5' - TGC TTG GTG GAT ACG GTC GAT GAT – 3', NAL-F-4: 5' – CCA TCA ATG AGA GAG CTT ATC CGC – 3', NAL-F-5: 5' – AGG CTT AGA GAG TGT GTG TGC TTG – 3', NAL-F6: 5' – CGC CAG TAA TAG GAG TCA TAT GCG – 3', NAL-R-1: 5'-TGT GAG TTG AGT ATC GTG CGG-3', NAL-R-2: 5'-ATG GCG TCC CGT CAT CAT ATT CTC-3', NAL-R4: 5' – AGA GAT GCG TAT AGT TGA CCG TGC – 3', NAL-R-4: 5' – ATG TCT AGT GTA GCT TGG CCT TCC – 3', NAL-F-x-1: 5'-CAG AGC AGG TCA CAA TGG AGA TGT-3', NAL-R-x-1: 5'-CCG TTT GGG TCA CAA TGT CAA CAG-3') were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). NZ amine and Casamino Acids were purchased from Difco (Sparks, MD).

II. Strains

The *E. coli* used in these experiments was obtained from the *E. coli* Genetic Stock Center (CGSC) at Yale University. Strain JM101 was selected to be the recipient of all constructed plasmid vectors. Cells were cultured in Luria Broth [LB] (1% bacto-tryptone, 0.5% yeast extract, 1% NaCl). Luria Agar [LA_{100}] (1% bacto-tryptone, 0.5% yeast extract, 1% NaCl, 100μg/ml ampicillin) plates containing ampicillin were used to select cells that received the plasmid DNA. Transformants that received target DNA were select using 100μL of 200mM IPTG and 50μL of 2% X-Gal in dimethyl formamide (DMF).

III. Bluescript

Bluescript phagmid pBluescript II KS (+) contains an ampicillin resistance site for selection. It also contains a portion of the *lacZ* gene which allows for blue/white color selection. Cloning into this phagmid into the multiple cloning region disrupts the *lacZ* gene causing the formation of white colonies in the presence of ITPG and X gal.

IV. Phage Titration

A three hour Luria Broth culture of JM101 was prepared by inoculating 50mls L-broth with 500μL from an overnight culture. The three hour culture was centrifuged for 10 minutes at 10,000g at 4°C (SA-600). The supernatant was discarded and the pellet was resuspended in about 25mls of 0.01M $MgSO_4$. A serial dilution of phage was prepared. Five hundred μl of cells were added to each dilution. A control tube was prepared that contained no phage. These tubes were incubated at 37°C for 15 minutes. Three mls of melted NCZYM soft agar (1% NZamine, 0.5% NaCl, 0.1% Casamino Acids, 0.5% Yeast Extract, 0.2% $MgSO_4 \cdot 7\ H_2O$, 0.7% Bacto-Agar) was added to each tube. The contents of the tube was poured onto an NCZYM hard agar plate (1% NZamine, 0.5% NaCl, 0.1% Casamino Acids, 0.5% Yeast Extract, 0.2% $MgSO_4 \cdot 7\ H_2O$, 1.5% Bacto Agar). This was allowed to harden for 10 minutes before incubating overnight at 37°C. Plaques were counted on the following day.

V. Phage Recovery

Based on the plaque count, enough phage was present to continue. More plates were diluted to the concentration that optimum yield. Four plates were plated and incubated overnight at 37°C. A sterile metal spatula was used to scrap the soft agar from the plate. This was placed into a centrifuge tube. About 4mls of SM buffer (0.58% NaCl, 0.2% $MgSO_4 \cdot 7\ H_2O$, 5.0% 1M Tris•Cl, 0.1% gelatin) total were used to rinse the plates. The plate washings were added to the centrifuge tube and centrifuged for 10 minutes at 10,000g at 4°C (SA-600). The supernatant was recovered and stored at 4°C.

VI. Lambda DNA Preparation

Ten mls of an overnight culture of Q358 cells were centrifuged down for 10 minutes at 10,000g at 4°C (SA-600). The supernatant was discarded and 10mls of SM buffer was used to resuspend the cells. Three mls of cells was placed in a new tube. $1.14X10^5$ phages were added to the cells. This was incubated at 37°C for 20 minutes. This culture was then used to inoculated 250mls of LB and incubated in a shaking incubator at 37°C until lysis. Once lysis occured 5mls of chloroform was added and incubated for 15 minutes at 37°C. This was poured into a two centrifuge bottles and centrifuged for 15 minutes at 9,600g (SLA-1500). Two hundred of μL of solution L1 (0.02g RnaseA, 100μL 0.06g/mL DnaseI, 20 μL 10mg/mL BSA, 20μL 0.5M EDTA, 100μL 1M Tris, 60μL 5M NaCl, 700μL Water) was added to each aliquot. This was incubated for 30 minutes at 37°C. The samples were removed from the incubator and

25mls of chilled L2 solution [30% polyethylene glycol (PEG 6000), 3M NaCl] was added to each sample. This was incubated for 1 hour on ice. The samples were again centrifuged at 9,600g for 10 minutes (SLA-1500). The supernatant was discarded and the samples were allowed to air dry while inverted. Nine mls of solution L3 (100mM NaCl, 100mM Tris•Cl pH 7.5, 25mM EDTA) was added to one aliquot and the sides of the tube were washed. This was then transferred to the other aliquot and the sides of the tube were again washed. This was then transferred to a 50ml centrifuge bottle and 9mls of solution L4 [4% sodium lauryl sulfate (SLS)] was added. This was incubated at 70°C for 10 minutes. After placing on ice to cool the bottle down, 9mls of solution L5 (3M potassium acetate pH 5.5) was added. This was mixed by inversion. This was then centrifuged at 23,000g for 30 minutes (SA-600). The supernatant was saved and the pellet was discarded. A QIAGEN column was prepared by pouring 10mls of QBT buffer (750mM NaCl, 50mM MOPS pH 7.0, 15% Isopropanol, 0.15% Triton® X-100) through it. The supernatant was then allowed to flow through the column. The column was then washed with 30mls of QC buffer (1.0M NaCl, 50mM MOPS pH 7.0, 15% isopropanol). The DNA was eluted into a clean tube with QF buffer (1.25M NaCl, 50mM Tris•Cl pH 8.5 15% isopropanol). The DNA was then washed with 70% ETOH. The pellet was allowed to air dry and resuspended in 2mls of 1XTE [Dilute from 100X solution (12.11% Tris Base, 3.72% EDTA, 50mL concHCl)].

VII. Gel Electrophoresis

DNA was analyzed using 1.0% agarose gels. The gels were made and run using 1x Tris Phosphate EDTA (TPE) buffer [0.08 M Tris, 0.005 M EDTA, 85% H_3PO_4 (1.679mg/ml)]. Ethidium Bromide (50µl of a 50mg/ml stock) was used to stain the gel. A transilluminator emitting UV light was used to view the gels. A two second exposure time was used in all photographs.

VIII. Southern Blot Digests

About 7 µg of Lambda DNA was digested with *EcoRI*, *HindIII*, *BamHI*, *EcoRI/BamHI*, *HindIII/BamHI*, and *EcoRI/HindIII* in separate eppendorf tubes. Four µL of 10xbuffer, 28 µL of De-ionized water, and 1 µL of enzyme was added to each digest and allowed to incubate for three hours at 37°C.

IX. Making Probe

Fifty ng of *N. crassa* DNA containing the *qa-2* gene was placed in about 5µL 1XTE and 6µL of water. This was denatured at 90°C for 5 minutes. 2.5µL of 10X buffer (6µL dNTPs, 2µL 1/100 dATP, 2.5µL [^{32}P] dATP, 1µL Klenow) was added and incubated for 2 hours. A spin column containing saline treated glass wool at the bottom and sephadex G-50 beads on top was prepared. The label mix was loaded into spin column and 150µL of 1XTE was added on top. The column was centrifuged for 5

minutes at about 7,000rpms (microfuge). Harvest probe and check for percent incorporation.

X. Southern Blot

Digested Lambda DNA was electrophoresed to separate fragments. After staining with Ethiduim Bromide, holes were punched out of the gel where the Lambda marker bands were located. The gel was then placed face down in a blotting tray. The tray was filled with 200mls of solution BI [10ml concHCl, 390mls water] to remove purine bases. This was placed on a rocker for 15 minutes. The BI was removed and 200mls more of BI was added. After 15 minutes, BI was removed and 200mls of solution BII [10M NaOH, 5M NaCl] was added to make the DNA single stranded. After 15 minutes, BII was removed and 200mls of solution BII was again added. After again rocking for 15 minutes, BII was removed and 400mls of solution BIII [1M Tris, NaCl] was added to rinse the gel and stabilized the DNA. While rocking for 30 minutes, three pieces of blotting paper, a piece of membrane and an abundance of paper towels are cut to a size equal to that of the gel. Once cut, wet the membrane and one sheet of blotting paper with some BIII. The wet sheet of blotting paper was placed on a tray the edges were covered and with parafilm. The gel was placed face down on the blotting paper followed by the membrane. Wrinkles were smoothed out and the other two sheets of blotting paper were placed on top of the membrane. The paper towels were stacked on top of the blotting paper and a weight was added. Blotting was allowed to proceed overnight. On the

following day, the location of the lanes and lambda marker was marked. UV was used to cross link the blot and the blot was stored wrapped in syran wrap in a freezer.

XI. Prehybridization and Hybridization

The membrane was placed in a hybridization tube and washed with about 30mls of 6XSCP [0.6M Na_2HPO_4, 2.0M NaCl] for 30 minutes at 60°C while rolling. Twelve mls of 6XSCP containing 300μL of single stranded salmon sperm DNA (5mg/ml) was brought to a boil. The probe was added and boiled for 5 minutes. The prehybridization washing was discarded and the probe mix was added. This was hybridized overnight at 60°C while rotating.

The probe was poured off and the membrane was washed with 2XSCP with 1% SLS for 30 minutes while rotating at 60°C. The wash was discarded and washed again with 0.2XSCP with 1%SDS for 15 minutes while rotating at 60°C. The blot was exposed to X-ray film (Kodak Biomax MR). The exposure time depended on the blot activity.

XII. Large Scale Isolation of Plasmid DNA

An overnight culture of *E. coli* containing the plasmid was used to begin a large (0.5mL in 250mL) culture in L-broth containing 100μg/ml of Ampicillin. The cells were harvested by centrifuging at 9,600g for 10 minutes (SLA-1500). The bottles were drained well and inverted to dry. The pellet was resuspended in 30mls of G-Buffer [500mls: 25mls 1M Dextrose, 12.5mls of 1M Tris, 25mls of 0.2M EDTA; pH ~8]. The

bottles were placed on ice and add 60mls of fresh denaturing solution [10ml 10N NaOH, 25ml 20% SLS, bring volume to 500]. The mixture was allowed to sit for 10 minutes on ice. Forty-five mLs of neutralizing solution [73g KOAc, 28.5mls conHOAc, bring volume to 500mls with water] was added and gently mix. The mixture was placed on ice for 45 minutes, then centrifuge at 9,600g for 15 minutes (SLA-1500). The supernatant was filtered through cheese cloth into a clean centrifuge bottle and 85mls of Isopropanol was added. The samples were centrifuged at 9,600g for 30 minutes (4°C) (SLA-1500). The tubes were drained well and the pellet was lyophilized to dryness. The DNA was resuspended in 6mls of 1XTE.

The DNA solution was used to make a CsCl gradient (.75g/ml CsCl). The total volume of the DNA solution was brought to 15mls with 1XTE. The sample was placed in a Sorval centrifuge bottle with 50 μL of 50mg/ml ethidium bromide. The sample was centrifuged at 300,000g (23°C) overnight (SORVALL® Stepsaver 65 V13). After a breathing hole was created in the top of the tube, the plasmid band was removed using a 5ml syringe with a 10 gauge needle. The ethidium bromide was removed from the sample with 1:5 water to isopropanol extraction. The ethidium bromide was on the top layer. This was repeated until the ethidium bromide was gone. The bottom layer was transferred to a dialysis bag. The DNA was dialyzed overnight in 500mls of 1x TE. The TE was changed two times. A Phenol Chloroform extraction (described below) was done and run a test gel was run.

XIII. Isolation and Digestion of Vector

Two separate digests of Bluescript, one with *BamHI* and the other with *HindIII*, were done. The 50 μL of Bluescript vector (50 μL Bluescript, 8 μL buffer, 1 μL of restriction enzyme) was incubated at 37°C. A test gel was run to determine if digestion was complete.

XIV. Isolation and Digest of Fragment

A 200 μL digest mix [100 μL (~7μg) of Lambda clone DNA, 8 μL 10x Buffer, 2 μL of restriction enzyme, bring volume to 200 with DI water] was prepared. The digest was incubated at 37°C overnight. The DNA was run overnight with size standards (Lambda DNA cut with *HindIII*). After the gel was stained and visualized, a razor blade was used to remove the desired bands. The excised gel pieces were placed into separate dialysis bags. The bags were filled with 400 μL of 0.5x Tris-Acetate (TAE) [0.04 M Tris, pH 7.9; 0.2M NaOAC; 0.002M Na_2EDTA, pH 7.9]. Air bubbles were removed and the ends were sealed. The dialysis bags were placed perpendicular to the electrical field and ran for 45 minutes. The current was reversed for 30 seconds and the gel pieces were removed from the dialysis tubes. The contents of the dialysis bags were decanted to an eppendorf tube. A phenol and chloroform extraction (described below) was done. The DNA was stored in 95% ETOH in the freezer.

