Imperial College London

Development of silencing vectors for *Aspergillus fumigatus* based on mycoviruses and short interspersed nuclear elements

LAKKHANA KANHAYUWA

A thesis submitted for the degree of Doctor of Philosophy and the Diploma of Membership of Imperial College

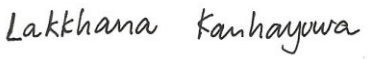
Imperial College London

Department of Life Sciences

Division of Cell and Molecular Biology

Declaration of Originality

I hereby declare that this thesis is the product of my own work unless otherwise acknowledged in the text or by references and has not been previously submitted for a degree in any university or other purposes.



Signed: Lakkhana Kanhayuwa

Date: September, 2015

Copyright Declaration

Abstract

Title of Thesis: Development of silencing vectors for *Aspergillus fumigatus* based on mycoviruses and short interspersed nuclear elements.

A novel mycovirus named Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1) was discovered and characterized in the human pathogenic fungus*, A. fumigatus* clinical isolate Af293. The virus reveals several unique features not previously found in double-stranded RNA (dsRNA) viruses and represents the first dsRNA that is infectious both as a purified entity and a naked dsRNA. The AfuTmV-1 is an unencapsidated dsRNA mycovirus comprised of four genomic segments, ranging in size from 1.1 to 2.4 kbp. The largest component encodes a putative viral RNA-dependent RNA polymerase (RdRP) where the sequence of the most highly conserved motif changes from GDDX to GDNQ. The third largest dsRNA encodes an S-adenosyl methionine-dependent methyltransferase (SAM) capping enzyme and the smallest dsRNA encodes a proline-alanine rich protein. Short interspersed nuclear elements (*SINEs*) were also identified in the fungal genome. Identification of the elements revealed tRNA-related and 5S rRNA-related *SINE* families which showed variation in transcription activity and copy number. AfuTmV-1 sequences together with *SINEs* were subsequently exploited to develop alternative tools for silencing genes in *A. fumigatus*. A truncated AfuTmV-1 based vector was successfully constructed and used as a prototype vector for generating a recombinant virus-induced gene silencing (VIGS) vector. Transcriptional fusion *SINE*-derived vectors were also developed to silence an *ALB1/PKSP* gene responsible for conidial pigmentation. With anticipation that the development will provide a powerful reverse genetic tool for functional genomics studies to identify key elements involved in fungal pathogenicity and also provide a medical benefit in exploiting mycoviruses as a future therapeutic agent against fungal infections.

RNA silencing | mycovirus | *SINEs* | *Aspergillus fumigatu*s

Table of Contents

List of Figures

List of Tables

Abbreviations

µg	Microgram
µJ	Microjoule
µl	Microliter
µm	Micrometre
µM	Micromolar
aa	Amino cid
ABPA	Allergic bronchopulmonary aspergillosis
ACM	*Aspergillus* Complete Media
ACN	Acetonitrile
AFM	Atomic force microscopy
AMM	*Aspergillus* Minimal Medium
APS	Ammonium persulfate
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary DNA
cfu	Colony-forming unit
Ci	Curie
cm^3	Cubic centimetres
CsCl	Caesium chloride
CSP	Cell surface protein
CTP	Cytidine triphosphate
$CuSo_4.5H_2O$	Copper sulphate pentahydrate
DEPC-H_2O	Diethylpyrocarbonate H_2O
dH_2O	Distilled water
DMSO	Dimethylsulphoxide
dNTP	Deoxynucleotide triphosphate
ds	Double-stranded

DTT	Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EN	Endonuclease
EtOH	Ethanol
$FePO_4.2H_2O$	Iron(III) phosphate dihydrate
g	Gram
GFP	Green fluorescent protein
GMP	Guanosine monophosphate
GTP	Guanosine-5'-triphosphate
h	Hour
HCl	Hydrochloric acid
HIGS	Host-induced gene silencing
HMW	High molecular weight
HR	Homologous recombination
IA	Invasive aspergillosis
IE	Intermediate element
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kbp	Kilobase pair
KCl	Potassium chloride
kDa	Kilodalton
KH_2PO_4	Potassium phosphate
l	Litre
LB	Luria-Bertani
LiCl	Lithium chloride
LINE	Long interspersed nuclear element
LLE	*LINE*-like element
LMW	Low molecular weight
LTR	Long terminal repeat
M	Molar
mg	Milligram

$MgCl_2$	Magnesium chloride
$MgSO_4.7H_2O$	Magnesium sulphate heptahydrate
min	Minute
miRNA	MicroRNA
ml	Millilitre
mm	Millimetre
mM	Millimolar
$MnSO_4.2H_2O$	Manganese sulphate dihydrate
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	Messenger RNA
MS	Mass spectrometry
MW	Molecular wight
N	Normal
$Na_2B_4O_7.10H_2O$	Di-sodium tetraborate decahydrate
NaCl	Sodium chloride
Na_2HPO_4	Disodium hydrogen phosphate
$NaH_2PO_4 .H_2O$	Sodium dihydrogen phosphate
$Na_2MoO_4.2H_2O$	Sodium molybdate
NaOH	Sodium hydroxide
ncRNA	Non-coding RNA
ng	Nanogram
nm	Nanometre
nt	Nucleotide
OD	Optical density
OM	Osmotic medium
ORF	Open reading frame
PASrp	Proline, alanine, serine-rich protein
PCD	Programme cell death
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMF	Peptide mass fingerprinting

rasiRNA	Repeat-associated small interfering RNA
rcf	Relative centrifugal force
RdRP	RNA dependent RNA polymerase
RISC	RNA-induced silencing complex
RLM-RACE	RNA linker-mediated rapid amplification of cDNA end
RNA pol III	RNA polymerase III
RNAi	RNA interference
rPCR	Random-PCR
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-PCR	Reverse transcription polymerase chain reaction
SAM	S-adenosyl methionine-dependent methyltransferase
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
sec	Second
SINE	Short interspersed nuclear element
siRNA	Small interfering RNA
shRNA	Short hairpin RNA
SMM	Stabilized minimal medium
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
SOC medium	Super optimal broth with catabolite repression
SSC	Saline Sodium Citrate
ss	Single-stranded
STC	Sorbitol/tris/calcium chloride solution
TAE	Tris acetate EDTA
TBE	Tris-borate EDTA
TE	Transposable element
TE buffer	Tris/EDTA buffer
TEM	Transmission electron microscopy

TEMED	N,N,N’,N’-tetramethylethylene diamine
TFA	Trifluoroacetic acid
Tm	Melting temperature
tRNA	Transfer RNA
TSD	Target site duplication
TSS	Transformation storage solution
U	Unit
UTP	Uridine-5'-triphosphate
UTR	Untranslated region
V	Volt
VCG	Vegetative compatibility group
vic	Vegetative incompatibility
VIGS	Virus-induced gene silencing
VLP	Virus-like particle
$ZnSO_4.7H_2O$	Zinc sulphate heptahydrate

Acknowledgment

First, I would like to express my deepest gratitude to the Royal Thai Government education programme for supporting my PhD.

I would like to express my special appreciation and thanks to my supervisor Dr. Robert H.A. Coutts for having been a tremendous mentor for me all through the entire research. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been precious. I would like to thank Professor Pietro Spanu, Professor Martin Buck and Dr. Colin Turnbull from the department of Life Sciences at Imperial College London for your valuable advice and guidance. I would also like to thank Professor Matthew Fisher from the department of Infectious Disease Epidemiology at Imperial College School of Public Health and Professor George Lomonossoff from John Innes Centre at Norwich Research Park for serving as my examiners. I also would like to thank you for letting my defense be an enjoyable moment and for your comments and suggestions. I would especially like to thank Dr. Elaine Bignell from the MRC Centre for Molecular Bacteriology and Infection at Imperial College London for an opportunity to use laboratory facilities and for your special advice. Besides, I gratefully thank Selin Ozkan from the department of Life Sciences and all of the staff in the *Aspergillus* Laboratory for their continuous support and assistance in my research. Additionally, I would like to express my special thanks to all the staff members in the department of Life Sciences, particularly to Ian Morris and Fiona May for your indispensable and kind help all through the entire research.

I would like to give special thanks to my family. Words cannot express how grateful I am to my mother, my father and my sister for all of the sacrifices that they have made on my behalf. I would also like to thank all of my friends who supported me in working and writing and encouraged me to strive towards my goal. At the end I would like express appreciation to Paul A. Wingfield who is always be my supporter.

Lakkhana Kanhayuwa

September, 2015

Introduction and Project Objectives

An opportunistic fungus, *Aspergillus fumigatus*, is the most commonly encountered fungal pathogen of humans. Increased incidence of infection and development of pulmonary aspergillosis in immunocompromised patients, coupled with developing fungal resistance to common antifungal agents (e.g. azoles) are currently a major concern. The completion of the *A. fumigatus* genome sequence provides an opportunity to investigate genes and molecules that regulate pathogenicity and drug resistance of the fungus. In order to investigate factors involved in fungal pathogenicity, genes associated with the process need to be identified and targeted. However most of the genes present within *A. fumigatus* genome are essential for its survival and knock out mutants of such genes would be lethal for pathogenicity. Thus, a method which down-regulates rather than deletes gene function would have wider application. To this purpose, the use of RNA silencing, a process associated with specific degradation of a target gene *via* sequence-homology-dependent mechanism, could substitute for a limitation of the conventional approach.

At present, hairpin RNA-expressing vector is a versatile tool commonly used to induce RNA silencing in fungi; however, the approach has drawbacks ranging from difficulty in plasmid preparation to instability of hairpin constructs in the fungal cells. As a consequence, an alternative tool which could overcome these constraints is required for fast and transient suppression of gene expression. In this regard, recently proposed approaches such as the use of virus sequences and short interspersed nuclear elements (*SINEs*) for construction of silencing vectors have been developed in several filamentous fungi. These demonstrated various degrees of silencing levels depending on the viruses, *SINE* sequences and target genes used (Suzuki *et al*., 2000; Vetukuri *et al*., 2011a; Panwar *et al*., 2013; Mascia *et al*., 2014).

RNA silencing mechanism has been described in several virus-infected fungi as a natural defense response against virus infection (Segers *et al*., 2007; Hammond *et al*., 2008b; Himeno *et al*., 2010). According to this concept following the principle of virus replication, recombinant virus and viral RNA derivatives corresponding to the virus target sequence have been developed to trigger the process responsible for RNA silencing to reduce virus accumulation in the fungal hosts. Similarly, engineering of an insert fragment homologous to the host endogenous target gene into the virus genome could be employed which would lead to down-regulation of the targeted transcripts. Thus, current knowledge of RNA silencing as an antiviral

defense response, coupled with a new insight on mycoviruses (the viruses which selectively infect fungi), have provided opportunities in developing tools which can be useful in the analysis of gene function and key elements involved in *A. fumigatus* virulence.

Prior to exploiting mycoviruses as a potential vector for silencing genes in *A. fumigatus*, a virus candidate suitable for incipient silencing by fungal silencing machinery and its genomic sequences including virus properties needed to be characterized. A novel Aspergillus fumigatus tetramycovirus 1 (AfuTmV-1) isolated from the *A. fumigatus* clinical isolate Af293 which appears to be an unencapsidated virus was selected for this purpose. The genome of AfuTmV-1 was cloned and sequenced with the long term goal of exploiting elements of its dsRNA genome in viral vector development. Thereafter, a prototype gene-silencing vector based on the AfuTmV-1 was constructed using dsRNA2 as a backbone. Then an untranslatable fragment of a target gene was inserted to replace an internal region of the virus. The modified virus vector was subsequently transfected into fungal protoplasts to trigger silencing of the host endogenous gene homologous to the insert fragment.

As the genome sequences of several fungi are annotated, it is clear that transposable elements, the activity of which may be controlled by RNA silencing, are key genomic components for genome organization and evolution. It has been demonstrated that small interfering RNAs (siRNAs) corresponding to the sequences of *SINEs* accumulated in the fungal cells, suggesting that siRNAs might be involved in silencing of *SINEs* (Whisson *et al*., 2005a, Vetukuri *et al*., 2011a). In addition, it has been described that retrotransposon transcriptional activation can modify the expression of an adjacent sequence (Anbar *et al*., 2005, Whisson *et al*., 2005a, Lerat and Semon, 2007). This notion has currently been exploited to develop a novel tool for gene silencing in fungi by engineering transcriptional fusions of initially marker genes to *SINE* sequence, for example, in *Phytophthora infestans* (Vetukuri *et al*., 2011a).

Initially, the abundance and distribution of *SINE* families in the *A. fumigatus* Af293 genome were identified by interrogating the genomic DNA sequence for the occurrence of well-characterized signature motif sequences using bioinformatics programmes. Based on the sequences of putative *SINEs* incorporating either 5'-tRNA-related regions containing RNA polymerase III A and B box internal promoter sequences or 5'-5S rRNA-related regions containing RNA polymerase III A, IE and C boxes, oligonucleotide primers were designed for RT-PCR amplification to examine transcriptional activity of the predicted *SINEs*. For those elements that are transcriptionally active, southern hybridisation was performed to estimate *SINE* copy number. Subsequently the most transcriptionally active *SINE* was selected to

construct *SINE*-based silencing vectors to develop a novel tool for gene silencing in *A. fumigatus*. In addition, to understand more on the regulation of retrotransposable elements and gene silencing in *A. fumigatus*, siRNAs homologous to *SINEs* was identified by northern blot analysis.

As a proof of the aforementioned principle that silencing of the target *SINE* will also target transcriptionally fused marker gene transcripts and in order to test whether the technique is applicable in *A. fumigatus*, construction of silencing vectors by transcriptional fusion of a portion of gene of interest to *SINE* sequence and examination of silencing efficiency of the vectors were performed. A proposed endogenous *A. fumigatus* gene used as a target for silencing was a gene involved in morphological development - the *alb1/pksp* gene which is responsible for melanin production and conidial pigmentation. Thereafter essential genes involved in *A. fumigatus* pathogenicity could be selected for further silencing study. *SINE*-based silencing constructs were then transformed into the fungal protoplasts. Subsequently, change in morphology in term of colony pigmentation was observed on the regenerated lines after exposure to the vectors.

Research Outcome

An outcome of the dsRNA-triggered gene silencing study would establish potential of the approaches to silence genes in order to determine gene regulation and gene function in the *A. fumigatus* and in other *Aspergillus* species. Simultaneously, discovery of a novel AfuTmV-1 mycovirus in *A. fumigatus* would provide an opportunity to develop the virus as a future therapeutic tool and biological control agent for human fungal infection and also would enhance the knowledge on virus diversity in fungi and virus-host evolution. Additionally, identification of *SINEs* would provide the knowledge on the contribution of the elements on genome organization and genome re-arrangement in *A. fumigatus* including an involvement of the elements on the evolution of fungal pathogenicity.

Project Objectives

1) To characterize a novel mycovirus, AfuTmV-1, isolated from the *A. fumigatus* Af293 clinical isolate.

2) To identify new families of short interspersed nuclear elements (*SINEs*) in the *A. fumigatus* genome.

3) To construct silencing vectors for gene silencing in *A. fumigatus* based on the sequence of AfuTmV-1 mycovirus and *SINE*.

Chapter 1

Literature Review

Chapter 1: Literature Review

1. Mycoviruses

1.1 General characteristics of mycoviruses

Mycoviruses are viruses which selectively infect fungi. The genomes of which are commonly double-stranded RNA (dsRNA) or single-stranded RNA (ssRNA) that are encapsidated in a variety of particle morphologies (including sphearical, flexuous rods, club-shaped particles, rigid rod, enveloped bacilliform and herpesvirus-like virus), but mainly in an isometric form. Besides, many unencapsidated mycoviruses have been described. Mycoviruses with ssDNA and dsDNA genomes are also known in phytopathogenic fungi such as Gemini-like DNA virus (Yu *et al*., 2010). The viruses are found to be widely distributed over various fungal phyla such as *Chytridiomycota*, *Zygomycota*, *Ascomycota* and *Basidiomycota* including plant pathogenic oomycetes such as *Phytophthora*. One or more unrelated mycoviruses can infect one particular fungal strain and can be accompanied by defective dsRNA and/or satellite dsRNA (Ghabrial and Suzuki, 2008). Generally, the host range of mycoviruses is restricted to closely related vegetative compatibility groups (VCGs) because cytoplasmic fusion and mating capability among different fungal groups can lead to programme cell death (PCD) event. However, some mycoviruses have been reported to be able to propagate in the fungi from different taxonomic groups.

The vast majority of mycoviruses characterized so far possess RNA genomes which makes the RNA viruses prone to high mutation rate. However, this occurrence is considered to be evolutionarily beneficial for the virus to adapt better in new environment and also contributed to mycovirus diversity. To date, information on mycovirus origins and evolution are insufficient. Thus, origins and evolutionary history of the virus are tentatively hypothesized into two aspects. First hypothesis is 'ancient co-evolution' which is proposed that ancient infection of the virus from an unknown source to the fungus have long co-evolved and given rise to present-day diversity. The second is 'plant virus hypothesis' which proposed that viruses from the host plant have migrated to the fungus (Ghabrial, 1998).

1.1.1 Double-stranded RNA (dsRNA) mycoviruses

Double-stranded RNA (dsRNA) mycoviruses have been reported in yeasts, filamentous fungi and mushrooms (Lee *et al*., 2011). They usually encapsidate in non-enveloped isometric or spherical virus particles with an average size of 25-50 nm in diameter or exist as unencapsidated virus associated with mitochondria or with endomembrane vesicle in the host cytoplasm. Currently, dsRNA mycoviruses are divided into seven major families based on their structures, genome segment numbers and genomic composition, which are *Totiviridae* (non-segmented virus with 4.6-7 kbp genome length), *Partitiviridae* (2 segmented virus with 1.4-2.3 kbp genome length), *Chrysoviridae* (4 segmented virus with 2.4-3.6 kbp genome length), *Reoviridae* (10-12 segmented virus with 0.7-5 kbp genome length), *Megabirnaviridae* (2 segmented virus with *ca.* 17 kbp genome length) and newly reported families *Quadriviridae* (4 segmented virus with 3.7-4.9 kbp genome length) and *Alternaviridae* (4 segmented virus with 1.4-3.6 kbp genome length) (Darissa *et al*., 2011; Lin *et al*., 2012). Recently, the increasing numbers of dsRNA mycoviruses have been described but many of them are still unassigned to a certain genus or a family.

RNA mycoviruses in the families *Hypoviridae* (13 kbp genome), *Narnaviridae* (2.5 kbp genome) and *Endornaviridae* (14 to over 17 kbp genome) which exist mostly as the dsRNA species in their hosts, are now proposed that the presence of dsRNA form is the replicative intermediate of ssRNA genome as same as observed in ss(+)RNA mycoviruses (Nuss, 2005a and 2005b). Most viruses from these families are described as unencapsidated virus as they do not form a typical virus particle.

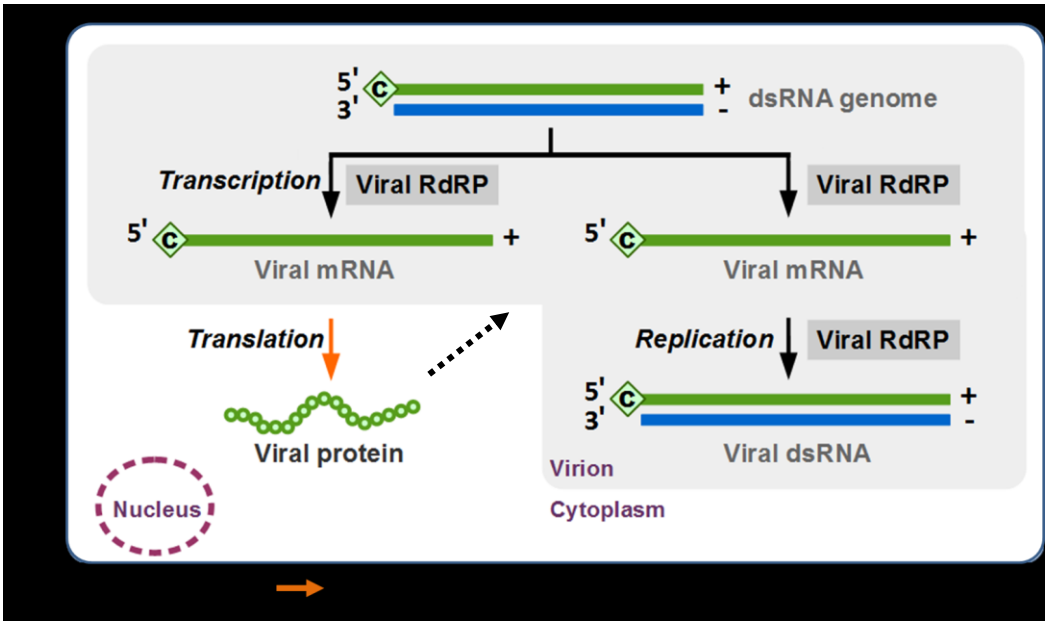
1.1.2 Replication of dsRNA mycoviruses

All RNA viruses with double stranded genomes contain an RNA-dependent RNA polymerase (RdRP) to catalyze the synthesis of the viral genome and mRNAs. The RNA polymerase is required for the virus because viral dsRNA cannot be directly translated or copied by the host cellular machinery. The mechanism by which viral mRNA of dsRNA viruses is synthesized and the RNA genome is replicated in the infected cells appears diverse due to differences in structure and organization of the viral RNA genomes. Two main strategies for efficient dsRNA replication have been proposed that first the RNA genome must be copied from one terminus to the other with no loss of nucleotide sequence. The second requirement is the viral mRNAs must be efficiently translated by the cellular protein synthetic machinery.

Following infection by encapsidated dsRNA viruses, virus particles are partially uncoated leaving the genome segments and transcriptional enzymes in the subviral core. Generally, the genomic dsRNAs do not readily serve as templates for translation, therefore the duplex genome needs to be partially unwound to serve as a template for (+) sense RNA synthesis. Messenger RNAs (mRNAs) are subsequently produced by copying of one strand of each duplex genome into a full-length strand *via* viral polymerase and are available for both translation and/or replication. In some dsRNA viruses, mRNAs may be naked, capped and methylated by viral enzymes such as a guanyltransferase and methyltransferase, or covalently linked to a viral protein at the 5'- terminus. Some dsRNA virsues have polyadenylated 3' termini while some have not e.g. reoviruses. Other viral enzymes required for transcriptional activity include a polymerase, a helicase and an RNA triphosphatase. In reoviruses, capped mRNAs are extruded from the viral core following translation of viral proteins necessary for replication and encapsidation in the cytoplasm *via* cellular ribosomes. After protein translation, structural proteins assemble around the newly synthesized (+) sense RNA followed by the synthesis of (-) antisense RNA *via* viral RdRP activity resulting in the formation of the dsRNA genome within the virion. The various genome segments can be transcribed/translated at different frequencies (Fig. 1a).

Eukaryotic hosts possess various antiviral defences that can degrade dsRNA. As a consequence many dsRNA viruses replicate their RNA inside the capsid and do not expose their dsRNAs to the cytoplasm protecting the genomic materials from celluar dsRNA defence responses. For unencapsidated dsRNA viruses, replication strategy and mechanisms used to avoid host antiviral response are even more sophisticated and not completely understood.

RNA can be replicated semi-conservatively or conservatively (Fig. 1b). In the semi-conservative model, each parental strand synthesizes a copy of itself. After one round of replication, the two daughter molecules each contain one original and one new strand. In conservative model, the parental molecule conducts synthesis of an entirely new double stranded molecule. After one round of replication, one molecule is conserved as two original strands and the other produce a copy comprised of two new strands. In reoviruses, RNA is transcribed conservatively where only (-) antisense strands are used, resulting in synthesis of (+) sense mRNAs.



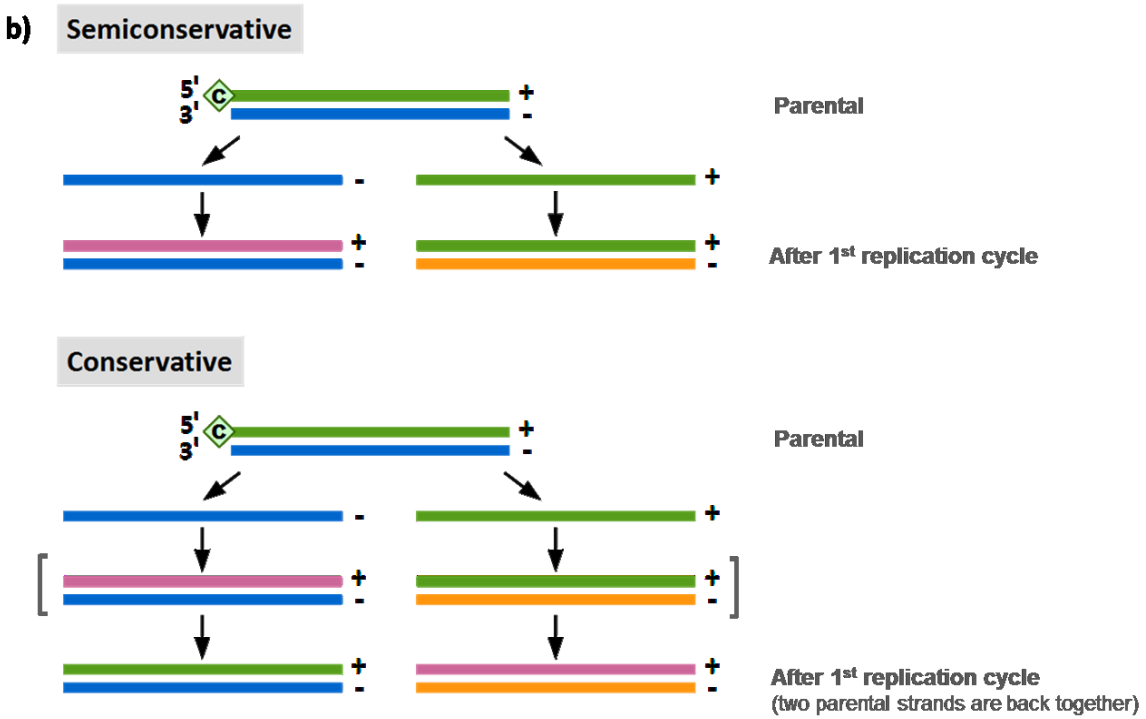


Fig. 1 Replication of dsRNA viruses. Replication cycle of encapsidated dsRNA viruses (a). Classical mode of RNA replication of the dsRNA viruses (b).

1.1.3 Mode of transmission

Typically, mycoviruses are transmitted through asexual spores whereas transmission through sexual spores differs with the fungal hosts. Mycoviruses are different from other viruses that infect plants or animals as they only possess an intracellular life cycle. Thus, they are not considered as an infectious unit. For this reason, mode of transmission of mycoviruses usually

depends on host cell division or *via* intracellular mechanisms. Vertical transmission through the formation of asexual spores (sporogenesis) is a common process for mycovirus spread. Horizontal transmission by fusion of the vegetatively compatible hyphae (anastomosis) is another well-established phenomenon where hyphae from different fungi fuse and exchange genetic and cytoplasmic substances (Nuss, 2011). Therefore, host ranges of the mycoviruses are limited to only a single species or to closely related species.

Recently, the vegetative incompatibility barrier has been overcome using the protoplast fusion technique to introduce mycovirus into different fungal species or strains. For instance, a hypovirulent-associated unencapsidated dsRNA virus (FgV1-DK21) from *Fusarium boothii* was successfully transmitted *via* protoplast fusion to *F. oxysporum* f. sp. *lycopersici*, *F. asiaticum*, *F. graminearum* and *Cryphonectria parasitica* and demonstrated reduction of pigmentation, growth rate and virulence of the recipient fungi (Lee *et al*., 2011). Also, the transfection approach has been successfully used to experimentally extend host range of the partitivirus RnPV1 and mycoreovirus MyRV3 from the white rot fungus *Rosellinia necatrix* in three new host range species- *Diaporthe* sp., *C. parasitica* and *Valsa ceratosperma* (Kanematsu *et al*., 2010). Since the nature of mycovirus replication does not rely on a lytic cycle (an extracellular route for transmission), it is not necessary for mycoviruses to form a rigid virus particle to protect their genomic materials. For this reason, some mycoviruses do not have a particular gene that encodes capsid proteins for particle formation, for example, some members in the families *Narnaviridae* and *Hypoviridae* (Nuss, 2005a and 2005b).

1.1.4 Effect of mycovirus on host phenotype

Mycovirus infection causes diverse effects on the fungal phenotype in both advantageous and disadvantageous ways but the majority of phenotypic alterations are cryptic or symptomless. External (environmental) factors and internal (cytoplasmic) factors can provide negative effects and elicit symptomatic features in the fungal host when certain conditions in the virus and fungi system are imbalanced. The viruses can cause morphological and physiological changes in the fungal hosts, generally resulting in changing the degree of host virulence. Mycoviruses which possess an ability to reduce fungal virulence are designated as “hypovirulent viruses”, whereas mycoviruses which possess an ability to enhance fungal virulence are designated as “hypervirulent viruses”.

A number of virus-mediated symptoms in the fungal hosts caused by the hypovirulent strains are diverse, such as reduction in toxin production, reduction in asexual sporulation, reduction

in pigmentation, decrease in growth rate, reduction of basidiospore germination, change in enzymatic activities and also loss of female fertility (Magliani *et al*., 1997; Dawe and Nuss, 2001). It has been believed that effects of mycovirus infection on host phenotype could result from virus-derived small interfering RNAs (siRNAs) *via* RNA silencing if there is complementarity between host mRNA and virus-derived siRNA, leading to disruption of both host and viral transcripts. Examples of mycoviruses associated with hypovirulence to their hosts mainly belong to the families *Hypoviridae*, *Reoviridae*, *Narnaviridae*, *Megabirnaviridae* and some unassigned viruses. It includes Cryphonectria parasitica hypovirus (CHV-1) (Anagnostakis, 1998; Dawe and Nuss, 2001), mycoreovirus infecting *C. parasitica* (Suzuki *et al*., 2004), Ophistoma novo-ulmi mitovirus (OnuMV) (Hong *et al*., 1999), Sclerotinia sclerotiorum debilitation-associated RNA virus (SsDRV) (Xie *et al*., 2006), S. sclerotiorum hypovirulent-associated DNA virus (SsHADV-1) (Yu *et al*., 2010), Alternaria alternata virus 1 (AaV-1) (Aoki *et al*., 2009), Aspergillus fumigatus partitivirus-1 (AfuPV-1) (Bhatti *et al*., 2011a and 2011b), unclassified Diaporthe ambigua mycovirus (Preisig *et al*., 2000), Botrytis cenerea mycovirus (Castro *et al*., 2003), B. porri RNA virus 1 (BpRV1) (Wu, *et al*., 2012), Fusarium graminearum mycovirus (Chu *et al*., 2002) and edible mushroom Agaricus bisporus mycovirus (Romaine and Schlagnhaufer, 1995). It can be notable that many mycoviruses associated with hypovirulence are phylogenetically close to (+) ssRNA plant viruses but distant from dsRNA viruses such as partitiviruses, totiviruses and chrysoviruses.

Besides the hypovirulent effect of mycoviruses, several studies on mycovirus-fungus interaction revealed that mycoviruses can exhibit some positive effects on the survival of endophytic fungi and their host plants. Marquez *et al*. (2007) reported that endophytic fungus *Curvularia protuberance* harboring Curvularia thermal tolerance mycovirus enhances thermal tolerance to the panic grass *Dichanthelium lanugonosum* which facilitates the survival of plants in an adverse habitat. Other mutual benefits between the fungal host and mycovirus have been reported. For instance, mycovirus associated with killer phenotypes in yeasts (*Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Hanseniaspora uvarum* and *Zygosaccharomyces bailii*) and corn smut fungus (*Ustilago maydis*) can live symbiotically with the their host and produce toxins to compete against sensitive fungi of closely related genera (Koltin, 1988; Tao *et al*., 1990; Schmitt and Neuhausen, 1994; Schmitt and Breinig, 2002).