XV. Phenol Chloroform Extraction and DNA Precipitation

An equal volume of phenol was added to the sample and mix well. The sample was centrifuged for 10 minutes at 13,000rpms (microfuge). The top layer was carefully removed to a fresh tube so that the interface was not disturbed. This step was repeated until the interface decreased in size. An equal volume of chloroform was added and mixed well. The sample was centrifuged for 5 minutes at 13,000rpms (microfuge) and the top layer was decanted to a fresh tube. A 1/10 volume of NaOAc was added and the tube was filled with 95% ethanol. The sample was centrifuged at 13,000rpms (microfuge) for 30 minutes (4°C). After the sample was decanted and 800μL of 70% ETOH was added. The DNA was resuspended in 1XTE (usually about 50μL).

XVI. CIP and Ligation

To increase ligation efficiency, calf intestinal phosphatase (CIP) was used on the vector. Ten μL of vector (cut with appropriate enzyme), 3μL 10xCIP buffer, 1μL CIP, 36μL of water was mixed and incubated at 37°C for 30 minutes. One hundred fifty μLs of 1xTE was added. A phenol chloroform extraction and an ethanol precipitation (as described above) was done. The sample was resuspended in 10μL of 1XTE.

The Ligation mix contained 1μL of vector DNA (cut with appropriate enzyme), 10μL fragments, 2μL of 10x ligation buffer, 1μL Ligase and 6μL of water. The sample was incubated at 15°C overnight.

XVII. Transformation

One hundred μLs of competent cells were placed into eppendorf tubes and 5μL of ligation mix was added to all tubes except one (control). The tubes were placed on ice for 20 minutes. The tubes were heat-shocked for 5 minutes at 37°C. One ml of Lauria broth was added to each tube and incubated at 37°C for 45 minutes. Fifty μL of X-gal and 50μL of ITPG was spread on Lauria plates containing 100μg/ml Ampicillin. Two hundred μL of the transformed cells were spread on the plates and incubated at 37°C overnight. Colonies were screened and picked the following day.

XVIII. Alkaline Plasmid Screen

After the white colonies were picked, they were grown in 2mls of L-broth (100μg/ml ampicillin) overnight. The samples were placed in eppendorf tubes (1.5mls). The samples were centrifuged for 15 seconds at 13,000rpms (microfuge). The samples were decanted and the pellets were allowed to air-dry. Two hundred μLs of G-buffer was added to each tube and the samples were placed on ice. Four hundred μLs of fresh denaturing solution was added and the samples were placed on ice for 5 minutes. Four hundred μLs of ice cold neutralizing solution was added to the tubes and allowed it to sit on ice for 15 minutes. The samples were centrifuged for 5 minutes at 13,000rpms (microfuge) and the supernatant was transferred to a clean tube. Five hundred ninety μLs of 99% isopropanol was added to the supernatant and the samples were centrifuged for 5 minutes at 13,000rpms (4°C) (microfuge). The samples were drained well and washed

with 80% ETOH. The samples were centrifuged for another 5 minutes at 13,000rpms (4°C) and drained well (microfuge). The pellets were dried in spin vacuum. The pellets were resuspended in 50μL of 1x TE and 9μL of each sample was digested. A test gel was run. If test gel showed positive results, the cultures were maintained. If negative results were found, another culture was picked and the screening was repeated. When a positive colony was found, a large scale recovery was done as previously described except the DNA was resuspended in 50μL of DI water. The DNA was digested and a test gel was run containing both digested and a known volume of non-digested DNA.

XIX. Sequencing Reactions

The concentration of DNA to use was determined by comparing a standard lambda ladder of known concentration to the unknown concentration of plasmid DNA. This number was then plugged into the following equation to determine the μL of sample to use:

$$\text{volume in } \mu L = \frac{[(100fmol)(649)(\text{size of plasmid with insert})(10^{-6})]}{\text{concentration of plasmid ng}/\mu L}$$

Once the amount of sample was determined the reaction mixture was prepared (8μL of Master Mix, 2μL of desired primer, bring total volume to 20μL with water). This was placed in the thermocycler (96°C for 20 seconds, 50°C for 20 seconds, 60°C for 4 minutes for 45 cycles and held at 4°C when done). Four μL of stop solution (40% NaOAc, 8% 0.5M EDTA, 12% glycogen, 40% water) was added to clean eppendorf tubes. The reaction mixes were then placed into the stop solution. This was mixed and 60μL of 95% ETOH was added to each tube. After mixing again, the tubes were

centrifuged for 15 minutes at 13,000rpm at 4°C. The supernatant was discarded and the pellets were washed 2 times with 70% ETOH. The pellets were allowed to air dry for about 20 minutes and resuspended in 40μL of sample loading solution. The samples were then sequenced on a Beckman CEQ 2000 dye terminator sequencer with the desired method. The reactions were repeated using the different primers to fill in gaps.

XX. Comparative Analysis

Using a bioinformatics program known as BioEdit, the various sequencing products were aligned. This allowed for the discrimination of any gaps. Once the gaps were filled, the sequence was then ready for comparison. Using the short nearly exact matches nucleotide blast search that is available on the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/), the sequenced DNA was compared to all the DNA available on the gene bank. It showed where the sequencing data begins to line up with other subjects, the location of *qa-2*, the orientation of the inserted DNA and gives a score for how well the sequences compare.

Results

I. Identification of DNA fragments containing the various *qa* genes from the homothallic species of *Neurospora*

The lambda clone λNT41 which is known to containing some or all of the *qa* cluster of *Neurospora terricola* was grown and prepared using the QAIGEN method. A gel was run to check the quality and quantity of λNT41 DNA obtained (Figure 8).

λNT41 was digested with *EcoRI*, *HindIII*, *BamHI*, *BamHI/EcoRI*, *HindIII/BamHI*, and *EcoRI/HindIII*. DNA was electrophoresed and a southern blot of the λNT41 DNA was prepared and probed with *qa-2* from *N. crassa*. It was found that λNT41 did not contain the *qa-2* gene (data not shown). The blot was stripped and probed again with the *qa-1S* (Figure 9) and *qa-y* genes of *N. crassa* (Figure 10). A 650bp band and a 1900bp band were selected from Lane 4 of the λNT41 for subcloning as they were shown to contain other parts of the *qa* cluster.

A lambda clone λNA41 containing the *qa* gene cluster of *N. africana* was also grown and DNA was isolated. A λNA41 blot was prepared. *Qa-2* was found in some of the smaller λNA41 fragments. A blot of a similar phage λNA43 is shown in Figure 11. The smallest fragment (~2400bp) from Lane 4 of the λNA43 blot was selected for subcloning based on small fragment size and high band intensity.

Figure 8: Test gel giving insight to the quality and quantity of λNT41 DNA collected.
Lane 1: *HindIII* digested lambda marker, Lane 2: λNT41 phage DNA.

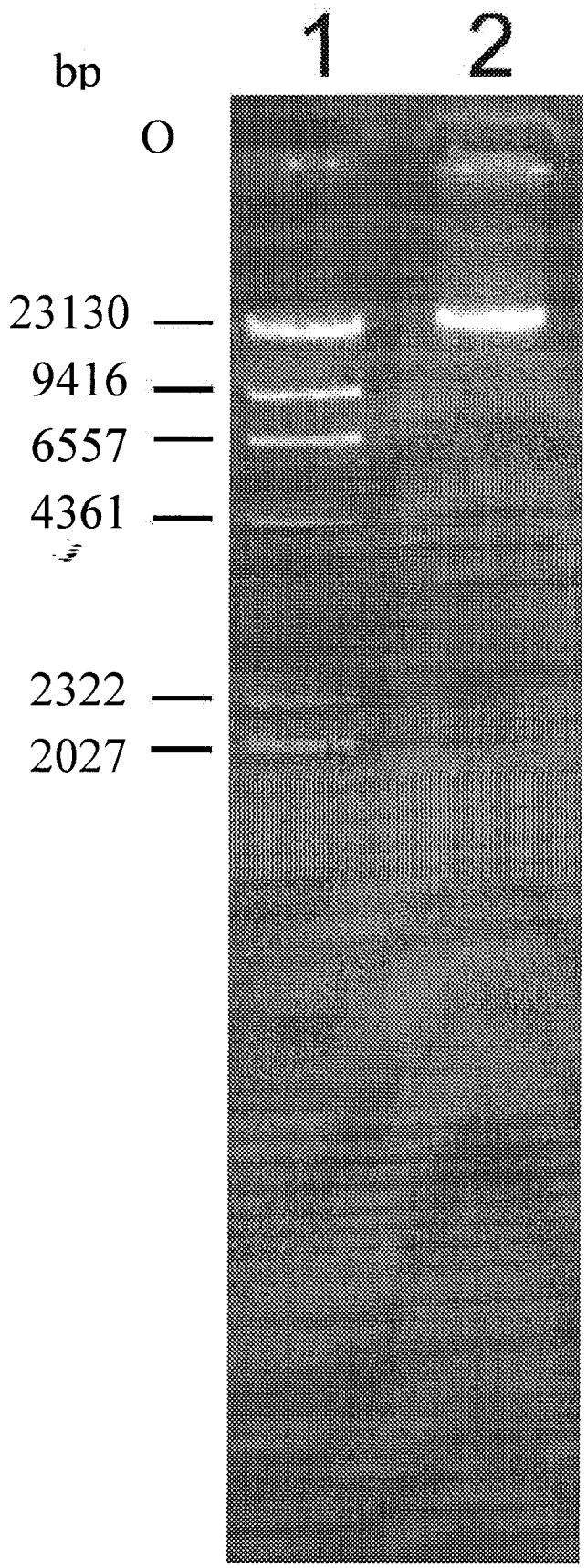

bp
1
2
O
23130
9416
6557
4361
2322
2027

Figure 9: Southern blot of lambda clone (λNT41) containing the *qa* gene cluster of *N. tericola* probed with the *qa-1S* gene of *N. crassa*.
Lane 1: New England Biolabs 1 kb ladder; Lane 2: λNT41/*EcoRI*; Lane 3: λNT41/*HindIII*; Lane 4: λNT41/*BamHI*; Lane 5: λNT41/*EcoRI* and *BamHI*; Lane 6: λNT41/*HindIII* and *BamHI*; Lane 7: λNT41/*EcoRI* and *HindIII*

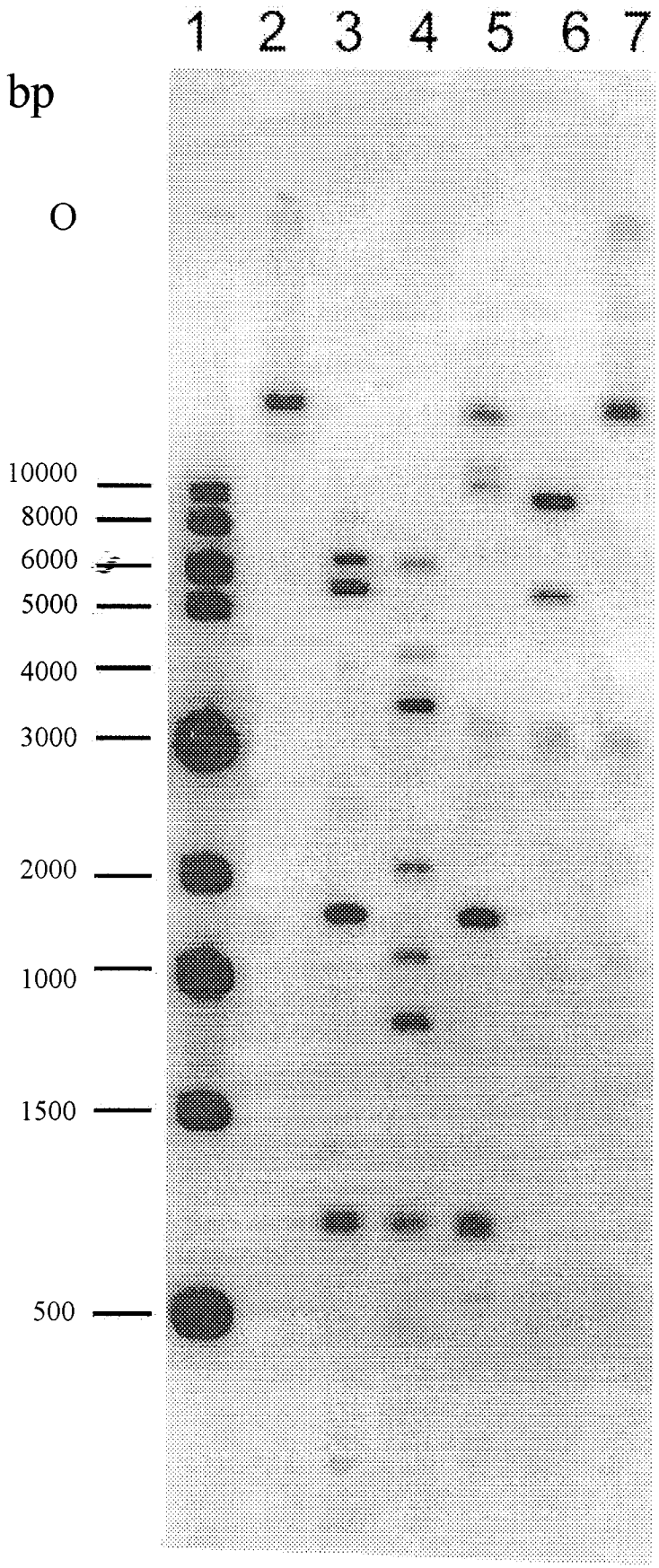

1 2 3 4 5 6 7
bp
O
10000
8000
6000
5000
4000
3000
2000
1000
1500
500

Figure 10: Southern blot of lambda clone (λNT41) containing the *qa* gene cluster of of *N. tericola* probed with the *qa-y* gene of *N. crassa*.
Lane 1: New England Biolabs 1 kb ladder; Lane 2: λ NT41/*EcoRI*; Lane 3: λNT41/*HindIII*; Lane 4: λNT41/*BamHI*; Lane 5: λNT41/*EcoRI* and *BamHI*; Lane 6: λNT41/*HindIII* and *BamHI*; Lane 7: λNT41/*EcoRI* and *HindIII*