1.2 Mycoviruses in *Aspergillus* spp. and other human pathogenic fungi

Several mycoviruses have been discovered from fungi in the genus *Aspergillus*. Most of them contain dsRNA molecules varying in dsRNA segment numbers and sizes, genomic composition and particle morphology. Occurrence of the *Aspergillus* mycovirus is rare as reported by Jamal *et al*. (2010) that the presence of dsRNA mycovirus from clinical and environmental *A. fumigatus* isolates was only 6.1%. Most of the *Aspergillus* mycoviruses are commonly found in asexual species and are vertically transmitted through asexual spores. However, horizontal transmission of mycoviruses between different *Aspergillus* species has been described, for instance, transferring of A. niger dsRNA virus to *A. nidulans* through protoplast fusion (van Diepeningen *et al*., 2000). It has been reported that a few *Aspergillus* mycoviruses show an ability to alter colony morphology of the fungal hosts such as Aspergillus fumigatus partitivirus-1 (AfuPV-1) and A. fumigatus chrysovirus (AfuCV) (Bhatti *et al*., 2011a).

To date, information on mycoviruses of human pathogenic fungi is still sparse. In addition to *Aspergillus* sp., some pathogenic fungal genera associated with invasive fungal diseases in humans and animals have been described to harbor mycoviruses such as *Candida* sp., *Penicillium* sp., *Fusarium* sp., *Saccharomyces* sp. and *Cladosporium* sp. These viruses are found to be widely distributed in various fungal phyla and contain different genome structures. A list of characterized mycoviruses found in the *Aspergillus* spp. and other human pathogenic fungi is shown in Table 1.

Table 1. Mycoviruses in the *Aspergillus* spp. and other human fungal pathogens (adapted from van de Sande *et al*. (2010) and ICTVdB (http://www.ncbi.nlm.nih.gov/genome/viruses/).

Virus family	Virus characteristics	Virus species
Chrysoviridae (dsRNA)	30-35 nm, 4 segmented RNA, Encapsidated virus, Non-enveloped virus, Multiple components	- Aspergillus flavus uncharacterized chrysovirus (Wood *et al*., 1974) - Aspergillus fumigatus chrysovirus (AfuCV) (Accession no. FN178512-FN178515) (Jamal *et al*., 2010) - Aspergillus niger virus F (AnV-F); poly adenylated dsRNA) (Buck *et al*., 1973; Hammond *et al*., 2008b) - Penicillium brevicopactus virus - Penicillium chrysogenum virus (Accession no. AF296439-AF296442) - Penicillium cyaneo-fulvum virus

Partitiviridae (dsRNA)	30-40 nm, 2 segmented RNA, Encapsidated virus, Non-enveloped virus, Multiple components	- Aspergillus ochraceous virus (AoV) (Kim and Bozarth, 1985) - Aspergillus ochraceous virus 1 (AoV1) (Accession no. EU118277-EU118279) (Liu *et al*., 2008) - Aspergillus fumigatus partitivirus-1 (AfuPV-1) (Bhatti *et al*., 2011a) - Fusarium poae virus (Accession no. AF15924, AF047013) - Fusarium solani virus 1 (Accession no. D55688-D55689) - Penicillium stoliniferum virus S - Penicillium stoliniferum virus F
Totiviridae (dsRNA)	40-43 nm, Non-segmented RNA, Encapsidated virus, Non-enveloped virus	- Saccharomyces cerevisiae virus L-A (L1) (Accession no. J04692 X13426) - Saccharomyces cerevisiae virus L-BC (La) (Accession no. U01060) - Aspergillus foetidus virus S (AfV-S; 2 segmented dsRNA and 1 satellite RNA) (Banks *et al*., 1970) - Aspergillus niger virus S (AnV-S) (Banks *et al*., 1970) - Aspergillus mycovirus 178 (AsV178) (Accession no. EU289894-EU289895) (Hammond *et al*., 2008b) - Aspergillus mycovirus 341 (AsV341) (Accession no. EU289897) (Hammond *et al*., 2008b) - Aspergillus mycovirus 1816 (AsV1816) (Accession no. EU289896) (Hammond *et al*., 2008b)
Alternaviridae (dsRNA)	4 segmented poly adenylated dsRNA (2.0-3.6 kbp), Encapsidated virus, Non-enveloped virus	- Aspergillus foetidus virus F (AFV-F) (Accession no. HE588144-HE588146, HE647818) (Banks *et al*., 1970; Kozlakidis *et al*., 2013)
Pseudoviridae (ss(+)RNA-RT)	30-50 nm, Non-segmented RNA, Encapsidated virus, Non-enveloped virus	- Candida albicans Tca2 virus (Accession no. AF050215) - Candida albicans Tca5 virus (Accession no. AF065343) - Saccharomyces paradoxus Ty5 virus (Accession no. U19263) - Saccharomyces cerevisiae Ty1 virus (Accession no. M18706) - Saccharomyces cerevisiae Ty2 virus (Accession no. M19542)

		- Saccharomyces cerevisiae Ty4 virus (Accession no. X67284)
Metaviridae (ss(+)RNA-RT)	50 nm, Non-segmented RNA, Encapsidated virus, Non-enveloped virus	- Cladosporium fulvum T-1 virus (Accession no. Z11866) - Fusarium oxysporum Skippy virus (Accession no. L34658) - Saccharomyces cerevisiae Ty3 virus (Accession no. M34549)
Narnaviridae (ss(+)RNA)	Non-segmented RNA, Non-encapsidated virus (Nucleoprotein complex)	- Saccharomyces 20S RNA narnavirus (Accession no. AF039063) - Saccharomyces 23S RNA narnavirus (Accession no. U90136)
Unassigned dsRNA		- Fusarium graminearum virus DK21
Unassigned	Sub-viral RNA encoding killer toxin	- M satellites of Saccharomyces cerevisiae virus L-A

1.3 Self-defense mechanisms in fungi and mycoviruses

1.3.1 Antiviral defence mechanisms in fungi

The thick and mechanically rigid cell walls of the fungal hosts make them impenetratable to mycoviruses. Once entering the cells, mycoviruses disperse within fungal colonies through hyphae of the filamentous fungi. Fungal hyphae are usually compartmentalised by the internal cross-wall structures called septa which are perforated by pores. These pores are large enough to allow free flow of nuclei, ribosomes, mitochondria and also viral genomic materials/virus particles between the cells. It has been reported that self/non-self genetic recognition system called vegetative incompatibility (*vic*) (or heterokaryon) in filamentous fungi plays an important role in controlling natural transmission of mycoviruses between different fungal species or strains (Lee *et al*., 2011; Nuss *et al*., 2011). The *vic* process prevents fusion events and prevents mating of a fungal isolate to another fungal isolate by triggering programmed cell death (PCD), resulting in the limitation of cellular component transfer. Thus, the system reduces the risk of transmission of viruses and transposons among genetically incompartible species, acting as an antiviral defence mechanism at the population level in the fungus.

Several fungi also exhibit the reduction of the mycovirus numbers at the cellular level *via* silencing mechanism (antiviral RNA interference) using fungal enzymes as silencing machinery.

Dicer and argonaute-RISC (RNA-induced silencing) complexes are involved in the process which triggers the degradation of the viral dsRNAs or viral structured RNAs. The resulting virus-derived small interfering RNAs (siRNAs) cleaved by dicer are incorporated into RISC, leaving the guide antisense siRNA strand to target the cognate viral RNAs which are eventually cleaved by argonaute-RISC complexes. Mycovirus-derived siRNAs were found in *Magnaporthe oryzae* infected with an encapsidated dsRNA virus, MoV2, indicating RNA silencing mechanism in the fungal host against mycovirus (Himeno *et al*., 2010). The same phenomenon has also been reported in *Cryphonectria parasitica* (Segers *et al*., 2007) and *Aspergillus nidulans* (Hammond *et al*., 2008a). In contrast, in some unicellular fungi such as yeast, an SKI1/XRN1 5'-exonuclease was found to play an important role in antiviral response by controlling the accumulation of dsRNA totivirus L-A (Wickner, 1996).

Interestingly, defective interfering dsRNA (DI-dsRNA) has been reported to reduce accumulation and replication of its parental virus. For instance, DI-dsRNA1 which is derived from dsRNA1 of the Rosellinia necatrix partitivirus2 (RnPV2) shows an ability to interfere with its cognate dsRNA1 by decreasing replication level of dsRNA1. The event could attribute to the competitiveness of the DI-dsRNA1 against its parental virus genome. In addition, the presence of DI-dsRNA1 in the fungus could alleviate the effect on the fungal host symptoms caused by its parental partitivirus (Chiba *et al*., 2013). In this regard, it indicates that DI-dsRNA could function in viral RNA interference and serve as an antiviral defense response.

1.3.2 Antifungal defence mechanisms in mycoviruses

Mycoviruses have developed a series of defence mechanisms to overcome antiviral defence of the fungal hosts. One example is the production of a protein called RNA silencing suppressor (RSS) to suppress fungal RNA silencing mechanism. Papain-like protease, p29, has been described in hypoviruses of the fungus *Cryphonectria parasitica (*Segers *et al.*, 2006) that this suppressor protein acts as a determinant which modulates sporulation and pigmentation, resulting in the hypovirulent trait of the fungus. It was also found that the protein has an ability to support an accumulation of viral RNAs by interfering with DCL-2 function which is one of the important silencing components in several fungi. Similarly, a mycoreovirus Rosellinia necatrix mycoreovirus 3 (RnMyRV3) has been reported to suppress RNA silencing in *Rosellinia necatrix* by interfering with host Dicer (Yaegashi *et al*., 2013). In addition, some of the mycoviruses from *Aspergillus* spp. were found to be capable of RNA silencing suppression. Hammond *et al*. (2008b) reported that Aspergillus virus 1816 could suppress RNA silencing in *A. nidulans aflR* IRT strain using RNA silencing suppressor encoded by one of the virus1816

elements. Interestingly, it has been described that a totivirus Magnaporthe oryzae virus 2 (MoV2), may possess a novel strategy to suppress host silencing mechanism (Himeno *et al*., 2010). However, information on the process that this virus employs to combat against host silencing is still unclear.

Proliferation of the membraneous lipid vesicles in the fungal cells induced by viruses is also known to be one of the antifungal defence mechanisms. As vesicles were reported to be a favourable site for viral replication, the formation of the vesicles and replication of the virus in the vesicle membranes could provide protection for the virus from host nucleases (Fahima *et al*., 1993). It is evident that numbers of membrane-enclosed vesicles was observed in Cryphonectria parasitica hypovirus 1 (CHV1) infected *C. parasitica* which an accumulation of the vesicles was attributed to virus infection. As CHV1 lacks a coat protein, host vesicles were found to be a replication site for this virus (Kazamierczak *et al*., 2012).

1.4 Application of mycoviruses

1.4.1 Application of mycoviruses in agriculture

Knowledge and understanding of virus-fungus relationship provide opportunities to develop various strategies to control fungal diseases using viruses as a control agent. Since several plant pathogenic fungi have been reported to be resistant to fungicides, including public concern on the environment and human health, mycoviruses have attracted plant pathologists to exploit mycoviruses as a source for novel fungal disease control. As a consequence, several mycoviruses possessing hypovirulent factors have currently been widely investigated. Cryphonectria parasitica hypovirus 1 (CHV1) is one of the best known mycoviruses in biocontrol applications since it has been used successfully as a biological control agent against chestnut blight disease (Anagnostakis *et al*., 1998; Robin and Heiniger, 2001; Milgroom and Cortesi, 2004). However, the exploitation of using mycoviruses to control phytopathogenic fungi is still restricted only in a particular area and small numbers of the pathogen groups due to incompatibility between fungi. It has been interestingly proposed that host range of mycovirus can be extended by introducing plasmids containing full-length mycovirus cDNAs fused with selectable marker genes into spheroplasts by transformation or protoplast fusion. Mycovirus-introduced pathogenic fungi can subsequently be fused compatibly, resulting in migration of mycoviruses into a new isolate. Another possible approach is transfection of synthetic transcripts corresponding to mycovirus coding RNAs by electroporation or by PEG-

mediated transfection of the recombinant virus into fungal spheroplasts. These methods allow replication of the mycoviruses to occur which can subsequently spread through hyphal network of the fungal colony (Nuss, 2011; Lee *et al*., 2011; Kanematsu *et al*., 2010). These proposed strategies can promote the development of exploiting mycoviruses as biological control agents and as tools to elucidate the mechanism of fungal pathogenicity due to reduction of the transmission hurdle caused by host range restriction to mycovirus infection.

In addition to the use of mycovirus as control agents, mycovirus dsRNAs have been exploited as a genetic marker to differentiate the regional populations of *Phytophthora infestans* between the US and Mexican isolates (Tooley and Hewings, 1989). However, this method may not be suitable in the fungi which frequently exhibit sexual reproduction in nature because loss of dsRNA can occur.

1.4.2 Application of mycoviruses in industry

Mycoviruses found in the industrially important yeast *Saccharomyces cerevisiae* encode toxic proteins, generating killer phenotypic yeasts. Generally, fungal cells harboring the killer viruses are immune to their toxic proteins, but other cells from the same or closely related strains lacking viruses are sensitive to toxins. Therefore, the viral killer system shows a strong potential in eukaryotic cell biology as well as virus-fungal cell interaction studies. Besides the beneficial aspect in biological studies, toxin-producing yeasts and killer toxins are also useful in food fermentation and preservation process since they have been shown to eliminate contaminants during bread, beer and wine fermentation (Schmitt and Breinig, 2002). Currently, dsRNA viruses involved in yeast killer toxin production have been used for expression of the heterologous proteins in yeasts (Schrunder and Meinhardt, 1995).

1.4.3 Application of mycoviruses in medicine

In medical aspects, killer strains have been reported in some medically important yeasts such as *Candida*, *Cryptococcus, Kluyveromyces* and *Pichia*. Therapeutic potential of the yeast killer toxins secreted by a dsRNA virus has been described that it exhibits natural antibiotic properties (reviewed by Magliani *et al*., 1997). In addition to *S. cerevisiae* toxin, killer toxins of virus-infected killer strains of the yeasts *Hanseniaspora uvarum* and *Zygosaccharomyces bailii* revealed their potential as an antifungal agents against several human fungal pathogens e.g. *C. albicans*, *Sporothrix schenkii* and *Fusarium* sp. (Schmitt *et al.*, 1997; Hodgson *et al.*, 1995). These have raised a practical idea to apply the killer yeasts/fungi including their toxins and synthetic derivatives as antimycotic agents against the opportunistic fungal pathogens. It is

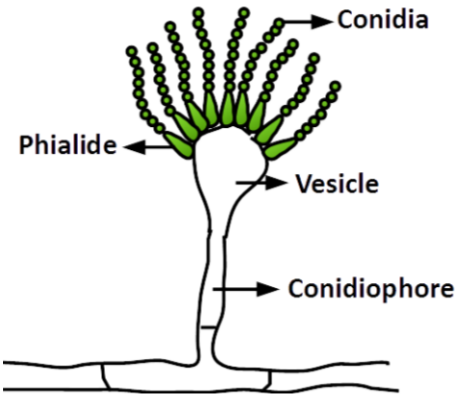
interestingly noted that since some dsRNA mycoviruses cause hypovirulence and can cytoplasmically infect fungi *via* transfection method using cDNA copy of mycovirus RNA, it could be possible to use these viruses as a therapeutic tools and biological control agents for the treatment of human fungal infection by manipulating virus as a gene vector.

---

2. *Aspergillus fumigatus*

2.1 General characteristics and genomic information

Aspergillus fumigatus is a saprophytic and thermotolerant filamentous fungus which produces numerous asexual spores termed conidia. A common habitat of the fungus is soil and decaying organic matter. The optimal temperature for the fungal growth is generally at 37 ˚C, however it can survive in a wide variety of environmental and high temperature conditions up to 50-70 ˚C. *A. fumigatus* cell wall is composed of cross-link polysaccharides (fibrillar branched β1,3-glucan core bound to chitin, galactomannan and β1,3-1,4-glucan embedded in an amorphous cement composed of α1,3-glucan, galactomannan and polygalactosamine) and several glycoproteins (Fontaine *et al*., 2000). Colonies are greyish-green in colour with a downy to powdery texture. Hyphae are septate and hyaline. Vesicle formation on the hyphal apex appears to be of a unique character for this fungal group in that it forms short, smooth, and colourless to greenish conidiophores. The conidiophores contain flask-shaped uniseriate phialides with radial chains of round-shaped conidia on the top (Fig. 2; Campbell *et al*., 1996).



Kingdom	**Fungi**
Phylum	***Ascomycota***
Class	***Ascomycetes***
Order	***Eurotiales***
Family	***Trichocomaceae***
Genus	***Aspergillus***
Species	***Aspergillus fumigatus***

Fig. 2 Illustration of asexual fruiting structure of *Aspergillus fumigatus* and its classification.

The complete genome sequence of *A. fumigatus* Af293 clinical isolate used in this study has been reported to contain a 29.4 megabase genome in eight chromosomes with 49.9% GC content (Nierman *et al.*, 2005). Approximately 9,783 predicted protein coding genes were described, *ca*. 65% of which have been identified and their functions are known. The total number of transfer RNAs (tRNAs) and 5S ribosomal RNAs (rRNAs) in the nuclear genome are 205 and 34, respectively. The number and other types of features annotated in *A. fumigatus* Af293 genome are shown in Table 2.

2.2 Pathogenicity and genes involving fungal virulence

Spores (conidia) of *A. fumigatus* are ubiquitous in the air, thus humans are routinely exposed to several hundred fungal spores following inhalation. Normally this process does not cause any harm to healthy individuals because spores are immediately eliminated from pulmonary areas by the host innate immune system. However, it can become a pathogenic or opportunistic fungus in immunocompromised patients whose immunity is debilitated, especially in bone marrow and organ transplant recipients including some patients with chemo- and radio-therapies for cancer treatment. Exposure to fungal spores can develop several serious illnesses or be fatal to patients as the spores escape immune responses and subsequently invade host tissues and penetrate blood vessels by the formation of filamentous hyphae.

Diseases caused by *A. fumigatus* can be divided into three categories: (A) allergic bronchopulmonary aspergillosis (ABPA) which is an allergic response to inhaled conidia in individuals with a hyperactive immune response, (B) chronic aspergillosis in generally immunocompetent persons with the proliferation of a large ball-like hyphal mass, termed aspergillomas or 'fungal ball' and (C) invasive aspergillosis (IA) in immunocompromised patients (Kavanagh, 2007). Approximately 90% of *Aspergillus* infection is IA which causes 50-95% mortality rate even when treated (Balloy and Chignard, 2009; Abad *et al.*, 2010). Neutrophil numbers and macrophage function is usually reduced in immunocompromised patients with IA, allowing the spores to germinate and proliferate in pulmonary tissues and alveolar walls. Fungal fragments can spread through the circulatory system and destroy other organs and brain, resulting in disseminated aspergillosis. The symptoms of IA are non-specific, but patients commonly show signs of fever, cough (sometimes with blood) and chest discomfort. Therefore, early diagnosis is vital for primary treatment to prevent high mortality.

Table 2. The number and types of features annotated in *A. fumigatus* Af293 genome (adapted from AspGD; Cerqueira *et al*., 2013).

Feature type	Total	Chromosome number									
		Chr# 1	Chr# 2	Chr# 3	Chr# 4	Chr# 5	Chr# 6	Chr# 7	Chr# 8	Nuclear genome	Mitochondrial genome
Total ORFs	9,783	1,635	1,637	1,387	1,243	1,363	1,241	642	619	9,767	16
Verified ORFs	468	61	80	65	69	57	74	22	40	468	0
Uncharacterised ORFs	9,315	1,574	1,557	1,322	1,174	1,306	1,167	620	579	9,299	16
tRNA	205	32	25	26	13	35	17	13	17	178	27
rRNA	34	4	3	9	3	4	6	1	4	34	0
snoRNA	29	11	2	2	7	1	4	2	0	29	0
ncRNA	12	4	1	2	3	0	0	2	0	12	0
snRNA	9	4	1	1	2	0	1	0	0	9	0
Pseudogene	1	0	0	2	0	0	1	0	0	1	0
Total	10,073	1,690	1,669	1,427	1,271	1,403	1,270	660	640	10,030	43
Chr. Length (bp)	**29,420,142**	**4,844,372**	**4,844,372**	**4,079,167**	**3,927,224**	**3,948,441**	**3,778,736**	**2,058,334**	**1,833,124**	**29,388,377**	**31,765**

It has been remarkable that the fungus lacks an ability to sporulate once it invades the human host. The event attributes to nutrient limitation in the host cells, resulting in preference of the fungus to form the vegetative stage rather than asexual stage to maintain its energy. The notion has raised an idea in developing a new therapeutic tool by activating the molecules involved in sporulation and hyphal growth to reduce fungal fitness and virulence.

A series of genes involving in fungal virulence and responsible for the disease development and survival adaptation of the fungus have been described. The fungal determinants or virulence factors (genes and molecules related to virulence) can be categorized according to the processes they participate in e.g. composition and maintenance of cell wall, thermotolerance, resistance to the immune response, toxins, nutrient uptake during invasive growth, signalling, metabolism regulation and response to stress conditions, and allergens (Abad *et al*., 2010). These determinants are likely to contribute to the overall fitness and invasive ability of the fungus. The small size of the spores (2-3 µm in diameter) facilitates the dispersion of asexual airborne spores into the lower respiratory tract and through pulmonary alveoli. Rapid growth rate and hyphal cell wall components also play a key role in the fungal virulence. In addition, various nutrient utility, thermotolerant and oxidative resistant ability are major factors that promote host cell invasion (Abad *et al*., 2010).

2.3 Treatment of invasive aspergillosis

Voriconazole is one of the drugs of choice which is effectively used to treat invasive aspergillosis, especially in immunocompromized patients. Liposomal amphotericin B, itraconazole and caspofungin are also used to treat the IA patients. However, treatment failure remains a major problem due to resistance of the fungus towards current azole antifungal agents. Therefore, new therapeutic strategies and identification of such essential genes involving in fungal virulence would facilitate further development of new antifungal drug targets to fight against the invasive fungal diseases. From several findings mentioned previously, it is very interesting to note that some A. fumigatus dsRNA mycoviruses which show deleterious effects of infection to the fungal host, have a great potential to be used as an alternative control agent to control *A. fumigatus* infection.

---

3. Transposable elements (TEs) in fungi

Introns and intergenic non-coding regions in the eukaryotic genes usually contain repetitive DNA sequences such as transposable elements (TEs). TEs (sometimes called transposons) are mobile, repetitive genetic elements which were first identified in maize in the 1940s by Barbara McClintock (Ravindran, 2012). In fungi TEs were first discovered in the yeast *Saccharomyces cerevisiae* (Boeke, 1989) and have been later found in various groups of unicellular and filamentous fungi. The elements are generally responsible for genome re-arrangement and evolution by duplication, insertion (into exons, regulatory regions and introns) and deletion owing to their repetitive nature. For instance the *Phytophthora infestans* genome consists of a large proportion of TEs and repeats which contributes to the increased size of the *P. infestans* genome (Whisson *et al*., 2005a). Interestingly, TEs of the ascomycetes *Aspergillus nidulans* and *Neurospora crassa* were found to be silenced to prevent invasion of the elements into the genome (Cogoni, 2001).

In *Aspergillus* species, TEs are often seen within secondary metabolite clusters which locate in the subtelomeric regions of the chromosomes (Fedorova *et al*., 2008). *A. fumigatus* TEs are believed to be associated with genetic diversity and regulation/change in structure of the secondary metabolite clusters, resulting in adaptations of the fungus to challenging harsh environment, especially during infection in the infected host (Perrin *et al*., 2007; Fedorova *et al*., 2008). TEs have been reported to play an important role in adaptive gene regulation in fungi. For instance, up-regulation of the TEs in *A. fumigatus* was observed during the early stages of lung infection where the fungal spores start the germination process (Lamarre *et al*., 2008). In *A. nidulans*, penicillin synthesis decreased by up to 75% when the TE-rich sequences flanking the penicillin cluster was deleted, signifying that TEs may serve as boundary elements or affect heterochromatin organization in regulatory process in penicillin synthesis (Shaaban *et al*., 2010).

TEs can be divided into two major groups according to their mode of transposition and structural organization. **Class I retrotransposable elements** transpose by reverse transcription of an RNA intermediate which are subdivided into different sub-classes based on their size, organization and structure (see section 3.1) while **Class II DNA transposable elements** transpose directly by the enzyme transposase (Finnegan, 1989).

3.1 Retrotransposable elements (Retrotransposons)

Retrotransposons are mobile genetics elements which are the most abundant class of TEs. They can amplify themselves in a host genome. The propagation of the elements occurs through the process called retrotransposition *via* RNA intermediates. Mode of replication of the elements has been recently elucidated that the retrotransposons are transcribed to RNA transcripts, then the transcripts are subsequently reverse-transcribed by reverse transcriptase (RT) encoded by a retrotransposon. The resulting cDNAs are re-integrated at diverse sites in the host genome resulting in an increased copy number of the elements and also an increase of the genome size. Generally, re-integration of retrotransposons can induce mutations if the insertion takes place within essential genes or nearby and also contributes to genome evolution.

Retrotransposons consist of two sub-types, the long terminal repeat (LTR) and non-long terminal repeat (non-LTR) retrotransposons.

3.1.1 LTR retrotransposons

LTR retrotransposons are 4 - 10 kbp long nucleotide sequences resembling retrovirus genomes and comprise an internal domain flanked by two direct repeats. They can be divided into several sub-groups e.g. *Ty1-copia*-like, *Ty3-gypsy*-like and *Pao-BEL*-like which are classified based on the level of sequence similarity, encoded gene products and order of the coding regions of structural proteins (*gag* – viral coat protein) and enzymatic proteins (*pol* – reverse transcriptase, RNase H, integrase, and protease). For example, in the *copia*-like element, the order of the *pol* genes is protease/integrase/reverse transcriptase/RnaseH while *gypsy*-like element the order of the *pol* genes is protease/reverse transcriptase/RnaseH/integrase (Fig. 3a) (Lima-Favaro *et al*., 2005). Both *Copia* and *gypsy*-like elements are widely distributed among eukaryotes and can be found in high copy numbers of up to few million copies in animals, plants, protists and fungi.

3.1.2 Non-LTR retrotransposons

Non-LTR retrotransposons consist of two sub-types, long interspersed nuclear elements (*LINEs*) and short interspersed nuclear elements (*SINEs*). They are ubiquitous components of many eukaryotic organisms (although not found in *S. cerevisiae* and *Schizosaccharomyces pombe*; Wood *et al*., 2002; Novikova *et al*., 2009) and are widespread in eukaryotic genomes in high copy numbers.

3.1.2.1 Long interspersed nuclear elements (*LINEs*)

LINEs are referred to as autonomous retrotransposons. They are present as a 3-7 kb in length and widely distributed in the genome. The composition of a functional full-length *LINE* consists of 5' untranslated region (5'-UTR) containing an RNA polymerase II internal promoter sequence and two non-overlapping open reading frames (ORF1 and ORF2) (Fig. 3b). ORF1 encodes an RNA-binding protein associated with stabilization of *LINE* RNA and enhancement of strand transfer during the reverse transcription process. ORF2 encodes two enzymes, endonuclease (EN) and reverse transcriptase (RT), which are necessary for retrotransposition. The 3'-UTR of the element contains a polyadenylation signal (AATAAA), a poly (A) tail as a terminal sequence, and short target site duplications of variable length of 9-20 bp on both ends (Ostertag and Kazazian, 2001).

The retrotransposition mechanism of *LINEs* has been described where elements are transcribed *via* RNA polymerase II and amplified *via* reverse transcription of their RNA intermediate. It has been reported that the N-terminal endonuclease domain of ORF2 protein has an ability to nick a double-stranded DNA target containing T_2A_4 signal. This excision creates a free 3'-OH single-stranded DNA which is used as a primer for first-strand cDNA synthesis through its RT (Cost *et al*., 2002). The DNA copies are then re-integrated into a new site on the genome. It has been described that inactivation of *LINEs* can be induced by a spontaneous mutation, internal rearrangement, and 5'-terminal truncations on the sequence (Soifer *et al*., 2005).

Seventeen clades of *LINE* families have been described to date (Kramerov and Vassetzky, 2011a and 2011b). The best characterized *LINE* is a *LINE-1* (*L1*) which can be found in the genome of all mammalian species. It is *ca*. 6 kbp in length and exhibits a well-organized structure of a typical *LINE* (Ostertag and Kazazian, 2001). The *Tad1-1* element of the fungus *N. crassa* was the first *LINE* identified in fungi. It is 6.9 kbp in length with 14-17 bp of target site duplication sequences. The element is found to be active showing approximately 40 copies and is inserted in the glutamate dehydrogenase *am* gene (Kinsey and Helber, 1989). In addition, several fungal *LINE* and *LINE*-like elements were characterized such as *MGR583* from *Magnaporthe grisea* (Valent and Chumley, 1991), *mgl* from *M. grisea* (Nishimura *et al*., 2000), *MARS1* from *Ascobolus immersus* (Goyon *et al*., 1996), *marY2* from *Tricholoma matsutake* (Murata *et al*., 2001), *CgT1* from *Colletotrichum gloesporioides* (He *et al*., 1996) and *I LINE* superfamily from *A. nidulans*, *A. fumigatus* and *A. oryzae* (Clutterbuck *et al*., 2007; Huber and Bignell, 2014). However, numbers of *LINE* elements are limited in the unicellular eukaryotes.

3.1.2.2 Short interspersed nuclear elements (*SINEs*)

SINEs are short repetitive, non-coding sequences which are 100 to 600 bp in length. However, the majority of *SINEs* (*c.a.* 66%) are 150 to 300 bp long. They are similar to *LINEs* but simpler and shorter whereas some of them possess complex *SINE* structures. Generally *SINEs* are established into an individual *SINE* family if their copies are not identical. The similarity of *SINEs* can be varied by 5-35% depending on the time the copies emerged. Commonly only one, two or several *SINE* families (often 2-4) are present in a particular genome. The abundance of *SINEs* in a particular genome usually ranges from ten to more than a million copies in each family. The elements are found to be randomly dispersed in intergenic regions - normally within gene-rich, GC-rich regions, while *LINEs* are accumulated in the gene-poor, AT-rich regions (Lander *et al*., 2001).

SINEs are referred to as non-autonomous retrotransposons because they do not possess ORFs to produce proteins necessary for retrotransposition. Thus, they usually depend on the enzymes encoded by their partner *LINEs* (EN/RT) for transcription and transposition. The amplification of *SINEs* occurs in the similar manner to *LINEs* but *SINE* transcripts are transcribed by RNA polymerase III. Thus, the pol III-*SINE* transcripts are capped and are not used as substrates for ribosomes in the translation process. Transcripts are then reverse transcribed through RT produced from the partner *LINE* into cDNAs and subsequently re-integrated into various sites in the genome. Some *SINE* copies which possess an intact RNA pol III promoter are functionally active. However the majority of *SINEs* are not actively transcribed.