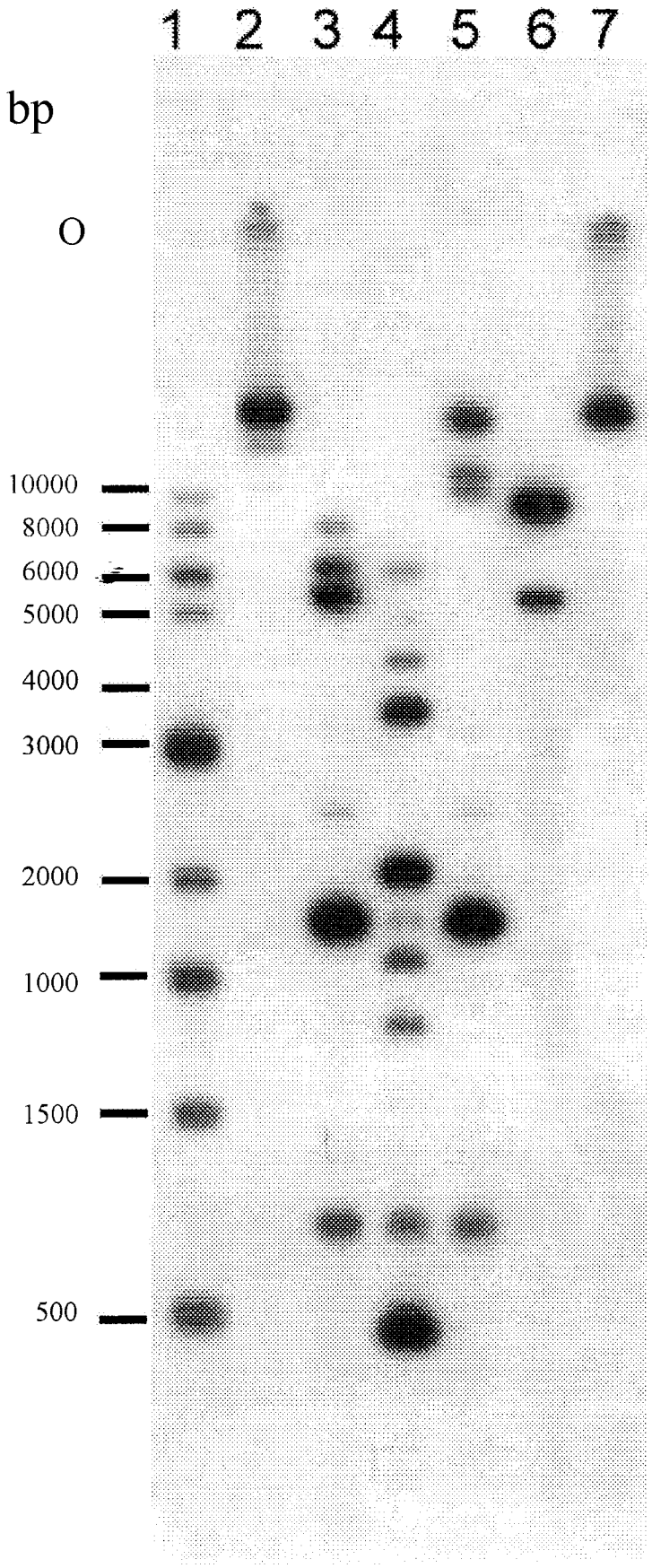
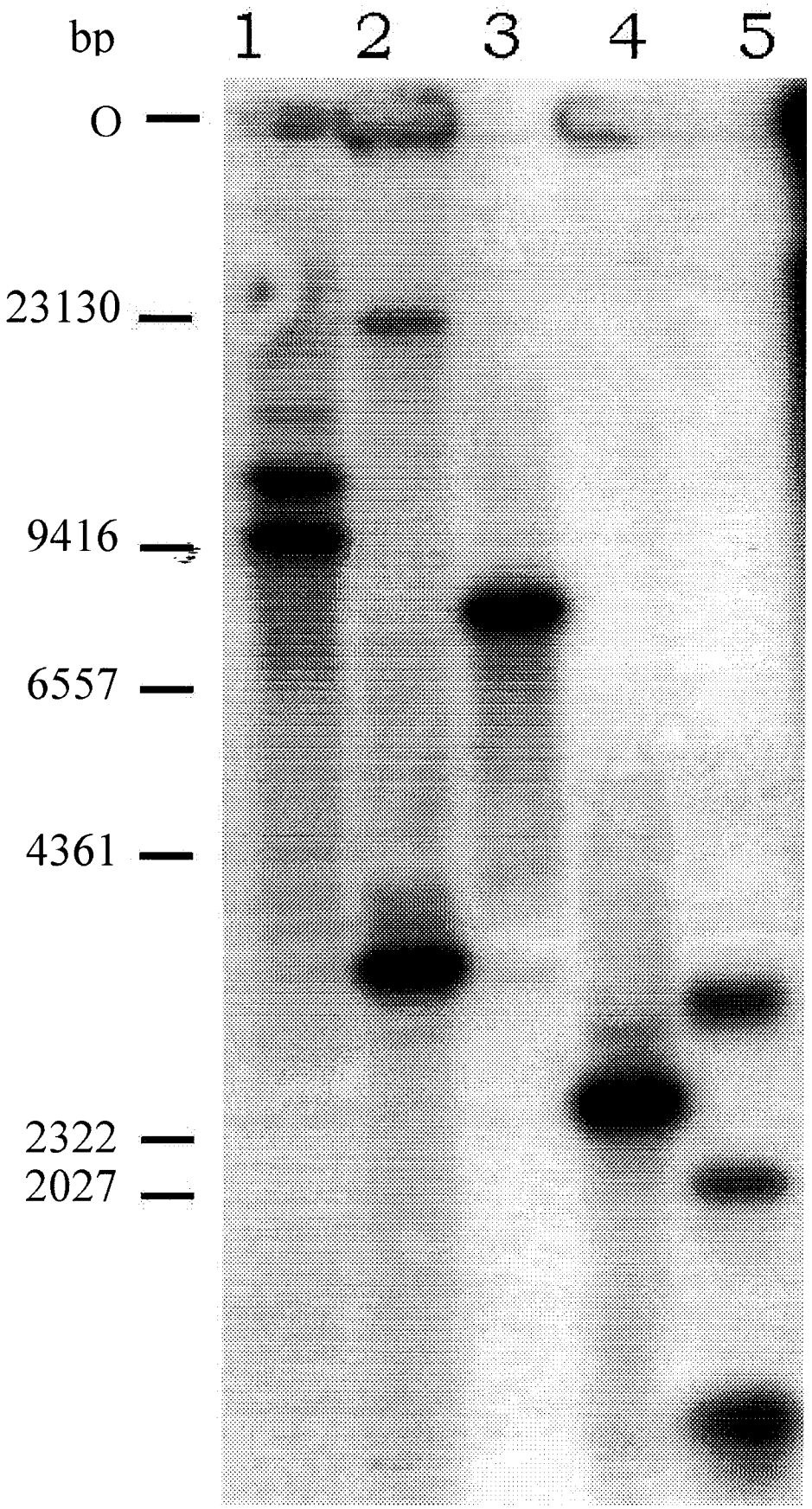

1 2 3 4 5 6 7
bp
O
10000
8000
6000
5000
4000
3000
2000
1000
1500
500

Figure 11: Southern blot of lambda clone (λNA43) containing the *qa* gene cluster of *N. africana* probed with the *qa-2* gene of *N. crassa*.
HindIII digested lambda marker is shown; Lane 1: λNA43 cut with *BamHI*; Lane 2: λNA43/*BglII*; Lane 3: λNA43/*EcoRI*; Lane 4: λNA43/*HindIII*; Lane 5: λNA43/*KpnI*.


bp
1
2
3
4
5
O
23130
9416
6557
4361
2322
2027

II. Subcloning of fragments containing *qa* genes from *N. africana* and *N. terricola*

DNA from phages λNT41 and λNA41 were digested at a larger amount in order to recover the desired DNA fragments. Test gels indicate that the desired fragments were successfully recovered (Figure 12 and 13).

The phagemid, pBluescript KS^+, was selected as the vector for subcloning. It contains both the *LacZ* operon and an ampicillin resistance gene. These two factors allow for selection of plasmid with insert and cells that have been transformed. Plasmid DNA was checked for quality and quantity (Figure 14). Vector digested with the appropriate restriction enzyme were ligated with the various fragments [(2400bp fragment of λNA41, the 500bp fragment of λNT41, and 1900bp fragment of λNT41) Figures 16, 17 and 18].

Ligation mixes were transformed into *E. coli* and white ampicillin resistant colonies were selected. Colonies were then quick screened. Promising clones were grown large scale and plasmid DNAs were isolated using a CsCl gradient.

The DNA was recovered and a confirmatory digests shows both the plasmid DNA and the inserted DNA traveling at the expected bp (Figure 19). The plasmids were now ready for sequencing.

Figure 12: Test gel showing successful recovery of the 2400bp fragment of λNA41. Lane 1: *HindIII* digested lambda marker; Lane 2: λNA41 fragment

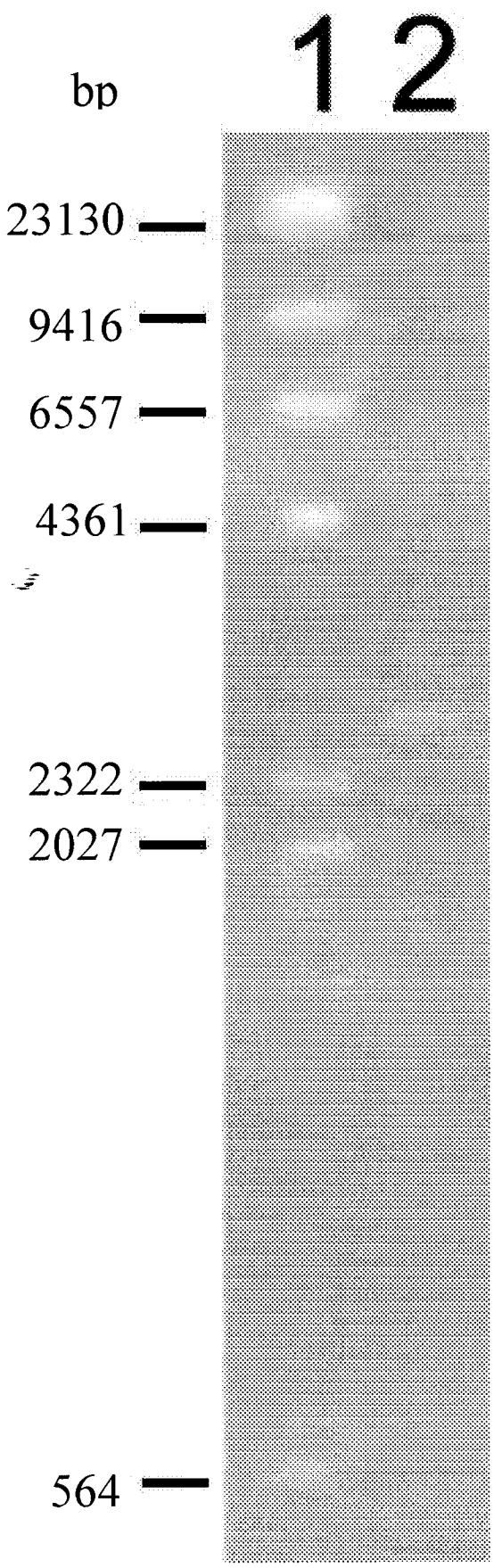

bp
1 2
23130
9416
6557
4361
2322
2027
564

Figure 13: Test gel showing that the 650bp and the 1900bp fragments of λNT41 were successfully recovered.
Lane 1: *HindIII* digested lambda marker; Lane 2: 650bp λNT41 fragment; Lane 3: 1900bp λNT41 fragment.

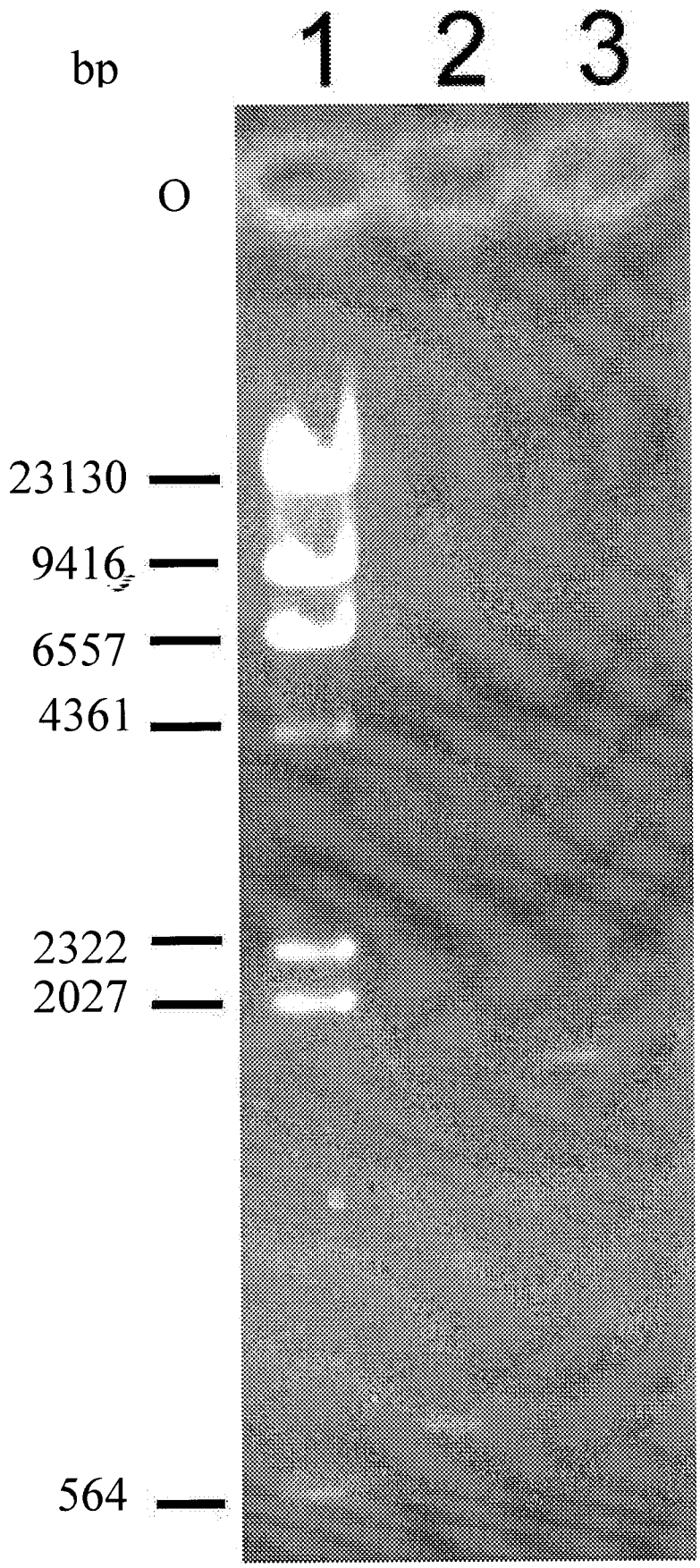

bp
1
2
3
O
23130
9416
6557
4361
2322
2027
564

Figure 14: Test gel showing that pBluescript KS^+ was successfully isolated. Lane 1: *HindIII* digested lambda marker; Lane 2: *HindIII* digested pBluescript KS^+

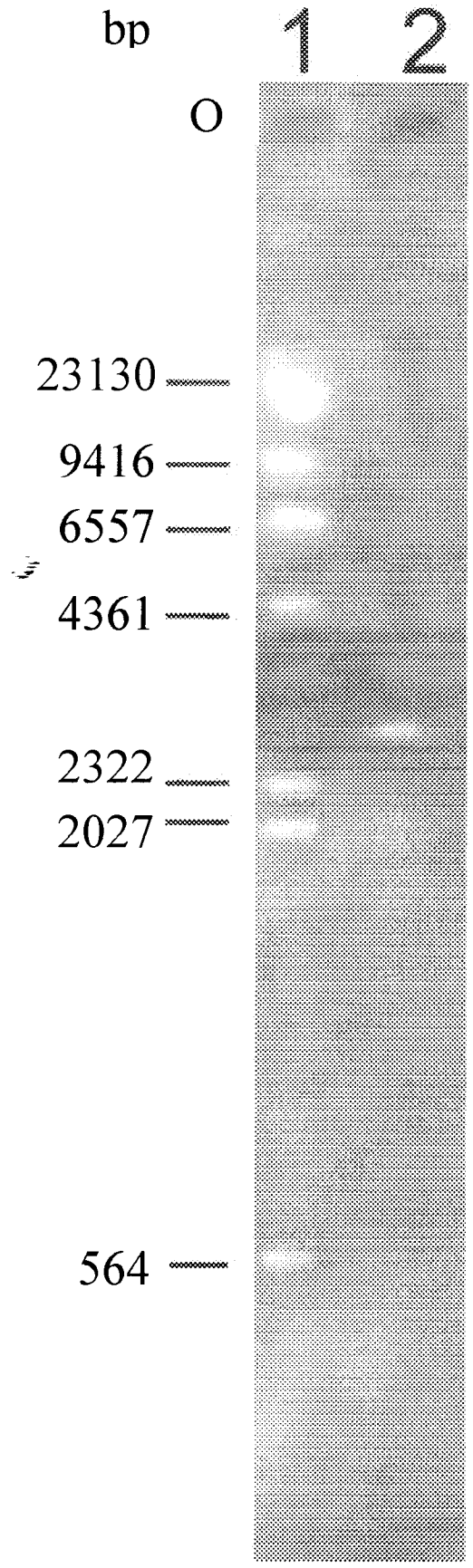

bp
1
2
O
23130
9416
6557
4361
2322
2027
564

Figure 15: Construction of the pNAL plasmid. Dark gray region indicates the *N. africana qa-2* coding sequence.

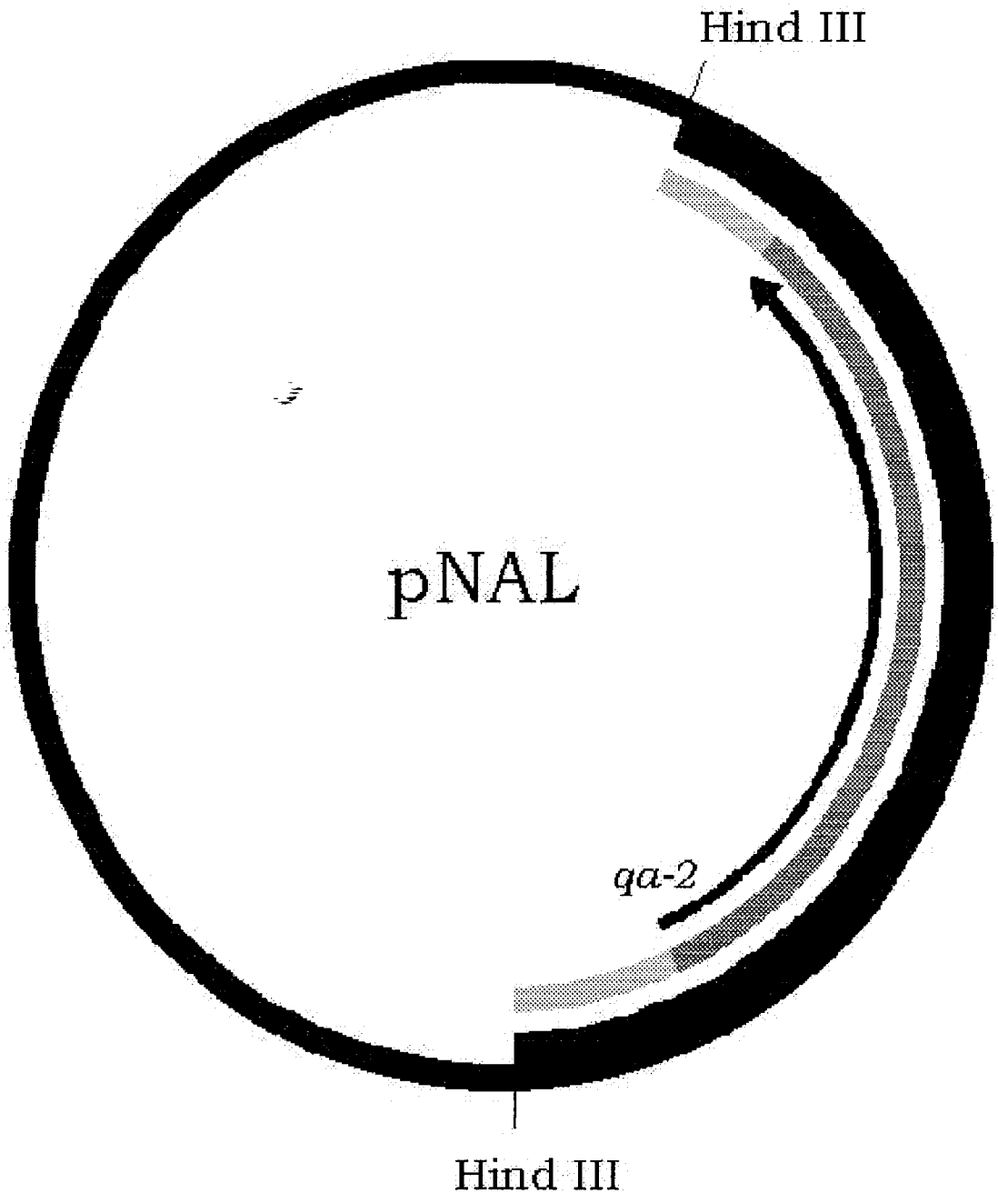

Hind III
pNAL
qa-2
Hind III

Figure 16: Construction of the pNTL plasmid. The dark gray area represents the inserted fragment of *N. terricola*

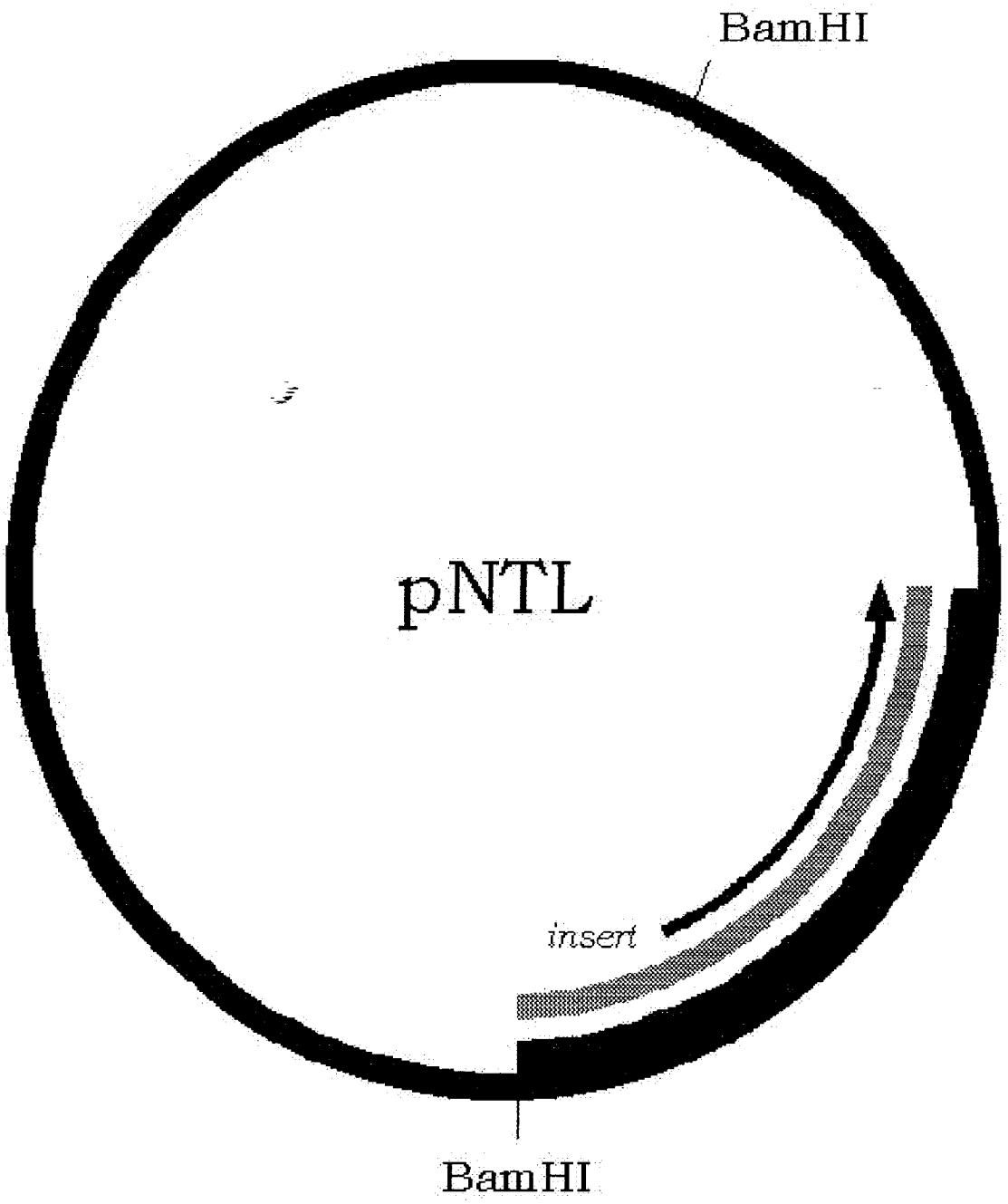

BamHI
pNTL
insert
BamHI

Figure 17: Construction of the pNTU plasmid. The dark gray area represents the inserted fragment of *N. terricola*.