It has been observed in several cases that 3' terminal regions of some tRNA-derived *SINEs* showed sequence and structural similarities to the corresponding *LINEs* found in the same genome (30-100 bp region of similarity) which is required for the recognition by *LINE* RT. This proves that retrotransposition of *SINEs* requires enzymatic sources from their partner *LINEs* (Fantaccione *et al*., 2008). Examples of *SINEs* which show significant homology at the 3-termini to the corresponding *LINEs* are for example *LINE2* and *MIR SINE* in mammals (Smit, 1996), *CR1 LINE* and *polIII/SINE* in tortoise (Ohshima *et al*., 1996) and *UnaL2 LINE* and *UnaSINE1* in eel (Kajikawa and Okada, 2002). However, the human *Alu SINEs* do not conform to this criterion as the sequences do not share any similarity at the 3'-terminus to the active *L1* partner *LINE* elements (Dewannieux and Heidmann, 2005).

1) Structure and classification of *SINE* elements

Multiple *SINE* structures and transposition mechanisms have been observed, resulting in the description of different *SINE* families. Main structure of *SINEs* consists of head, body and tail regions.

a) 'Head' at 5'-terminus is commonly similar to one of the three types of RNA (tRNA, 7SL RNA, or 5S rRNA), which are synthesized by RNA pol III.

b) 'Body' consists of the central domain of unknown origin.

c) 'Tail' at 3'-terminus usually consists of the repeated mono-, di-, tri-, tetra- or pentanucleotides of the poly(A) or A-rich sequence.

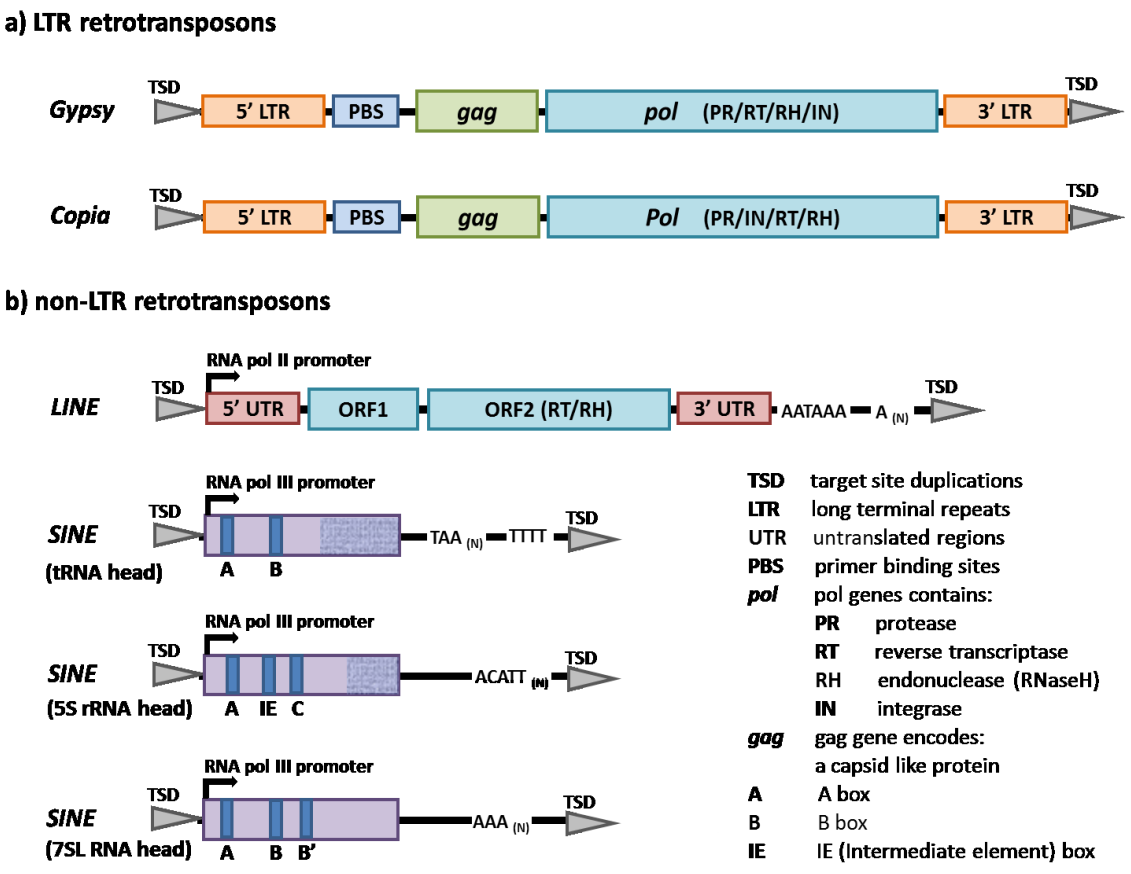


Fig. 3 Structural features of retrotransposons

SINEs are commonly found in monomeric structures. However, complex *SINE* structures, such as in a dimeric or trimeric forms, have been investigated. The *Alu* family in primates is a dimeric *SINE* composing of two 7SL RNA-derived *SINEs* from the same *SINE* family. This type of complex *SINE* is called homodimer. The complex *SINE* structures which are formed by two or

more different *SINE* families have also been described such as a dimeric *SINE* composed of tRNA/7SL RNA or tRNA/5S rRNA-derived *SINEs*. This type of complex *SINE* is called heterodimer (Kramerov and Vassetzky, 2011).

tRNA-related *SINEs*

A high abundance of *SINE* elements in eukaryotic species are derived from tRNAs while some *SINEs* are found to be derived from 5S rRNAs or 7SL RNA. The tRNA-related *SINEs* (or *SINE2* class) have an internal promoter (type II promoter) recognized by RNA polymerase III (pol III), therefore they are transcribed *via* this RNA polymerase type. The pol III promoter contains two boxes of *ca*. 11-bp A box and *ca.* 10-bp B box in which are spaced by a 30-45 bp sequence, resulting in an internal control region (ICR) running from +10 to +65 (Borodulina and Kramerov, 1999; Kramerov and Vassetzky, 2011a and 2011b) (Fig. 3b).

Since most *SINEs* are derived from tRNA, their 5'- upstream region show significant similarity to the sequence of tRNAs (commonly lysine and arginine). However, a tRNA origin of some *SINEs* cannot be identified because of nucleotide substitution during evolution of the elements. The 3'-downstream region of *SINEs* is composed of a 3'-tRNA unrelated sequence of variable length (usually 100-200 bp) which is very diverse between families followed by a short stretch of A or T residues at their 3'-ends. Polythymidine (poly T) tracts downstream act as pol III termination signals in functional *SINEs* while polyadenine (poly A) tracts varying from 8-50 bp function as templates for the initial step of reverse transcription. The presence of flanking direct repeats of short target site duplication (TSD) for retrotransposition is one of the unique characteristics of *SINEs*. Generally, transcripts of most *SINEs* cannot fold into a tRNA cloverleaf structure (e.g. *Fusarium oxysporum Foxy*; Fig. 4a), however a few *SINEs* have been reported to form tRNA-like cloverleaf secondary structures such as *MgSINE* (Fig. 4b) and *MIR* (Smit and Riggs, 1995; Kachroo *et al*., 1995).

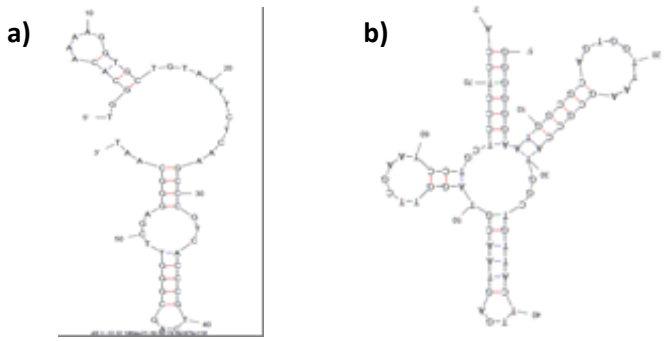


Fig. 4 Secondary structure of tRNA-related region of *Foxy* (a) and *MgSINE* (b).

7SL RNA-related *SINEs*

The 7SL RNA-related *SINEs* (or *SINE2* class) have an internal promoter (type II promoter) recognized by RNA pol III the same as observed in the tRNA-derived *SINEs*. However, the internal promoter of 7SL RNA-derived *SINEs* contains an additional B' box after the typical B box (Fig. 3b). The number of *SINEs* derived from 7SL RNA is low. To date, the 7SL RNA-related *SINEs* are dominant in primates such as the *Alu* in human (Ullu and Tschudi, 1984). However some have been identified in several rodent and three shrew genomes such as *B1* families (Nishihara *et al*., 2002; Kapitonov and Jurka, 2003).

5S rRNA-related *SINEs*

SINEs derived from 5S rRNA are classified into the *SINE3* class which is transcribed from a type I internal RNA polymerase III promoter. Type I promoters are comprised of *ca.* 15-bp A box, *ca.* 5-bp intermediate element (IE), and *ca.* 18-bp C box, which form *ca.* 50-bp ICR running between 150 and 190 (Fig. 3b). Unlike tRNA-related *SINEs*, the sequence of 5S rRNA-derived *SINEs* are not bordered by target site duplications and their 3'-termini contain (ACATT)n and (ATT)n microsatellites (Kapitonov and Jurka, 2003).

To date, 5S rRNA-derived *SINEs* are reported in the genome of *A. oryzae* (Galagan *et al*., 2005; Kapitonov and Jurka, 2006), zebrafish (Kapitonov and Jurka, 2006) and a few mammalian species such as springhare (Gogolevsky *et al*., 2008) and fruit bats (Gogolevsky *et al*., 2009).

2) Origin and evolution of *SINEs*

Traceability for the origin of *SINEs* is difficult. It has been elucidated that most *SINEs* are an integration product of a tRNA/7SL RNA/5S rRNA-related gene and an unrelated sequence. In tRNA-derived *SINEs*, a new *SINE* transcript emerges when a tRNA gene containing RNA pol III promoter integrates close to the terminus of a *LINE*. The integration event generates a 3'-end recognition site for the reverse transcriptase of a *LINE*, leading to the amplification of new *SINE* elements (Smit and Riggs, 1995; Smit, 1996). The origin of *SINE3* has also been described by Kapitonov and Jurka (2003). It has been proposed that the 5'-terminus of *SINE3* identified from the zebrafish genome was derived from a 5S rRNA gene or pseudogene while the 3'-teminus was derived from a *CR1*-like non-LTR retrotransposon which is required for retrotransposition of *SINE3*. In addition, it has been found that the termini of the elements are spaced by a short fragment of the host genome.

Evolutionarily, new *SINE* families/subfamilies emerge due to modifications in their structures such as point mutation. The activity of these elements is also a crucial factor for the emergence of new *SINE* families. Subsequently, some *SINEs* lost their function and activity with time by degradation of the copies, leading to extinction of the elements (Kramerov and Vassetzky, 2011a and 2011b).

3) *SINE* elements in eukaryotes

Several *SINE* families have been found in many eukaryotic species such as mammals, reptiles, fish, mollusks, ascidia, insects, flowering plants and fungi. Approximately 120 *SINE* families have been reported to date (Kramerov and Vassetzky, 2011b). The first described *SINEs* is the *Alu SINE* family which is present in the human genome consisting of $>10^6$ copies and covering more than 10% of the genome (International Human Genome Sequencing Consortium, 2001). The insertion of new human *Alu* element takes place *ca.* every 20 human births where the elements are found to be inserted on every 3 kb within the human genome. Besides *Alu, MIR* is also the most common *SINE* found in primates. The element is a tRNA-derive *SINE* of 260 bp in length. They are present in a high copy numbers of *ca.* 300,000 copies and can be found in all mammalian genomes (Smit and Riggs, 1995).

Novel *SINE* families have been described from the genome of bats such as *VES* ($tRNA^{arg}$ -derived *SINE*) from *Myotis daubentoni* (Borodulina and Kramerov, 1999), *B1* and *B2* families from rodent genomes (Nishihara *et al*., 2002; Kapitonov and Jurka, 2003), *Rhin-1* and *Das-1* from the great horseshoe bats *Rhinolophus ferrumequinum* (Borodulina and Kramerov, 2005) and *P.k.SINE* ($tRNA^{arg}$-derived *SINE*) from *Pipistrellus kuhli* (Fantaccione *et al*., 2008). However, it is still obscure whether the presence of *SINEs* can be investigated in the genomes of most eukaryotic organisms. Eickbush and Furano (2002) suggested that some eukaryotes do not have *SINEs* in their genomes such as yeasts, *Drosophila* species, or in the nematode *Caenorhabditis elegans*.

4) *SINE* elements in fungi

To date, several *SINEs* have been reported in fungi. *MgSINE* isolated from the rice blast fungus, *Magnaporthe grisea*, is a 472 bp tRNA-derived *SINE* representing approximately 100 copy numbers in the genome and possesses similar features to mammalian *SINEs*. The element was isolated from an inverted-repeat transposon called *Pot2* and was found as an insertion element within the transposon (Kachroo *et al*., 1995). Another *SINE* family identified from the

M. grisea was the *mgsr1 SINE* which is an 800 bp element representing *ca.* 40 copies in the genome (Sone *et al*., 1993).

A few *SINEs* have been described in the powdery mildew fungus *Erysiphe graminis*. *Egr1* is a 700 bp long *SINE* element which is present *ca.*50 copies in the genome (Wei *et al*., 1996). Another *SINE* family found in this fungus is *egh1 (EGH24)* which is a 900 bp *SINE* (Rasmussen *et al*., 1993). *Foxy* is an active *SINE* family found in *Fusarium oxysporum* f. sp. *lycopersici* strain Fo1007 with approximately 160 copies in the genome (Mes *et al*., 2000). The element exhibits a unique feature of terminal tetranucleotide repeats present at the 5' end and found to be active after exposure to gamma radiation resulting in the insertion of the element to a new site in the genome. In addition, various *SINEs* have been identified in other filamentous fungi such as *nrs1* (500 bp, 11 copies) from *Nectria haematococca* (Kim *et al*., 1995), *SINE2-1_BG* (tRNA-derived *SINE*) from barley powdery mildew *Blumeria graminis* (Spanu *et al*., 2010; Bao and Jurka, 2011) and a family of *infSINEs* from *Phytophthora infestans* (Whisson *et al*., 2005a).

A family of *SINE3-1_AO* (5S rRNA-derived *SINE*-like retrotransposons) was found in *A. oryzae* genome (Galagan *et al*., 2005; Kapitonov and Jurka, 2006). The element is 206 bp in length and shows a 63-bp 5' termini derived from 5S rRNA (position 3-65) with 98% identity to the *A. oryzae* 5S rRNA.

5) Roles of *SINE* elements in cellular and molecular levels

SINEs are considered as selfish DNA and genome parasites since they exist in very high copies and possess the ability to propagate in order to secure their existence in the genome. Thus, *SINEs* have a crucial role in genome shaping, genome re-arrangement and modulation of gene expression. *SINEs* may be involved in cell survival during physiological stress conditions such as heat shock, DNA damage, irradiation, oxidative stress, toxic agents, low temperature and infection by pathogens or protoplast isolation (Bradshaw and McEntee, 1989; Paquin and Williams, 1988; Wessler, 1996; Grandbastien, 1998; Capy *et al*., 2000). These abiotic and biotic stresses have been identified to activate the abundance and transcription activity of the elements.

As the sequences of some *SINEs* resemble the functional signals of the genes, *SINEs* can act as enhancers or silencers to modulate gene expression. The insertion inactivation into coding or regulatory regions of genes or introns can cause RNA splicing abnormality event and disruption of the genes (Brosius, 1999; Deininger *et al*., 2003). In addition, they have been demonstrated to play an important role in translational inhibition and protein structure modification which

may result in genetic disorder (Wichman *et al*., 1992; Amariglio and Rechavi, 1993; Schmid, 1996; Britten, 1996). However, an impact of integration events by *SINEs* is limited because the integration site is usually far from the functional regions. Interestingly, the inactivation of *SINEs* by silencing mechanisms triggered by the host genome has been investigated. The silencing event was believed to be a consequence of mutagenic activity of the elements which can alter/block gene expression and transcription pattern due to insertion into important genes.