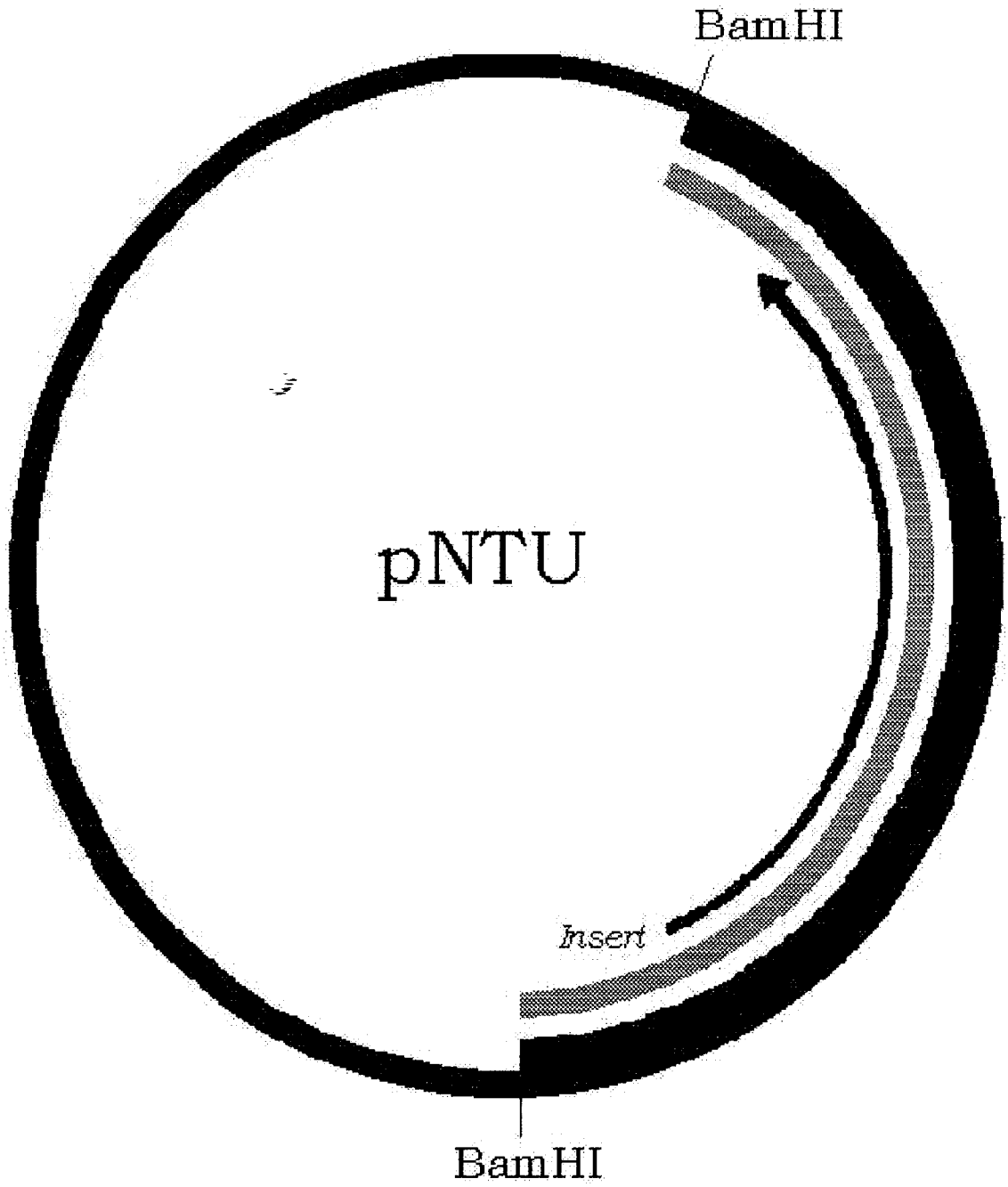
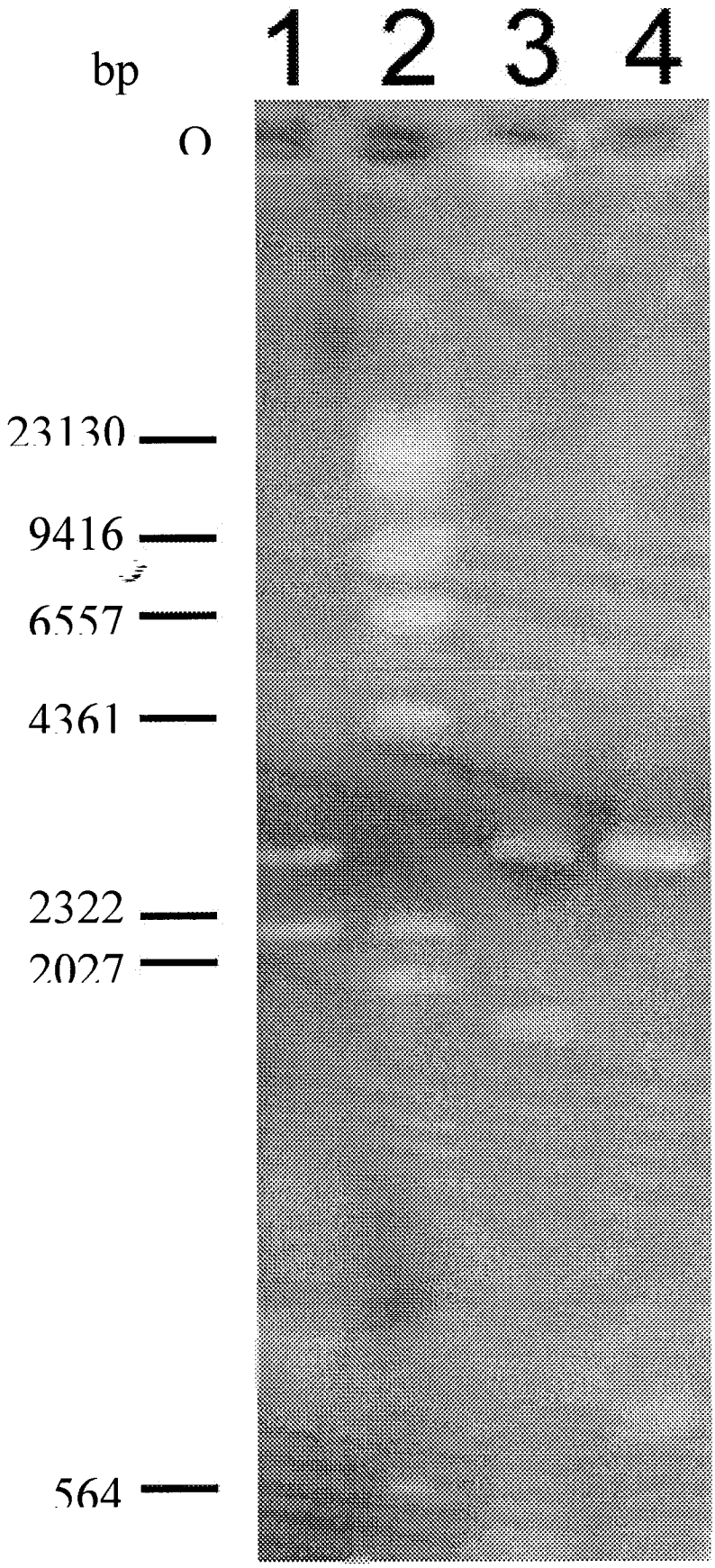

BamHI
pNTU
Insert
BamHI

Figure 18: Test gel of the constructed clones.
Lane 1: *HindIII*/NAL; Lane 2: *HindIII* digest of lambda marker; Lane 3: *BamHI*/NTU; Lane 4: *BamHI*/NTL.


bp
1 2 3 4
23130
9416
6557
4361
2322
2027
564

III. Sequencing of the plasmid subclones pNAL, pNTL and pNTU

Following the Beckman Coulter sequencing protocol, the inserts of all three plasmids were sequenced. Typical sequence data is shown in Figure 19. A nucleotide blast search on the National Center for Biotechnology Information was conducted in order to determine the sequence homologies, if any (Figure 20). The BioEdit program was then used to align and compare sequences. New primers were selected in the forward and reverse directions based on the sequence data. More sequencing reactions were done.

IV. Alignment of *Neurospora crassa qa-2* and *Neurospora africana qa-2*

Sequence data from pNAL lines up with the *N. crassa qa-2* gene in reverse orientation. There is a low e value meaning that there is high homology between the two sequences containing a 91% identity (Figure 21). Since homology between pNAL and part of the *qa-x* gene was seen with this run, primers for the reverse direction were based upon already published sequence data by Dr. David Asch (Asch, Orejas et al. 1991). There is high sequence homology between *N. crassa* and *N. africana*. Once aligned, they were found to be 88% identical and 91% similar.

Figure 19: Sample of typical sequencing data.