Some active *SINEs*, such as the *Alu* elements, involve in the disease-causing *de novo* insertions because they are described as a human endogenous mutagen (Deininger and Batzer, 2002). In addition, an abundance of *SINE* transcripts has been observed in many malignant tumors and has been reported to be correlated to DNA methylation (Feber *et al*., 2011; Gualtieri *et al*., 2013). Akiyama and the colleagues (2007) investigated the alteration of *Foxy* transposition and transcription activity in the *F. oxysporum* isolates treated with 5-azacytidine (5azaC) a chemical that causes a decrease in DNA methylation. It has been found that the transposition frequency was enhanced in the 5azaC-treated isolates compared to the non-treated ones. In contrast, the alteration in transcription level was not observed on both treatments.

3.2 Retrotransposons in *A. fumigatus*

Various classes of retrotransposons in the genome of *A. fumigatus* Af293 have been described. Repetitive DNA families (e.g. *Mariner*) and LTR-retrotransposable elements including two *Copia* and five *Gypsy* families have been reported in the fungus (Goldman and Osmani, 2007). *Gypsy* type retrotransposons have been identified such as *Afut1_AF* and *Afut2_AF* which are ancient retrotransposon-like elements consisting of RT, RNase H and EN pseudogenes (Neuveglise *et al*., 1996). The *Afut1_AF* is 6914 bp in size (>10 copies), bounded by two long terminal direct repeats of 282 bp, and flanked by 5 bp (TCCTT) sequence on both ends. The element is characterized as a defective retrotransposon because mutations in the conserved motifs in the RT domain as well as the interruption by many stop codons in the RT amino acid sequence have been observed (Neuveglise *et al*., 1996).

Numbers of non-LTR retrotransposons have been described in the *A. fumigatus* Af293 genome (Novikova *et al*., 2009). Non-LTR retrotransposon, *I-1_AF* which belongs to the *Tad* clade of *LINE*-like elements (LLEs) has been found in several copies. The elements encodes DNA/RNA-binding protein, EN, RT and RNase H and insert randomly in the genome or precisely at the

same target site in *Afut2_AF* (Galagan *et al*., 2005; Kapitonov and Jurka, 2006). Recently, a full description of *LINE*-like elements (*ca*. 24 copies) in the *A. fumigatus* genome of several clinical and environmental isolates has been described to belong to the *Tad1* clade of retrotransposons and have sequence similarity to the previously characterized *I-1_AF LINE* (Huber and Bignell, 2014). Among the 14 clinical isolates tested, one isolate was found not to contain any *LINEs*, suggesting that *LINE* activity is not required for infection in humans. In addition, distribution, expression and degree of insertional heterogeneity of LLEs in *A. fumigatus* varies among the isolates tested. For example, LLEs in the *A. fumigatus* Af293 are actively transcribed but not in the CEA17 isolate. However, the existence of *SINE* elements in the *A. fumigatus* genome is as yet not recorded.

3.3 Application of *SINEs*

3.3.1 Phylogenetic marker and genomic tag

Single copies and entire *SINE* families are useful in phylogenetic studies as they can be used as an alternative phylogenetic marker and also used for genotyping the middle rank of the taxa e.g. orders and families (Kramerov *et al*., 1999). Distinct features of *SINEs* were described to confirm the reliability of employing the elements as phylogenetic markers as proposed by Serdobova and Kramerov (1997). Firstly, it is unlikely that the same *SINE* family in the genome of two different species can be formed. Secondly, *SINE* families are transmitted vertically, inherited from parent to the offspring and can persist for a very long time (more than 10^7 years, approximately) in the genomes. Thirdly, since no identical tRNA-derived *SINE* is observed in two unrelated taxa proving that the element cannot be transferred horizontally among species.

Deragon and Zhang (2006) exploited these intriguing properties of the elements using *SINEs* from the young SB1 family (*SINE* families from the flowering plant family *Brassicaceae*) as a molecular marker to study the phylogenetic relations among wild species of the *Brassica oleracea* cytodeme and cultivated *B. oleracea* species. In addition, the molecular marker from *SINEs* can also be used to monitor gene flow among closely related species in plants. The application of *SINEs* as molecular markers was also examined in mammal species using the data on the species distribution of the *B2 SINE* family to investigate phylogenetic information among rodents (Serdobova and Kramerov, 1997). In fungi, the *Foxy* element found in *F.*

oxysporum could be used as a phylogenetic marker since the element is present among the *Fusarium* tested in the study (Mes *et al*., 2000).

In addition, *SINEs* can play an important role in the agricultural industry since PCR amplification of *SINEs* has been applied for species-specific detection and quantitation of DNA in bovine-, porcine-, chicken- and ruminant-derived products (Walker *et al*., 2003). Since *SINE* families in the different genomes arise independently and randomly dispersed in a large number, the elements can be exploited as a specific marker/ genomic tag to identify DNA from the species in different order from mixed DNA materials.

3.3.2 Regulation of gene expression

Several studies on fungal gene expression, which is regulated by retrotransposons, have been reported (Anbar *et al*., 2005, Whisson *et al*., 2005a, Lerat and Semon, 2007) and demonstrated that retrotransposon transcriptional activation can modify the expression of an adjacent sequence. A study on *P. infestans* retrotransposons revealed the accumulation of small interfering (si)RNAs corresponding to the sequence of *SINEs*, suggesting that the activity of *SINEs* may be controlled by RNA silencing (Whisson *et al*., 2005a, Vetukuri *et al*., 2011a). From these findings, *SINE* elements were then exploited to develop a silencing vector by transcriptional fusion to a *PiAvr3a* gene encoding an RXLR avirulence effector. The investigation on the transcript levels in transgenic *P. infestans* lines showed that the expression of the RXLR effector could be silenced through the spread of silencing from the adjacent *SINE* (Vetukuri *et al*., 2011a). Therefore it could be possible to exploit *SINEs* to develop a novel tool for gene silencing in *A. fumigatus* and other *Aspergillus* species. In addition, another application of *SINE* might involve exploiting the element for insertional mutagenesis analysis and gene knock-out.

---

4. RNA silencing in fungi

4.1 Cellular mechanism of RNA silencing

RNA silencing is an RNA dependent process of sequence-specific gene regulation which can modulate gene activities in living cells, commonly in eukaryotes. This phenomenon has been described as RNA interference (RNAi) in animals, post-transcriptional gene silencing in plants

and quelling in fungi and algae. The process takes place in the cytoplasm. Protective and defensive roles of RNA silencing have been established and clearly described that it plays a crucial role in gene expression and regulation, including defence against parasite elements such as viruses (exogenous threats) and transposons (endogenous threats).

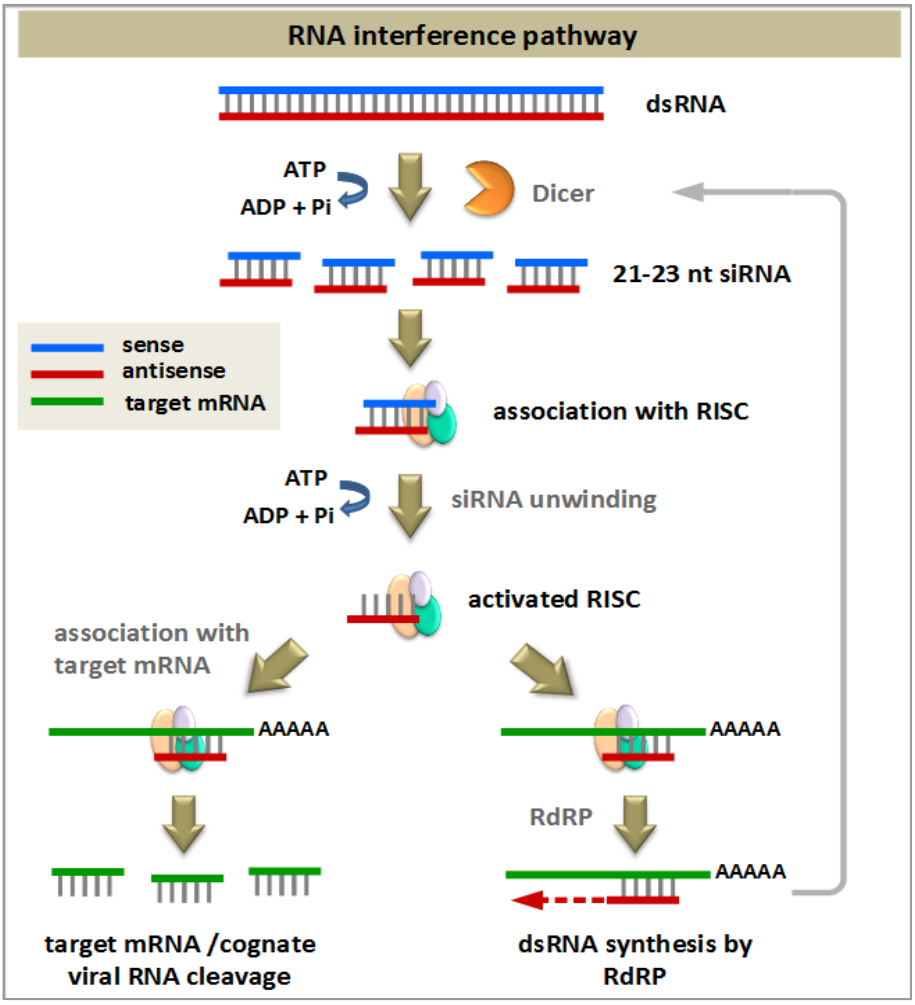


Fig. 5 Schematic diagram of RNA silencing pathway.

Core components of RNA silencing consist of three main silencing proteins; 1) Dicer protein families, 2) Argonaute protein families which are ribonucleases; and 3) RNA-dependent RNA polymerase (RDRP) (Hammond, 2005). Multi-domain proteins of the dicer family are composed of two RNase III domains, a dsRNA binding domain, a PAZ (Piwi-Argonaute-Zwille) domain, a DEAD box helicase domain, a helicase domain and a domain of unknown function. This protein is responsible for processing long dsRNAs into small RNAs of 21-23 bp fragments. These small RNAs are classified based on their dsRNA origins e.g. small interfering RNA (siRNA), microRNA (miRNA) and repeat-associated small interfering RNA (rasiRNA). Argonaute proteins exhibit RNase H-like activity. They are essential elements in the RNA-inducing silencing

complex (RISC) involving in mRNA cleavage. The argonaute proteins are composed of a central PAZ domain possessing small RNA binding activity and C-terminal PIWI domain containing slicer activity which degrades mRNAs. RdRPs are believed to participate in either dsRNA formation for Dicer processing or participate in small RNA formation for RISC complexes in the silencing process

The mechanism of RNA silencing (Fig. 5) is initially induced by the presence of >200 nt long endogenous (e.g. pre-microRNA expressed from an RNA coding gene that forms hairpin structure) or exogenous (virus infection or laboratory introduced) dsRNAs homologous to a gene targeted for silencing. The dsRNA molecules are recognized by the enzyme called dicer nucleases in an ATP-requiring step and then cleaved into small RNAs in the host cytoplasm. This process generates short fragments of 21-23 nucleotides (nt) in length, termed small interfering RNAs (siRNAs), with 2-3 nt overhangs at the 3'-end. The siRNAs are subsequently incorporated into the effector complex, RISC, with co-operation of argonaute, and then are unwound. The sense strand of siRNA is degraded during RISC activation, leaving the anti-sense strand or guide strand to guide RISC to bind the complementary mRNA or cognate viral RNA. The duplex is then endonucleolytically cleaved by argonaute, leading to the inhibition of protein translation or viral replication (Ding, 2010; Nuss, 2011).

In some organisms, RNA silencing is maintained by the reuse of siRNAs which are cleaved from long dsRNAs. This mechanism is referred to as transitive RNAi which involves incorporation of an RdRP. In this process, antisense siRNA complementary to the cellular mRNA is used as a primer for RdRP to re-amplify dsRNA interference from the targeting mRNA (Melnyk *et al*., 2011). Newly synthesized dsRNAs then serve as substrates for dicer to produce secondary siRNAs, leading to persistence of RNA silencing. The presence of RdRP required for RNA silencing has been described in *Caenorhabditis elegans*, plants and some fungal species e.g. *Neurospora*, but not in humans and *Drosophila* (Zamore, 2001).

4.2 RNA silencing machinery in fungi

RNA silencing in fungi was first studied in *Neurospora crassa* which showed the silencing components of two dicer-like proteins (DCL-1 and DCL-2 which *dlc-2* encodes endonuclease that cleaves dsRNAs into siRNAs), two Argonaute-like proteins (QDE-2 and SMS-2), RdRPs (QDE-1 and SAD-1) and RecQ DNA helicase (QDE3) (Cogoni and Macino, 1999). *Cryphonectria parasitica* and *Phytopthora infestans* also encodes the same families of RNA silencing core

components as *N. crassa* (Vetukuri *et al*., 2011a). However, RNA silencing pathways and silencing components seem to be complex and diverse among fungal species or even closely related species. For example, dicer proteins and RdRPs are not found in the opportunistic human pathogenic fungi, *Candida albicans* and *C. tropicalis* (Nakayashiki, 2005b) whereas some fungal species lost both dicer and argonaute RNA silencing machinery e.g. in budding yeast *Saccharomyces cerevisiae* and the basidiomycete *Ustilago maydis* (Aravind *et al*., 2000; Kamper *et al*., 2006).

An RNA silencing mechanism has been described in several *Aspergillus* species. The human pathogen *Aspergillus fumigatus* was first studied for RNA silencing by targeting two endogenous genes *alb1* and *fks1* using inverted repeat transgene construct (IRT) which resulted in range of silenced phenotypes (Mouyna *et al*., 2004). Subsequent similar investigations have been conducted in the mycotoxin-producing fungi, *A. flavus* and *A. paraciticus* which demonstrated silencing activity in fungi by suppression of aflatoxin biosynthesis pathway (McDonald *et al*., 2005). These experimental studies led to the investigation on RNA silencing components in six *Aspergillus* sp. including *A. oryzae*, *A. nidulans*, *A. fumigatus*, *A. flavus*, *A. clavatus* and *A. terreus* including *N. fischeri*. The analysis of Dicers, Argonautes and RdRPs on these fungal genomes revealed complexity of the RNA silencing gene evolution and silencing component. Two dicers (*dclA* and *dclB*) and two argonautes (*ppdA*/*rsdA* and *ppdB*), which are orthologous to *N. crassa,* have been identified in the majority of *Aspergillus* species. However, an additional dicer and argonaute (*dclC* and *ppdC*) have been identified in *A. oryzae* and *A. flavus* where duplication of these genes is likely to attributable to an evolutionary event. In contrast, a dicer (*dclB*) and argonaute (*ppdA*) are absent from *A. nidulans*. The analysis of RdRP genes revealed that three general classes of RdRPs (*RrpA, RrpB* and *RrpC*) are present in the Aspergilli. However *RrpC* is absent from *A. fumigatus* and its closely related species *N. fischeri* (Hammond and Keller, 2005; Hammond *et al*., 2008b).

4.3 RNA silencing in fungi in response to virus infection

Fungi respond to invasive viruses at the cellular level by exhibiting an innate immune defense response called RNA silencing (or quelling). Evidence confirming the presence of a silencing mechanism in fungi against mycoviruses has been identified in e.g. *C. parasitica* (Segers *et al*., 2007), *A. nidulans* (Hammond *et al*., 2008a) and *M. oryzae* (Himeno *et al*., 2010). The effect of dicer-like gene disruptions was examined on virus-infected *C. parasitica* by individual or double

disruption of *dcl1* and *dcl2* (Segers *et al.,* 2007). In virus-free dicer gene deletion mutants, no distinct phenotypic change was observed in all three mutants (Δ*dcl1*, Δ*dcl2* and Δ*dcl1*/Δ*dcl2*) as relative to the virus-free wild-type strain. Conversely, hypovirus (CHV1)-infected Δ*dcl2* and Δ*dcl1*/Δ*dcl2* mutants showed obvious reduced growth rates whereas the virus-infected Δ*dcl1* mutant showed a growth rate comparable to the wild-type. The same experiment was also conducted to study function of four argonaute proteins (*agl1-agl4* genes) which demonstrated that phenotypic change appeared only in the Δ*agl2* mutant infected with hypovirus CHV-1/EP713 (Sun *et al*., 2009). In addition, it has been observed that transcript levels during viral infection of two *C. parasitica* Dicers, *dcl1* and *dcl2* respectively were enhanced approximately 1.5 and 15 fold while *agl2* transcripts showed approximately a 2 fold accumulation. This finding indicated that *dcl-2* and *agl-2* are crucial for RNA silencing in the fungus.

In addition to a role of silencing mechanism against mycovirus infection in *C. parasitica*, RNA silencing of the uncharacterised Aspergillus virus 341 has also been described. Suggestions that the virus was targeted by the *A. nidulans* RNA silencing machinery was confirmed when mycovirus-derived siRNAs were detected in the infected strains which only required a single dicer and single argonaute gene in the process (Hammond *et al*., 2008b). Likewise in *M. oryzae*, MoV2-derived siRNAs from both positive and negative strands of MoV2 viral RNA were detected, suggesting the existence of RNA silencing in this fungus (Himeno *et al*., 2010). However, an accumulation of the MoV2-derived siRNAs appeared to be relatively low as compared to the total small RNAs and this observation is similar to the finding reported for virus-derived siRNAs in *A. nidulans.* On the other hand, a finding from Zhang *et al.* (2008) revealed that accumulation of siRNAs detected in CHV1-EP713-infected *C. parasitica* appeared to be the majority (73%) of total small RNAs isolated. This could be attributed to the unencapsidated nature of CHV1-EP713 which makes the virus more susceptible to host defence response, in contrast to MoV2, which is fully encpsidated in virions. Overall, it clearly indicates that the antiviral defence response of fungi require only DCL-2 and AGL-2 for RNA silencing. This is opposed to the silencing mechanism in plants which requires multiple dicers and argonautes. In addition to mycoviruses, RNA silencing in fungi can be induced in response to transposable elements and hairpin RNA structures.

4.4 Application of RNA silencing

RNA silencing has been adopted to examine gene function for functional genomic studies in many organisms. In filamentous fungi, the approach shows several advantages over

conventional gene knockout strategies such as gene disruption, gene deletion and gene/promoter replacement by homologous recombination (HR) technology which are currently used to generate fungal mutants. The RNA silencing technique can induce transient expression of a specific gene using an inducible promoter thus the effects of RNA silencing are reversible. In addition, the approach also show advantage over conventional gene targeting techniques where multiple copies of a gene could be targeted using a single silencing construct and it is less time-consuming. In contrast to RNA silencing, the conventional approach needs permanent deletion of the target genes which may be lethal to the organism. Also, gene targeting by traditional methods is less efficient due to low frequency of homologous recombination. However, the application of RNA silencing is not efficient since it usually causes non-specific inhibition (off-target effect) of mRNA for example in mammalian cells, resulting from induction of an interferon in response to long dsRNA. In this regard, the problem was resolved using chemically synthesized siRNA or short hairpin RNA homologous to the gene target instead of long dsRNA (Sandy *et al*., 2005; Silva and Schnierle, 2010).

RNA silencing can facilitate the identification of important components in a particular cell under a particular condition. To knockdown or decrease the expression of the target genes, a sequence complementary to the gene of interest is synthesized to make dsRNA or siRNA and then introduced into target cells, subsequently leading to the process of RNA silencing. In clinical applications, RNA silencing has been used in antiviral therapies as well as cancer treatment by interfering with genes that are up-regulated in tumor cells or involved with cell division. In addition, the approach has also been applied to biotechnology and agricultural industries such as reducing natural toxin and allergen levels in food plants (Pai *et al*., 2006; Morandini, 2010; Wang *et al*., 2011).

Application of RNA silencing in human pathogenic fungi e.g. *Cryptococcus neoformans*, *Candida albcans* and *Aspergillus fumigatus* has provided enormous information regarding gene function, pathogenicity and interaction between fungal pathogens and their hosts. For example in the encapsulated pathogenic fungus *C. neoformans* which causes fatal meningoencephalitis, an episomal tandem construct folded into a hairpin structure was used to individually and simultaneously induce silencing of a gene involved in capsule biosynthesis and an auxotrophic marker (Liu *et al*, 2002). In agriculture RNA silencing technology can be possibly adopted to control fungal pathogens of plants. For instance, construction of the vectors to control mycotoxin production and population of *A. parasiticus* and *Fusarium graminaerum* which are economically important plant pathogens respectively infecting corn

and wheat. To this end, if an effective dsRNA delivery system is developed, this could lead to breakthroughs in treating fungal infections in both medical and agricultural fields. Other examples of gene silencing in pathogenic fungi are further discussed in section 5.2.

In recent years, study on fungal RNA silencing has provided an in-depth understanding of silencing pathways of fungi in response to mycovirus infection. Newly discovered knowledge and information have a significant potential to be practically applied in many fields such as in agriculture, biotechnology including therapeutic manipulation. As mycoviruses have currently been used as biological control agents, it might be possible to adopt the knowledge on RNA silencing mechanism to generate recombinant mycovirus as a silencing tool which can be used for functional genomics or controlling virulence of fungal pathogens.

---

5. Development of gene silencing approaches in fungi

RNA silencing has been used as a powerful approach to study gene function in many organisms. In fungi, different gene silencing approaches have been widely developed particularly for fungal plant pathogens. Commonly used technique to stimulate gene silencing *via* derivation of siRNAs is the construction of either dsRNA expressing- or short hairpin RNA (shRNA) expressing plasmids. This technique is termed host-induced gene silencing (HIGS) where plasmid constructs have an ability to either integrate their sequences into the fungal genome or episomally self-replicate. Direct introduction of siRNAs or short dsRNAs has also been described. In recent years, virus-induced gene silencing (VIGS) has become one of novel and powerful reverse genetic tools for functional genomic studies in fungi. Details of individual vector constructs and silencing techniques are discussed below.

5.1 Direct delivery of siRNA/dsRNA

Gene silencing in fungi can be achieved by direct delivery of short dsRNAs/siRNAs or synthesized long dsRNAs into fungal protoplasts (Fig. 6a and 6b). This technique avoids multiple cloning steps of silencing vector construction and has been successfully used to introduce synthesized siRNA in mammalian systems but hardly reported in fungi. Problem on introduction of the sequence-specific siRNA is an 'off-target' effect which causes down-regulation of non-targeted gene. The silencing effect of the approach is usually transient as

concentration of siRNAs is diluted during cell division. However, an advantage of using siRNA for gene silencing is that the approach does not require functional Dicer enzyme in the host cells for degradation of long dsRNAs.

Direct introduction of dsRNA has been successfully applied in a few fungal species. To generate dsRNA elements, sense and antisense strands homologous to the target gene are *in vitro* transcribed and combined to form dsRNA molecules. This technique has been applied in the oomyces (diploid) fungus, *Phytopthora infestan*s, using lipofectin-mediated transfection where transient silencing was observed in most lines and persistent silencing was observed in some lines (Whisson *et al.*, 2005b). Additionally, dsRNA-induced gene silencing was successfully performed in the fungus causing witches' broom disease in cacao, *Moniliophthora perniciosa*, by direct PEG-mediated transformation of the *gfp*dsRNA and endogenous gene homologous dsRNAs (Caribe dos Santos *et al*., 2009).

5.2 Host-induced gene silencing vectors (HIGS)

5.2.1 Short hairpin RNA (shRNA) expressing/silencing plasmid vectors

One effective technique used for gene silencing in fungi is by delivery of gene expression vectors that intracellularly express short hairpin RNAs (shRNAs) controlled by an inducible or constitutive promoter (Fig. 6c). The silencing effect of this approach is more stable, mediating long-term down regulation of the target mRNA (Nakayashiki *et al*., 2005a; Silva and Schnierle, 2010). Silencing vectors for fungi which express shRNA structures have been widely developed, for instance for *Cryptococcus neoformans* (Liu *et al*., 2002); *Neurospora crassa* (Goldoni *et al*., 2004); *Magnaporthe oryzae* (Nakayashiki *et al*., 2005a); *Fusarium culmorum* (Scherm *et al*., 2011); *Phytopthora infestans* (Ah-Fong *et al.,* 2008); *Aspergillus niger* (Oliveria *et al*., 2008) and *A. fumigatus* (Mouyna *et al*., 2004; Henry *et al*., 2007; Khalaj *et al*., 2007) which commonly carry a gene resistance cassette and a transcriptional unit for shRNA expression. However, some drawbacks of the approach have been reported for instance an instability of hairpin structures in the fungal cells. In addition, preparation of the constructs is time-consuming because two-step cloning of inverted orientation fragments is required and transformation of the transgenic fungi is still technically difficult. The use of integrative silencing construct showed various degree of interference (Mouyna *et al*., 2004; Nakayashiki *et al*., 2005a; Khalaj *et al*., 2007). Inconsistent efficiency of silencing may result from random integration of the vector in different locations in the genome which could affect transcriptional activity of the

RNA. Another factor could result from rearrangement of the vector during integration. To overcome problems caused by the integrative vectors, replicative apical membrane antigen 1 (AMA1)-based vector expressing hairpin structures under the control of inducible *cbhB* promoter was successfully developed to silence the *alb1* gene in *A. fumigatus* which demonstrated high frequency of transformation and high efficiency in down-regulation of the *alb1* gene (Khalaj *et al*., 2007).

5.2.2 Bidirectional-dual promoter system

Construction of silencing vectors which express dsRNA by a bidirectional-dual promoter system has been reported for human pathogenic fungi *Histoplasma capsulatum* and *Cryptococcus neoformans* (Rappleye *et al*., 2004), including in *M. oryzae* (Nguyen *et al*., 2008). DsRNA molecules are synthesized from two-opposing promoter cassette vector where sense and antisense RNA strands homologous to the target gene are bidirectionally transcribed and independently form dsRNA in the cells (Fig. 6d). For example, pSD1 vector developed for silencing in *M. oryzae* is composed of opposing *A. nidulans trpC* and *gpd* promoters flanking the multiple cloning site. The advantage of this system is that there is no requirement for two-steps of orientated cloning but the silencing efficacy of this technique is lower than that of shRNA-expressing vector.

Another vector construction which used the same principle of opposing constitutive promoters has been developed. This approach relies on the construction of two separate plasmids containing sense and antisense fragments (Fig. 6e). DsRNA interference can be intracellularly synthesized in the transformed cells from the two constructs and subsequently trigger RNA silencing.

5.3 Virus-induced gene silencing vectors (VIGS)

Virus-induced gene silencing vector (VIGS) is a technique which exploits recombinant viruses to transport sequence homologous to a host gene fragment and subsequently induce gene silencing of the targeted gene. The principle of the approach is based on a defence response against viral infection of the host which initiates degradation of viral RNAs (Fig. 6f). Initiation of the VIGS construction is by engineering the virus genome to carry fragments of host genes or transgenes to be silenced by sequence-homology-dependent mechanisms (Senthil-Kumar and Mysore, 2011). Replication of the recombinant virus vector by the viral RdRP subsequently induces the synthesis of dsRNAs which are cleaved into virus-derived siRNAs in the host

cytoplasm by the enzymes involving in RNA silencing mechanism including help from host RdRP to down-regulate expression of the target gene. Thus, an ability of the vector to systemically infect and replicate in the host cells are key factors for successful application of the technique. Another important factor is that any particular fungal host should possess the basic components of the RNA silencing pathway. VIGS can also be designed to contain a promoter sequence which drives expression of short hairpin RNAs (shRNAs) to induce RNA silencing.

VIGS was originally developed for gene silencing in plants mainly in dicotyledonous plant species. Silencing of endogenous plant genes with VIGS has also been demonstrated in agriculturally important monocot plant species for instance the use of a modified barley stripe mosaic (BSMV)-VIGS system carrying an untranslatable phytoene desaturase fragment downstream of the γ*b* gene to silence genes involving in carotenoid synthesis in barley (Holzberg *et al*., 2002).

To date, application of VIGS for functional genomic studies in fungi has been widely used and several protocols have been developed. For instance, the BSMV-based vector expressing a dsRNA fragment derived from the fungus *Puccinia triticina* genes has been successfully used to silence gene involved in fungal pathogenicity (Panwar *et al*., 2013). VIGS was also successfully used to induce RNA silencing in the plant pathogenic fungus *Collectotrichum acutatum* (strain C71) with recombinant a plant virus vector, tobacco mosaic virus (TMV), carrying the green fluorescent protein (GFP) (Mascia *et al*., 2014) where the vector induced stable silencing in the transgenic fungus for at least two months after infection.

The advantage of using VIGS strategy is remarkable. Firstly, VIGS allows knockdown and avoids knockout of the gene of interest which could lead to lethal phenotypes. Secondly, the method is easy to perform by using only direct infection of the recombinant/engineered virus and without the need to develop transformation techniques as compared to a silencing plasmid vector. Thirdly, the technique is less time-consuming and inexpensive because preparation of the construct and method of transfection are simple which provides the generation of rapid phenotype. However, the discovery of antiviral defense responses of the fungal host becomes an important obstacle that hampers to the use of recombinant viruses as expression and silencing vectors to induce gene silencing in fungi. This results in instability of the chimeric virus genome in host cytoplasms as reported by Suzuki *et al*., (2000). Fortunately and intriguingly, the recent finding of Mascia *et al*. (2014) on stability and silencing efficiency of the TMV-based VIGS vectors has revealed a strong evidence that the approach could be practically

exploited as a reverse genetic tool to transiently silence specific genes without complex transformation technology and stability and silencing efficiency of the vectors could be improved. One important strategy to improve efficiency of the VIGS vector is the selection of gene fragments for silencing. Successful VIGS-induced gene silencing has been commonly reported when fragments of 300-1500 bp were employed. Gene fragments <300 bp can reduce silencing efficiency whereas fragment >1500 bp results in vector instability as the insert is usually lost from the recombinant. In addition, it has been stated that different gene fragments can influence variability of the silencing efficiency, thus, inserts from different regions in the same gene should be examined for effective silencing (Padmanabhan and Dinesh-Kumar, 2009).