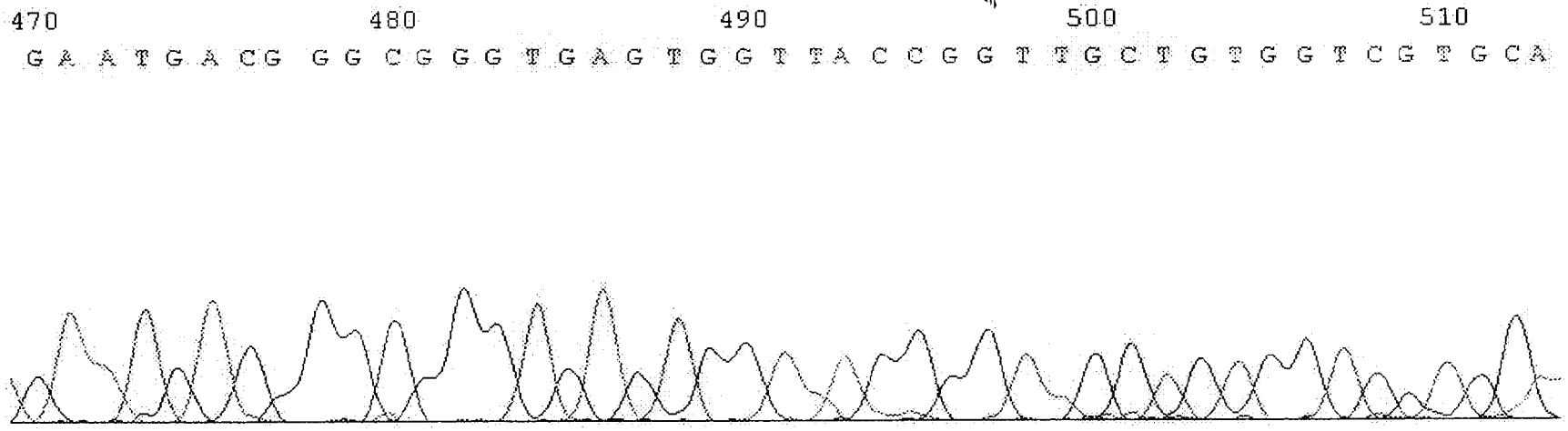
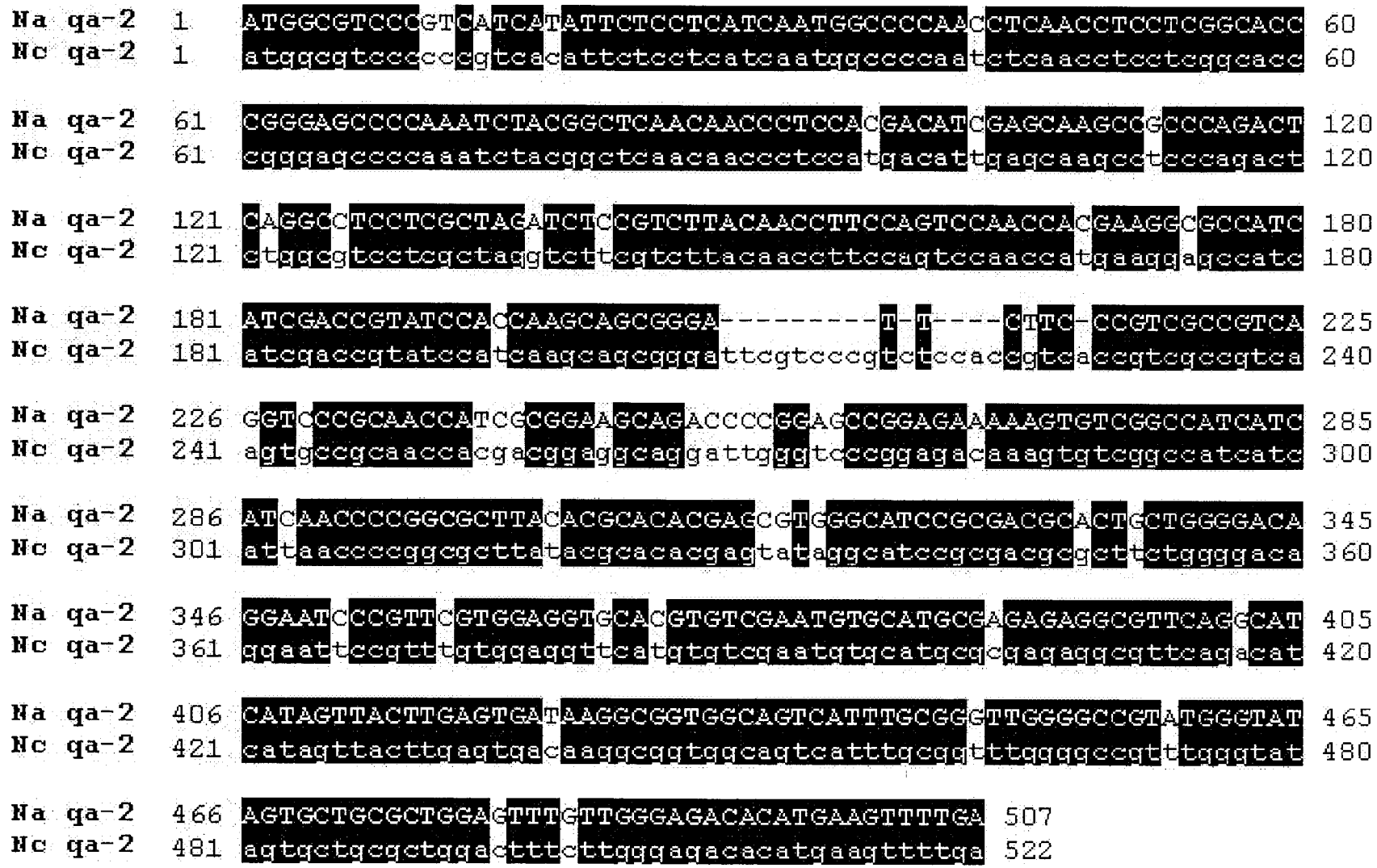

470
480
490
500
510
GAATGACG GGCGGTGAGTGGTTACCGGTTGCTGTGGTCGTGCA

Figure 20: Results of NCBI BLAST search. The sequence data (query) is in reverse orientation to the *qa-2* gene of *Neurospora crassa*.

```
gi|3060|emb|X14603.1|NCQA   Neurospora crassa qa gene cluster
          Length = 18120

 Score =  341 bits (172), Expect = 8e-91
 Identities = 229/248 (92%)
 Strand = Plus / Minus


Query: 1     tcaaaacttcatgtgtctccaaacaaactccagcgcagcactatacccatacgggcccaa 60
              ||||||||||||||||||||||| ||| ||||||||||||||||||| ||||||||||
Sbjct: 3384  tcaaaacttcatgtgtctccaagaaagtccagcgcagcactatacccaaacgggcccaa 3325


Query: 61    ccgcaaaatgactgccacgccttatcactcaagtaactatgatgccgaacgcctctct 120
               |||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Sbjct: 3324  acgcaaaatgactgccacgccttgtcactcaagtaactatgatgtctgaacgcctctcg 3265


Query: 121   cgcatgcacattcgacacgtgcacctccatgaacgggattcctgtccccagcagtgcgtc 180
              |||||||||| ||||||| || |||||||| ||||| |||||||||||||| || |||||
Sbjct: 3264  cgcatgcacattcgacacatgaacctccacaaacggaattcctgtccccagaagcgcgtc 3205


Query: 181   gcggatgcccacgctccgtgtgcgtgtaagcgccggggttgatgatgatggccgacacttt 240
              ||||||||| |  |||||||||| |||||||||||||| ||||||||||||||||||
Sbjct: 3204  gcggatgcctatacttcgtgtgcgtataagcgccggggttaatgatgatggccgacacttt 3145


Query: 241   ttctccgg 248
               |||||||
Sbjct: 3144  gtctccgg 3137



 Score =  305 bits (154), Expect = 4e-80
 Identities = 202/218 (92%)
 Strand = Plus / Minus


Query: 290   gacgggaagaatcccgctgcttggtggatacggtcgatgatggcgcccttcgtggttggac 349
              ||||||| |||||||||||||| ||||||||||||||||||| ||||| |||||||||
Sbjct: 3080  gacgggacgaatcccgctgcttgatggatacggtcgatgatggctcccttcatggttggac 3021


Query: 350   tggaaggttgtaagacggagatcttagcgaggaggcctgagtctgggcggcttgcttcgatg 409
              |||||||||||||||| ||| |||||||||| ||| ||||||||| |||||||| |||
Sbjct: 3020  tggaaggttgtaagacgaagacctagcgaggacgccgagtcctgggaggcttgcttcaatg 2961


Query: 410   tcgtggagggttgttgagccgtagatttgggggctccgggtgccgaggaggttgaggttg 469
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct: 2960  tcatggagggttgttgagccgtagatttgggggctccgggtgccgaggaggttgagattg 2901


Query: 470   gggccattgatgaggagaatatgatgacgggacgccat 507
              |||||||||||||||||| ||| |  ||||||||||
Sbjct: 2900  gggccattgatgaggagaatgtgacgggggacgccat 2863
```

Figure 21: Sequence alignment of the coding regions of the *qa-2* genes of *N. africana* and of *N. crassa*.


Na qa-2 1 ATGGCGTCCCGTCATCATATTCTCCTCATCAATGGCCCCAACCTCAACCTCCTCGGCACC 60
Nc qa-2 1 atggcgtcccccccgtcacattctcctcatcaatggccccaatctcaacctcctcggcacc 60
Na qa-2 61 CGGGAGCCCCAAATCTACGGCTCAACAACCCTCCACGACATCGAGCAAGCCGCCCAGACT 120
Nc qa-2 61 cgggagccccaaatctacggctcaacaaccctccatgacattgagcaagcctcccagact 120
Na qa-2 121 CAGGCCTCCTCGCTAGATCTCCGTCTTACAACCTTCCAGTCCAACCACGAAGGCGCCATC 180
Nc qa-2 121 ctggcgtcctcgctaggtcttcgtcttacaaccttccagtccaaccatgaaggagccatc 180
Na qa-2 181 ATCGACCGTATCCACCAAGCAGCGGGA---------T-T----CTTC-CCGTCGCCGTCA 225
Nc qa-2 181 atcgaccgtatccatcaagcagcgggattcgtcccgtctccaccgtcaccgtcgccgtca 240
Na qa-2 226 GGTCCCGCAACCATCGCGGAAGCAGACCCCGGAGCCGGAGAAAAAGTGTCGGCCATCATC 285
Nc qa-2 241 agtgccgcaaccacgacggaggcaggattgggtcccggagacaaagtgtcggccatcatc 300
Na qa-2 286 ATCAACCCCGGCGCTTACACGCACACGAGCGTGGGCATCCGCGACGCACTGCTGGGGACA 345
Nc qa-2 301 attaaccccggcgcttatacgcacacgagtataggcatccgcgacgcgcttctggggaca 360
Na qa-2 346 GGAATCCCGTTCGTGGAGGTGCACGTGTCGAATGTGCATGCGAGAGAGGCGTTCAGGCAT 405
Nc qa-2 361 ggaattccgtttgtggaggttcatgtgtcgaatgtgcatgcgcgagaggcgttcaggcat 420
Na qa-2 406 CATAGTTACTTGAGTGATAAGGCGGTGGCAGTCATTTGCGGGTTGGGGCCGTATGGGTAT 465
Nc qa-2 421 catagttacttgagtgacaaggcggtggcagtcatttgcgggttgggggccgttatgggtat 480
Na qa-2 466 AGTGCTGCGCTGGAGTTTGTTGGGAGACACATGAAGTTTTGA 507
Nc qa-2 481 agtgctgcgctggactttcttgggagacacatgaagttttga 522

V. Comparison of the *qa-2* protein of *N. africana*, the *qa-2* protein of *N. crassa*, and the *qutE* protein of *Aspergillus nidulans*

The *qa-2* sequences were then translated using BioEdit and compared (Figure 22). There is a high degree of homology between the *qa-2* proteins of *N. africana* and *N. crassa*. A short sequence protein blast search revealed a bit score of 453 bits. At the protein level a high degree of homology is again found with an 87% sequence homology and a 94% similarity meaning that certain amino acids were replaced with similar amino acids. An example of this is the 81st residue of *N. africana* is similar to the 86th residue of *N. crassa*. Although they are different amino acids, a threonine in *N. crassa* is substituted with an alanine in *N. africana*, the amino acid used for substitution chemically similar.

Because similarities have existed in the past, a comparison to another fungus was then made. *Aspergillus nidulans* is from another family but retains the ability to utilize quinic acid through a slightly different pathway. The protein sequence for the QutE protein was pulled from the NCBI website. A short sequence protein blast search revealed a bit score of 270 bits. When the *qa-2* gene of *N. africana* is aligned in BioEdit to *QutE* of *A. nidulans*, there is a 56% amino acid sequence homology and a 70% amino acid sequence similarity (Figure 23). When SBASE was used to search the protein for domain, the only domain found was as expected in the acetyltransferase family (Data Not Shown).

Figure 22: Comparison of the *qa-2* protein of *N. africana* and the *qa-2* protein of *N. crassa*

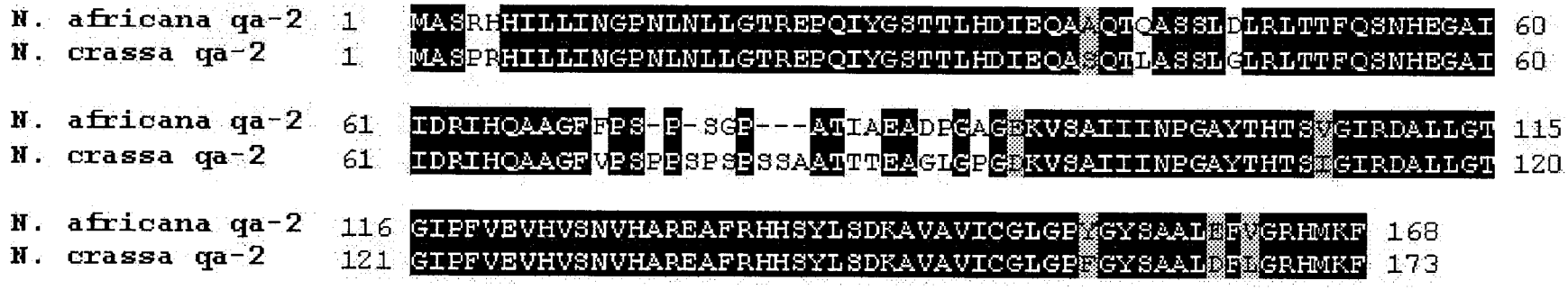

N. africana qa-2 1 MASRHHILLINGPNLNLLGTREPQIYGSTTLHDIEQAAQTQASSLDLRLTTFQSNHEGAI 60
N. crassa qa-2 1 MASPRHILLINGPNLNLLGTREPQIYGSTTLHDIEQASQTLASSLGLRLTTFQSNHEGAI 60
N. africana qa-2 61 IDRIHQAAGFFPS-P-SGP---ATIAEADPGAGFKVSAIIINPGAYTHTSYGIRDALLGT 115
N. crassa qa-2 61 IDRIHQAAGFVPSPPSPSPSSAATTTEAGLGPGFKVSAIIINPGAYTHTSIGIRDALLGT 120
N. africana qa-2 116 GIPFVEVHVSNVHAREAFRHHSYLSDKAVAVICGLGPLGYSAALFFYGRHMKF 168
N. crassa qa-2 121 GIPFVEVHVSNVHAREAFRHHSYLSDKAVAVICGLGPFGYSAALDFLGRHMKF 173

Figure 23: Comparison of the *qa-2* protein of *N. africana*, the *qa-2* protein of *N. crassa*, and the *qutE* protein of *Aspergillus nidulans*.


N. africana qa-2 1 MASRHHILLINGPNLNLLGTREPQIYGSTTLHDIEQAAQTQASSLDLRLTTFQSNHEGAI 60
N. crassa qa-2 1 MASPRHILLINGPNLNLLGTREPQIYGSTTLHDIEQASQTLASSLGLRLTTFQSNHEGAI 60
A. nidulans QutE 1 --MEKSILLINGPNLNLLGTREPHIYGSTTLSDVEESSKGHAASLGASLQTFQSNHEGAI 58
N. africana qa-2 61 IDRIHQAAGFFPS-P-SGP---ATIAEADPGAGEKVSAIIINPGAYTHTSVGIRDALLGT 115
N. crassa qa-2 61 IDRIHQAAGFVPSPPSPSPSSAATTTEAGLGPGDKVSAIIINPGAYTHTSIGIRDALLGT 120
A. nidulans QutE 59 VERIH----------AARG-NTDAIIINPGAYTHTSVAIRDALLGV 93
N. africana qa-2 116 GIPFVEVHVSNVHAREAFRHHSYLSDKAVAVICGLGPYGYSAALEFVGRHMKF 168
N. crassa qa-2 121 GIPFVEVHVSNVHAREAFRHHSYLSDKAVAVICGLGPFGYSAALDFLGRHMKF 173
A. nidulans QutE 94 EIPFTELHVSNVHAREPFRHHSYFSDKASGIIVGLGVYGYKVAVEHVA--LNFKPLEKKA 151
N. africana qa-2 168 168
N. crassa qa-2 173 173
A. nidulans QutE 152 AL 153

VI. Blast Search using *qa-2* protein sequences

When highly conserved areas of the *qa-2* protein of *N. africana* are used in a blast search, different fungal relationships begin to arise. A blast search revealed that *Aspergillus fumigatus*, *Candida albicans* and *Magnaporthe grisea* all contain this highly conserved protein segment (Figure 24).

VII. Sequence of entire Insert

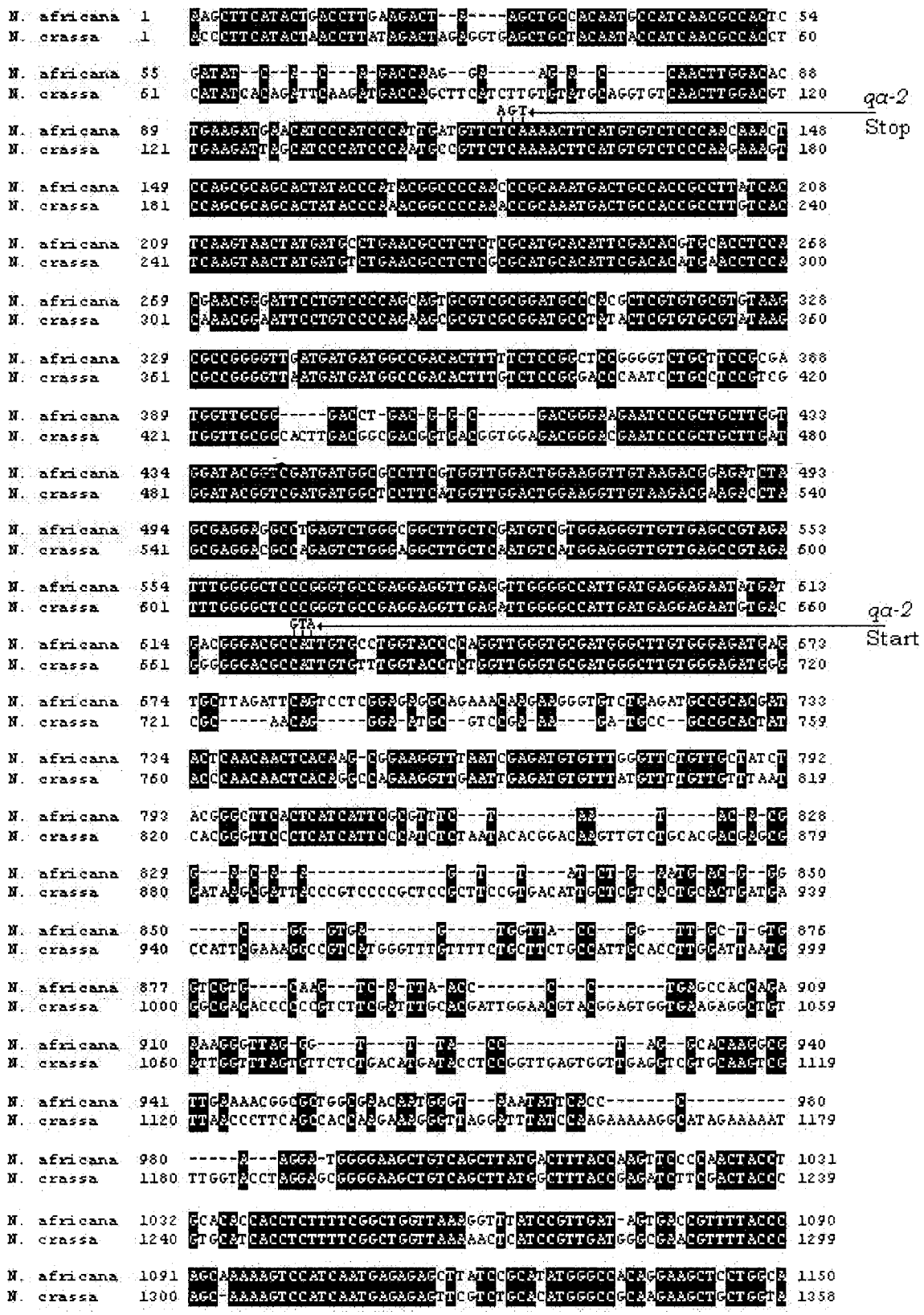
The entire NAL clone was sequenced. Some of this data, the *qa-2 – qa-4* intergenic region has previously been published (Asch, Orejas et al. 1991). Figure 25 is the results when you align the entire sequence to *N. crassa*.

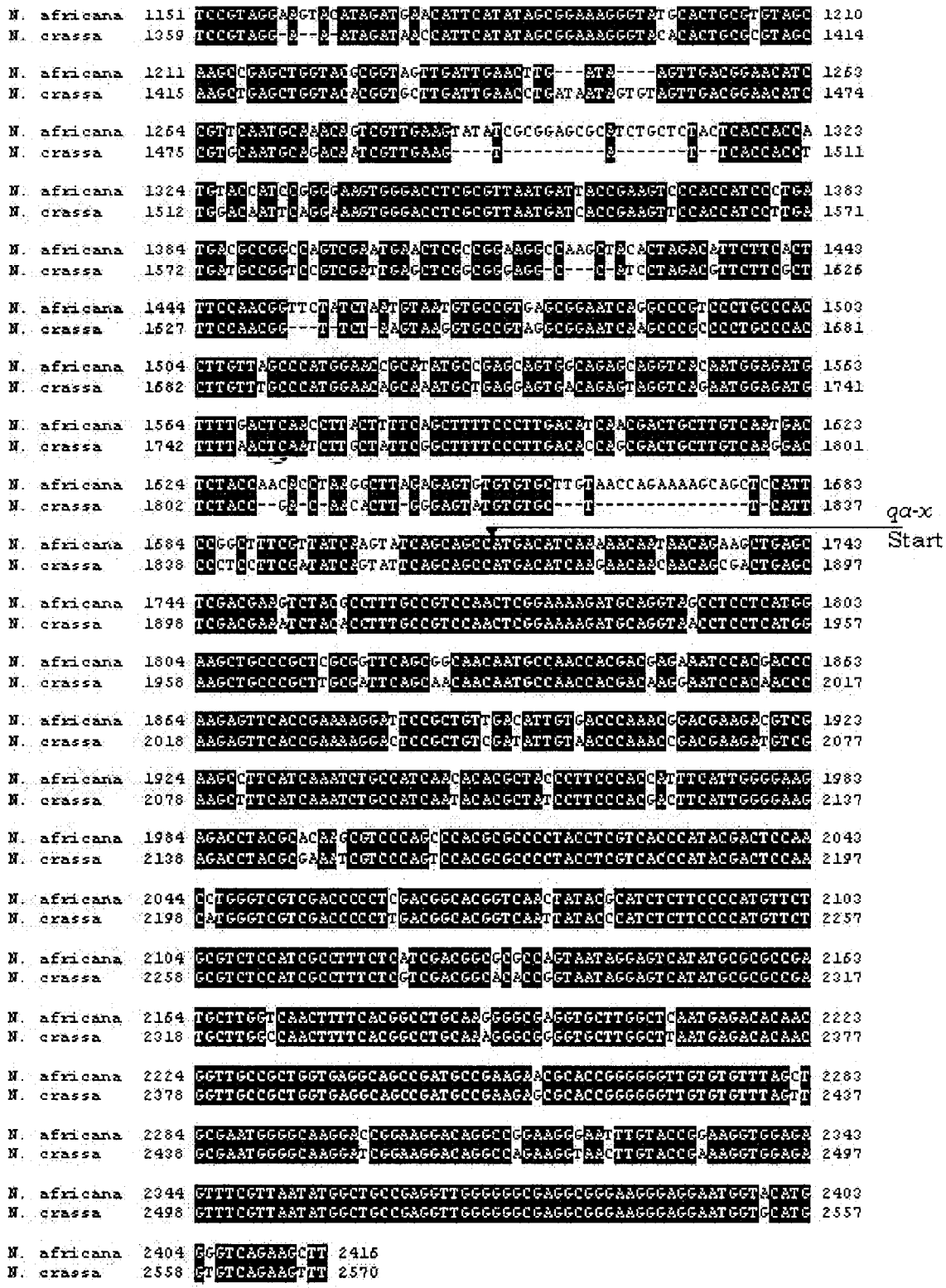
Figure 24: Blast results of interest when using a high homology segment of the *qa-2* protein of *N. africana*.

```
N. africana     1  ILLINGPNLNLLGTREP 17
A. fumigatus    1  ILLINGPNLNLLGTREP 17
C. albicans     1  VLLINGPNLNLLGTREP 17
M. grisea       1  ILLVNGPNLNLLGTREP 17
```

```
gi|41581237|emb|CAE47886.1|   catabolic 3-dehydroquinase,
putative [Aspergillus fumigatus]
          Length = 150

 Score = 56.6 bits (126), Expect = 6e-08
 Identities = 17/17 (100%), Positives = 17/17 (100%)

Query: 1  ILLINGPNLNLLGTREP 17
          ILLINGPNLNLLGTREP
Sbjct: 5  ILLINGPNLNLLGTREP 21


gi|46436470|gb|EAK95832.1|   hypothetical protein
CaO19.2283 [Candida albicans SC5314]
          Length = 146

 Score = 54.1 bits (120), Expect = 3e-07
 Identities = 16/17 (94%), Positives = 17/17 (100%)

Query: 1  ILLINGPNLNLLGTREP 17
          +LLINGPNLNLLGTREP
Sbjct: 5  VLLINGPNLNLLGTREP 21


gi|38107314|gb|EAA53505.1|   hypothetical protein MG07782.4
[Magnaporthe grisea 70-15]
          Length = 157

 Score = 52.8 bits (117), Expect = 8e-07
 Identities = 16/17 (94%), Positives = 17/17 (100%)

Query: 1  ILLINGPNLNLLGTREP 17
          ILL+NGPNLNLLGTREP
Sbjct: 8  ILLLNGPNLNLLGTREP 24
```

Figure 25: Entire NAL clone aligned to *N. crassa*. (Sequence displayed on next two pages.)


qa-2
Stop
qa-2
Start


qa-x
Start

Discussion

Different insights into structure and function relationship may be elucidated through comparative sequence analysis. Furthermore, light may be shed on the evolution of the particular subject at a molecular level. Well characterized systems like the *qa* gene cluster provide an excellent opportunity for comparative analysis.

A previous study on comparing *Neurospora crassa* to *Neurospora africana* (Asch, Orejas et al. 1991) has shown that there is a good deal of sequence homology. This means despite evolutionary changes, the *qa* cluster has remained highly conserved. Although quinic acid is not a highly palatable carbon source and *Neurospora* prefer numerous sources to qa, the *qa* cluster does serve a purpose to the organism as it has been maintained at a relatively high level.

Since the *qa-x – qa-2* intergeneic region is already sequenced; the next logical step is to continue sequencing the next gene in line, the *qa-2* genes from homothallic *Neurospora* species in an attempt to completely sequence the *qa* cluster so it may latter be used for more in-depth comparisons. In order to do this, a probe was used which contained the *qa-2* gene of *Neurospora crassa*. Lambda clones containing the *qa* gene clusters from *N. africana* and *N. terricola* were probed. The clone from *N. terricola* did not contain any *qa*-2 sequences; however, *qa-2* containing fragments were identified from *N. africana*. This allowed a fragment to be selected for subcloning, as smaller pieces of DNA are easier to work with than larger ones. A 2.4 Kb was selected because it had small size and appeared to hybridized the probe really well.

Once the fragment was selected, it was cloned into the bluescript vector. A test gel showed that the desired fragment had been cloned. The gel also allowed for a good estimation of how concentrated the DNA sample was.

Once the plasmid was isolated and purified, the next step was to sequence it. The first run of the sequencer using a pUC reverse primer and an M-13 forward primer revealed the location and orientation of the inserted fragment. Figure 26 illustrates the location of the insert within the *qa* gene cluster. Since the insert contained part of the *qa-x* and all of the previously sequenced intergenic region, two methods were used to select the second set of primers. The goal was to select a primer that bound close to the end of the data so the sequence could be extended. The forward primer was selected based on the sequencing data. Previously published data (Asch, Orejas et al. 1991) was used to determine the reverse primer. This one was picked so that it was near the start of *qa-2*.

An overall comparison between *N. africana* and *N. crassa* shows that the DNA sequences are 91% identical to each other. Comparative analysis of the intergenic region is an excellent way to search for upstream elements involved in gene regulation (Asch, Orejas et al. 1991). Conserved intergenic regions are a good indication of a binding site for a DNA binding protein such as the activator-binding site for *qa-4*. Since the sequenced DNA itself is a gene, this sequence information becomes useful when looking for species differentiation or overall importance of that particular gene.

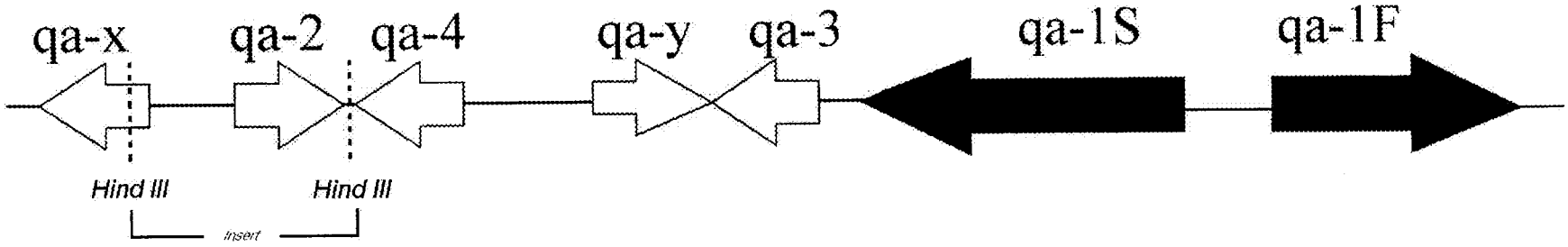
When the amino acid sequences of *N. africana* and *N. crassa* are compared, they are found to be 87% identical. When similar amino acid substitutions are taken into consideration the two *Neurospora* sequences show a 94% level of similarity. The overall of this protein comparison is 453 bits. This further supports the need that

Neurospora has for the *qa-2* gene. Since other fungi are known to metabolize qa, comparing this protein to other organisms may prove beneficial.

Because past comparisons between the two organisms had been done and it is known to metabolize qa, *A. nidulans* was chosen for assessment. The comparison between *N. africana* and *A. nidulans* reveals that 56% of the amino acid sequence is conserved; moreover, that most of the identity is found at the N and C terminals. This suggests that the N and C terminals are the functional part of the protein. When similar amino acids are substituted, the two sequences show a 70% level of homology. The overall score of this protein comparison is 270 bits. A domain search showed that this protein belongs to the acetyltransferase family. This is expected, as it is a dehydroquinase. No portions of this protein were analogous to any other protein families. This ruled out the possibility that this conserved region served a function other than a dehydroquinase.

A blast search of the highly conserved regions of the *qa-2* protein revealed that other fungi such as *Aspergillus fumigatus*, *Candida albicans* and *Magnaporthe grisea* also contain this highly conserved region somewhere within their proteome; moreover, the function of the protein containing this segment may be as a dehydroquinase. Future studies may further this finding by probing these other proteomes with antibodies that are specific to that conserved region. The protein may then be isolated and sequenced. A similar protein sequence would suggest similar function.

Figure 26: Map of the constructed clone in relation to the *qa* cluster.


qa-x
qa-2
qa-4
qa-y
qa-3
qa-1S
qa-1F
Hind III
Hind III
insert

Bibliography

Arnett, D. (MS Thesis 2000). Site-directed mutagenesis of the -127 activator binding site of the qa-2 gene of neurospora crassa: 1 v. (various pagings).

Asch, D. K., M. Orejas, R. F. Geever and M. E. Case (1991). "Comparative studies of the quinic acid (qa) cluster in several Neurospora species with special emphasis on the qa-x-qa-2 intergenic region." Mol Gen Genet **230**(3): 337-44.

Battogtokh, D., D. K. Asch, M. E. Case, J. Arnold and H.-B. Schuttler. (2002). "An Ensemble Method for Identifying Regulatory Circuits with Special Reference to the qa Gene Cluster of *Neurospora crassa*." Proceedings of the National Academy of Sciences **99**: 16904-16909.

Baum, J. A., R. Geever and N. H. Giles (1987). "Expression of qa-1F activator protein: identification of upstream binding sites in the qa gene cluster and localization of the DNA-binding domain." Mol Cell Biol **7**(3): 1256-66.

Baum, J. A. and N. H. Giles (1985). "Genetic control of chromatin structure 5' to the qa-x and qa-2 genes of Neurospora." J Mol Biol **182**(1): 79-89.

Bevan, P. and H. C. Douglas (1969). "Genetic control of phosphoglucomutase variants in Saccharomyces cerevisiae." J Bacteriol **98**(2): 532-5.

Blank, T. E., M. P. Woods, C. M. Lebo, P. Xin and J. E. Hopper (1997). "Novel Gal3 proteins showing altered Gal80p binding cause constitutive transcription of Gal4p-activated genes in Saccharomyces cerevisiae." Mol Cell Biol **17**(5): 2566-75.

Borkovich, K. A., L. A. Alex, O. Yarden, M. Freitag, G. E. Turner, N. D. Read, S. Seiler, D. Bell-Pedersen, J. Paietta, N. Plesofsky, M. Plamann, M. Goodrich-Tanrikulu, U. Schulte, G. Mannhaupt, F. E. Nargang, A. Radford, C. Selitrennikoff, J. E. Galagan, J. C. Dunlap, J. J. Loros, D. Catcheside, H. Inoue, R. Aramayo, M. Polymenis, E. U. Selker, M. S. Sachs, G. A. Marzluf, I. Paulsen, R. Davis, D. J. Ebbole, A. Zelter, E. R. Kalkman, R. O'Rourke, F. Bowring, J. Yeadon, C. Ishii, K. Suzuki, W. Sakai and R. Pratt (2004). "Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism." Microbiol Mol Biol Rev **68**(1): 1-108, table of contents.

Case, M. E., R. F. Geever and D. K. Asch (1992). "Use of Gene Replacement Transformation to Elucidate Gene Function in the qa Gene Cluster of Neurospora crassa." Genetics **130**(4): 729-736.

Davis, R. H., and F.J. deSerres (1970). "Genetic and microbiological research techniques for Neurospora crassa." Meth. Enzymology **17**: 79-143.

Douglas, H. C. and C. D. Hawthorne (1972). "Uninducible mutants in the gal i locus of Saccharomyces cerevisiae." J Bacteriol **109**(3): 1139-43.

Douglas, H. C. and D. C. Hawthorne (1964). "Enzymatic Expression and Genetic Linkage of Genes Controlling Galactose Utilization in Saccharomyces." Genetics **49**: 837-44.

Galagan, J. E., Sarah E. Calvo, Katherine A Borkovich, Eric U. Selker, et al. (2003). "The Genome Sequence of the Filamentous Fungus *Neurospora crassa*." Nature **422**: 859-869.

Geever, R., L Huiet, J.A. Baum, BM Tyler, VB Patel, BJ Rutledge, ME Case, NH Giles. (1989). "DNA Sequence, Organization and Regulation of the qa Gene Cluster of *Neurospora crassa*." Journal of Molecular Biology **207**: 15-34.

Geever, R. F., J. A. Baum, M. E. Case and N. H. Giles (1987). "Regulation of the QA gene cluster of Neurospora crassa." Antonie Van Leeuwenhoek **53**(5): 343-8.

Geever, R. F., M. E. Case, B. M. Tyler, F. Buxton and N. H. Giles (1983). "Point mutations and DNA rearrangements 5' to the inducible qa-2 gene of Neurospora allow activator protein-independent transcription." Proc Natl Acad Sci U S A **80**(23): 7298-302.

Giles, N. H., M. E. Case, J. Baum, R. Geever, L. Huiet, V. Patel and B. Tyler (1985). "Gene organization and regulation in the qa (quinic acid) gene cluster of Neurospora crassa." Microbiol Rev **49**(3): 338-58.

Grant, S., C. F. Roberts, H. Lamb, M. Stout and A. R. Hawkins (1988). "Genetic regulation of the qunic acid utilization (*QUT*) gene cluster in *Aspergillus nidulans*." J. Gen. Microbiol **134**: 347-350.

Hashimoto, H., Y. Kikuchi, Y. Nogi and T. Fukasawa (1983). "Regulation of expression of the galactose gene cluster in Saccharomyces cerevisiae. Isolation and characterization of the regulatory gene GAL4." Mol Gen Genet **191**(1): 31-8.

Hawkins, A. R., H. K. Lamb, M. Smith, J. W. Keyte and C. F. Roberts (1988). "Molecular organization of the quinic acid utilization (*QUT*) gene cluster in *Aspergillus nidulans*." Mol. Gen. Genet **214**: 224-231.

Huiet, L. (1984). "Molecular analysis of the Neurospora qa-1 regulatory region indicates that two interacting genes control qa gene expression." Proc Natl Acad Sci U S A **81**(4): 1174-8.

Huiet, L. and M. E. Case (1985). Gene Manipulations in Fungi, Academic Press.

Johnston, M. (1987). "A model fungal gene regulatory mechanism: the GAL genes of Saccharomyces cerevisiae." Microbiol Rev **51**(4): 458-76.

Johnston, M. and J. Dover (1987). "Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA binding domain." Proc Natl Acad Sci U S A **84**(8): 2401-5.

Kang, T., T. Martins and I. Sadowski (1993). "Wild type GAL4 binds cooperatively to the GAL1-10 UASG in vitro." J Biol Chem **268**(13): 9629-35.

Keller, N. P. and T. M. Hohn (1997). "Metabolic Pathway Gene Clusters in Filamentous Fungi." Fungal Genetics and Biology **21**: 17-29.

Leloir, L. F. (1951). "The enzymatic transformation of uridine diphosphate glucose into a galactose derivative." Arch Biochem **33**(2): 186-90.

Lue, N. F., D. I. Chasman, A. R. Buchman and R. D. Kornberg (1987). "Interaction of GAL4 and GAL80 gene regulatory proteins in vitro." Mol Cell Biol 7(10): 3446-51.

Mishra, N. C. (1991). "Genetics and molecular biology of Neurospora crassa." Adv Genet **29**: 1-62.

Patel, V. B. and N. H. Giles (1985). "Autogenous regulation of the positive regulatory qa-1F gene in Neurospora crassa." Mol Cell Biol **5**(12): 3593-9.

Patel, V. B., M. Schweizer, C. C. Dykstra, S. R. Kushner and N. H. Giles (1981). "Genetic organization and transcriptional regulation in the qa gene cluster of Neurospora crassa." Proc Natl Acad Sci U S A **78**(9): 5783-7.

Perkins, D. D. (1992). "Neurospora: the organism behind the molecular revolution." Genetics **130**(4): 687-701.

Seale, T. (1973). "Life cycle of Neurospora crassa viewed by scanning electron microscopy." J Bacteriol **113**(2): 1015-25.

Silver, P. A., L. P. Keegan and M. Ptashine (1984). "Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization." Proc Natl Acad Sci U S A **81**(19): 5951-5.

Stroman, P., W. R. Reinert and N. H. Giles (1978). "Purification and characterization of 3-dehydroshikimate dehydratase, an enzyme in the inducible quinic acid catabolic pathway of Neurospora crassa." J. Biol. Chem. **253**(13): 4593-4598.

Taylor, J. W., Barbara H. Bowman, Mary L. Berbee, and Thomas J. White (1993). "Fungal Model Organisms: Phylogenetics of *Saccharomyces*, *Aspergillus* and *Neurospora*." *Syst. Biol.* **42**(4): 440-457.

Torchia, T. E., R. W. Hamilton, C. L. Cano and J. E. Hopper (1984). "Disruption of regulatory gene GAL80 in Saccharomyces cerevisiae: effects on carbon-controlled regulation of the galactose/melibiose pathway genes." Mol Cell Biol **4**(8): 1521-7.

Vapnek, D., J. A. Hautala, J. W. Jacobson, N. H. Giles and S. R. Kushner (1977). "Expression in Escherichia coli K-12 of the structural gene for catabolic dehydroquinase of Neurospora crassa." Proc Natl Acad Sci U S A **74**(8): 3508-12.

Whittington, H. A., S. Grant, C. F. Roberts, H. Lamb and A. R. Hawkins (1987). "Identification and isolation of a putative permease gene in the quinic acid utilization (QUT) gene cluster of Aspergillus nidulans." Curr Genet **12**(2): 135-9.