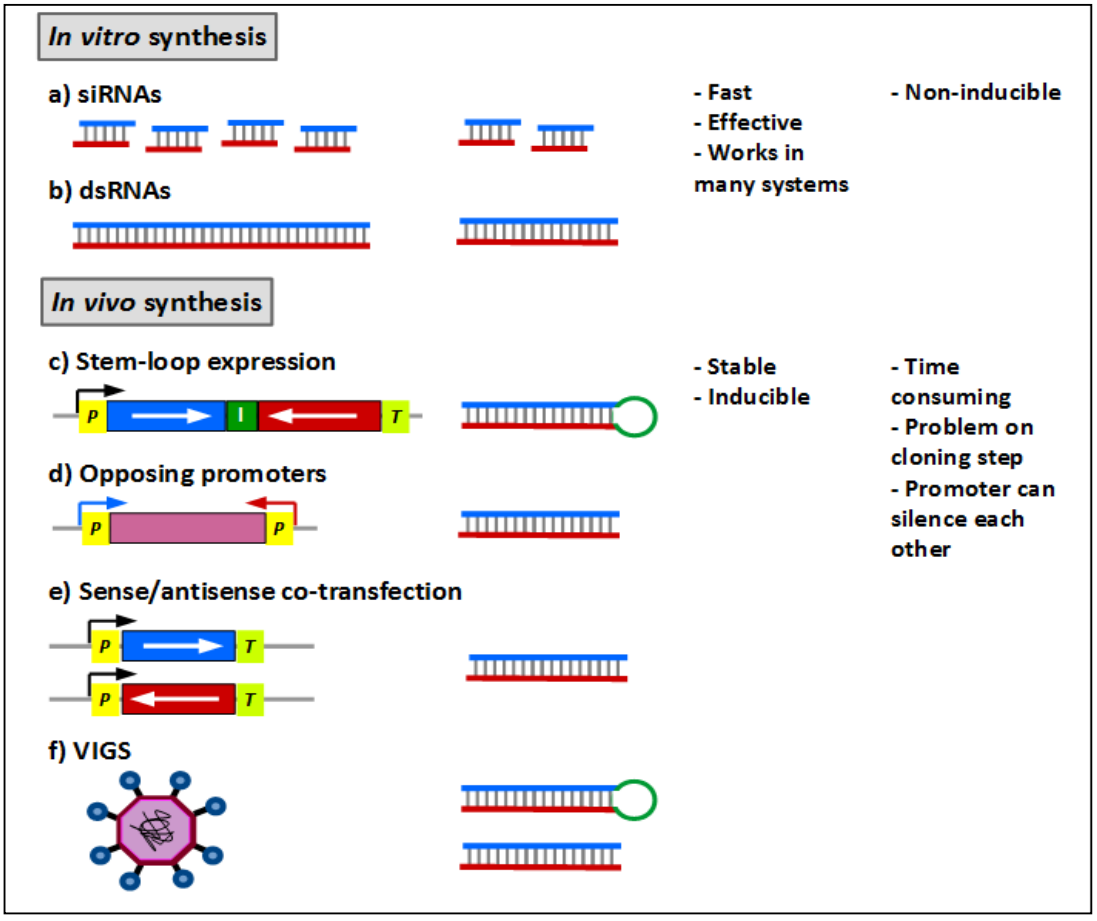


Fig. 6 Methods to generate double-stranded RNA for RNA silencing (Adapted from Hammond *et al*., 2001).

Chapter 2

Materials and Methods

Chapter 2: Materials and Methods

1. Fungal maintenance and harvesting

Aspergillus fumigatus Af293 Nancy Keller (NK) isolated at autopsy from a patient with invasive pulmonary aspergillosis was used throughout this study. The isolate was maintained in 20% (v/v) glycerol solution in the storage tubes and kept at -80 ˚C. The project was registered under the assessment ID: GMIC-1442 and DPIC – 2506. Manipulation of the fungus was carried out in class-II Microbiological Safety Cabinets (CL2-MSC; Envair class-II MSC) in the Flowers Building, CMBI5 *Aspergillus* Laboratory. Working areas were disinfected with fungicidal 10% Distel (formerly Trigene) following UV sterilisation prior to experimental work.

A 20 µl of conidia suspension was inoculated on a centre of an *Aspergillus* Complete Media (ACM) agar plate and incubated for 4 days at 37 ˚C. Ampicillin (100 µg/ml) was added to the growth medium when required. After 4 days incubation, conidia were harvested by scraping off the agar surface using a sterile glass rod. During this step, arms were covered and a face mask was worn to protect from spore attachment and inhalation. Gloves were frequently changed and surface cleaned together with maintained aseptic procedures to avoid cross contamination. A 10 ml suspension of *ca.* $5x10^8$ conidia/ ml was transferred to 500 ml ACM broth and incubated at 37 ˚C for 7 days on a rotary shaker at 150 rpm. After 7 day cultivation, fungal hyphae were harvested by filtration through autoclaved Miracloth™ (Calbiochem, Darmstadt, Germany), rinsed with water and then immediately snap-frozen in liquid nitrogen and then kept in -80 ˚C prior to nucleic acid extraction procedures.

---

2. Nucleic acid manipulation

2.1 Extraction of *A. fumigatus* genomic materials for small scale preparation

A. fumigatus genomic DNA was prepared using the DNeasy® Plant Mini Kit (QIAGEN, UK) with mini protocol provided by the manufacturer. For total RNA extraction, the RNA was prepared using the RNeasy® Plant Mini Kit (QIAGEN, UK) with small scale protocol provided by the manufacturer. DNA contaminants were removed from the total RNA sample using TURBO™ DNase (Ambion, UK) prior to cDNA synthesis.

Quantity and purity of DNA and total RNA were determined by measuring the absorbance at 260 and 280 nm (A_{260}/A_{280}) using a Nanodrop 2000C spectrophotometer (Thermo Fischer). DNA samples showing an A_{260}/A_{280} ratio of 1.7-1.9 indicated high purified DNA whereas RNA sample showing A_{260}/A_{280} ratio of 1.9-2.1 indicated high purified RNA. The quality of the DNA was checked by electrophoresis in a 1% (w/v) agarose gel mixed with 1X TAE buffer in distilled water and stained using SYBR safe for DNA stain (Invitrogen ™). To check the integrity of total RNA, 1% (w/v) agarose gel mixed with 1X TBE buffer in Diethylpyrocarbonate (DEPC)-treated water was used and stained using SYBR Gold Nucleic Acid Gel Stain (Invitrogen ™). For long term storage, purified DNA and total RNA samples were stored at -80 ˚C.

2.2 Extraction of *A. fumigatus* total RNA for large scale preparation

This protocol is designed for extraction of large amounts of RNA. Frozen mycelia were ground in liquid nitrogen to a fine powder using a pre-cooled mortar and pestle. Then 100 mg of the powder was transferred to a pre-cooled sterile 2.0 ml microcentrifuge tube which was transferred immediately to liquid nitrogen. Then 1 ml of TRizol® Reagent (Invitrogen) was added to the homogenate and the contents mixed thoroughly in the fume cabinet. Homogenisation was performed at room temperature for 2-3 min to ensure complete disruption of the tissue. The homogenate was centrifuged at 8000 rpm for 10 min at 4 ˚C and then the supernatant containing RNAs was pipetted into a new pre-cooled microcentrifuge tube (avoid pipetting the interphase fraction which contains DNA). Then 200 µl of ice-cold chloroform was added, the whole mixed vigorously by hand for 5 sec and then centrifuged at 8000 rpm for 10 min at 4 ˚C. The upper phase was transferred into a new pre-cooled microcentrifuge tube and then 200 µl of ice-cold chloroform was added again, mixed vigorously by hand for 5 sec and then centrifuged at 8000 rpm for 10 min at 4 ˚C (two phases should be clearly separated without an interphase. If an interphase was present the chloroform clarification step was repeated). The upper phase was transferred into a new pre-cooled microcentrifuge tube and then 1 vol of ice-cold isopropanol (*ca*. 400 µl) was added, mixed and RNA was precipitated from the mixture following overnight storage at -20 ˚C. The precipitate was collected following centrifugation at 8000 rpm for 30 min at 4 ˚C and the supernatant was discarded. The pellet was washed with 400 µl of 70% (v/v) pre-cooled ethanol (prepared with DEPC-treated water) and then centrifuged at 8000 rpm for 30 min at 4 ˚C. The supernatant was removed and the tube was briefly centrifuged again, removed any remaining ethanol. The pellet was briefly air-dried for 10 min and then resuspended in 50 µl of DEPC-treated water. The solution was kept on ice to dissolve RNA.

2.3 Extraction of viral dsRNA (LiCl fractionation); Adapted from Hull and Covey (1983).

Approximately 25 g of frozen fungal mycelia were ground in liquid nitrogen to a fine powder using a pre-cooled mortar and pestle and transferred to a pre-cooled sterile Nalgene® centrifuge tube (Thermo Fisher Scientific, Inc). Then 30-40 ml of extraction buffer (50 ml/l Tris-HCl buffered phenol (Sigma), 60 g/l 4-aminosalicylic acid (Sigma)), made up to the desired volume with water which was added to the powder and the mixture was shaken on an iced plate for 60 min to homogenize to smooth paste. The mixture was then centrifuged at 10,000 rpm for 40 min at 4˚C (Beckman Coulter Avanti-J-26xP, rotor ID JA 14). The aqueous supernatant containing total nucleic acids was transferred into a new pre-cooled, sterile Nalgene centrifuge tube. Then 15 ml of Tris-saturated phenol was added and the mixture was centrifuged again at 10,000 rpm for 40 min at 4˚C. The upper phase was collected and transferred to a new pre-cooled sterile centrifuge tube by pipetting, 15 ml of chloroform was added and the mixture centrifuged at 10,000 rpm for 40 min at 4˚C. The supernatant was transferred into a new pre-cooled centrifuge tube and the volume measured.

After extraction, LiCl (Sigma Aldrich) fractionation was carried out according to Diaz-Ruiz and Kaper (1978). An equal volume of 4 M LiCl solution was added (final conc. = 2M), mixed and nucleic acids were precipitated overnight at 4˚C. The solution was then centrifuged at 4˚C (10,000 rpm, 50 min) to pellet the precipitated ssRNA and DNA (The pellet containing ssRNA was resuspended in 1 ml DEPC-H_2O and then kept at -80 ˚C). The supernatant containing dsRNA was transferred to a pre-cooled sterile tube and an equal volume of 8 M LiCl solution was added (final conc. = 4M), mixed and nucleic acids were precipitated overnight at 4˚C. The solution was centrifuged at 10,000 rpm for 50 min to collect the dsRNA. The supernatant was discarded immediately to prevent dissolution of the pellet. At this step, dsRNA formed a clear precipitate against the side of the tube. A sample (1 ml) was also taken from solution before discarding. This sample was used as negative control. The pellet was then resuspended in 2 ml of sterile distilled water and stored at -80 ˚C for further use. Prior to the next nucleic acid manipulation, dsRNA samples were purified by phenol-Sevag extraction and cleared of DNA and ssRNA by sequential DNase I and S I nuclease treatment. The quality of the dsRNA was checked by 1% (w/v) agarose gel electrophoresis in 1xTAE buffer using ethidium bromide staining.

2.4 Isolation of mycovirus particles (Virus-like particles, VLPs)

Approximately 50 g of frozen fungal mycelia were homogenized in 2 volumes (w/v) of TE buffer for 3 min in a blender. The homogenate was filtered through Miracloth™ (Calbiochem, Darmstadt, Germany). The filtrate was then transferred into a pre-cooled sterile Nalgene® centrifuge tube (Thermo Fisher Scientific, Inc) and centrifuged at 10,000 x g for 20 min at 4˚C (Beckman Coulter Avanti-J-26xP, rotor ID JA 14). The supernatant was collected, precipitated with 10% PEG-6000 (Sigma Aldrich) and 0.6 M NaCl (Sigma Aldrich) and slowly stirred at 4˚C overnight. After overnight precipitation, the solution was centrifuged at 10,000 x g for 20 min at 4˚C and the supernatant was discarded. The precipitate containing virus was resuspended in 60 ml TE buffer and clarified by additional centrifugation at 10,000 x g for 20 min at 4˚C. The supernatant containing virus was transferred into a 50 ml centrifuge bottle (Beckman; polycarbonate bottle with cap; 29x104 mm) and then subjected to ultracentrifugation (Beckman Coulter Optima™ L-100 xP, rotor ID 45 Ti) at 105,000 x g (30,000 rpm) for 90 min at 4˚C. The glass like pellet was resuspended in 1 ml TE buffer and clarified by low-speed centrifugation at 10,000 rpm for 20 min (AccuSpin™ Micro, Fisher Scientific). The supernatant containing virus particles was collected.

On occasion sucrose gradienting was performed using 1 ml of the supernatant obtained as above. Sucrose solutions (10%, 20%, 30%, 40% and 50% sucrose in TE buffer) were sequentially overlaid in Beckman ultra-clear centrifuge tubes (14x89 mm) and kept at 4 ˚C overnight prior to layering the virus on the top. The tubes were then centrifuged at 33,700 rpm for 3 h at 4 ˚C (Beckman Coulter Optima™ L-100 XP, rotor SW-41Ti). Individual fractions were collected in 2 ml volumes by puncturing the base of the tubes and each dialysed against TE buffer for 24-36 h. Virus particle was then collected by centrifugation at 130,000 x g (40,500 rpm) for 90 min at 4 ˚C using rotor MLS-SO (tube no. 326819).

Pellets were then resuspended in 1 ml TE buffer and further purified by CsCl density centrifugation. A 0.5 ml of virus suspension was layered onto a cushion of 4.5 ml CsCl solution (1.45 g/cm^3 density in TE buffer) using polyallomer centrifuge tube (13x51 mm, #326819, Beckman) and centrifuged at 55,000 rcf (xg) (21,300 rpm) for 90 min at 4 ˚C (SW 55 Ti rotor, Beckman). The virus fraction was collected and dialysed against 10 mM $MgCl_2$ in TE buffer (Slide-A-Lyzer® dialysis cassette, Thermo Scientific, USA) or diluted 10-fold with TE buffer, following the centrifugation at 120,000 x g (45,000 rpm) for 2 h at 4 ˚C to pellet the virus (SW 55 Ti rotor). Pellets were resuspended overnight at 4 °C in 200 µl TE buffer. The purified virus was kept at -80 ˚C

To investigate virus particles, VLPs suspension was stained with 1% (w/v) uranyl acetate on carbon-coated 400-mesh copper grids and examined by transmission electron microscopy (LEO 906E, Zeiss, Germany).

Virus precipitate was also observed under an atomic force microscopy (AFM) as previously described in Magae (2012). Sample was filtered through a 0.2 µm filter (Millipore) and then 1 µl of the filtrate was diluted with 50 µl of 10 mM $MgCl_2$ in TE buffer. RNase III - treated sample was also included to confirm the AFM image was dsRNAs.

2.5 Phenol: Chloroform extraction (Sevag extraction)

Phenol was prepared by melting at 68 ˚C, mixed with 0.1% (w/v) parahydroxy-quinoline (Sigma) and then equilibrated five times with 0.1 M Tris-HCl (pH 8.0). The solution was stored in a dark bottle at 4 ˚C. For extraction, 200 µl of aqueous solution of nucleic acid was mixed with sterile distilled water to a volume of 400 µl. An equal volume of phenol (400 µl) was added in the suspension (prepared in a fume cabinet) and then vortexed for 6-8 sec to denature proteins. The suspension was centrifuged at 13,000 rpm for 5 min to separate the two phases. The upper aqueous phase was collected into a new microcentrifuge tube and 400 µl of Sevag solution (chloroform: iso-amyl alcohol = 24: 1 (v/v)) was added, vortexed and then centrifuged. The upper phase containing dsRNA was collected and nucleic acids were precipitated.

2.6 Purification of viral dsRNAs

The dsRNA was treated sequentially with DNase I and SI nuclease in order to remove contaminant DNA and ssRNA endonucleolytically to yield 5’-phosphoryl-terminated products.

(i) DNase I Treatment

For the removal of dsDNA, 40 µl of dsRNA sample was added to 10 µl 10xRQ1 DNase Buffer (Promega), 25 µl RQ1 DNase (Promega) and 25 µl DEPC-H_2O to make a total 100 µl reaction. The mixture was incubated at 37 °C for 1 h, stored on ice and nucleic acids were isolated as above within 1 h of the procedure being initiated. The mixture was then kept on ice. Phenol/Sevag extraction was performed following the addition of 300 µl water to make a total volume of 400 µl solution.

(ii) S1 Nuclease Treatment

Single stranded RNA was removed from the sample by treating with S1 nuclease (Promega). In a typical reaction 40 µl RNA was incubated at 37°C for 1 h in the presence of 2 µl Promega S1 nuclease (100 U/µl), 10 µl 10 x buffer and the volume adjusted with dH_2O to make a total volume of 100 µl. The treated nucleic acids were isolated and purified as described above.

2.7 Purification of *A. fumigatus* total RNA

A. fumigatus total RNA was treated with TURBO™ DNase (Ambion). The enzyme is used to degrade genomic DNA in RNA preparations prior to reverse transcription-PCR (RT-PCR).

Reaction components – 50 µl reaction

Prior to the reaction preparation, the nucleic acid solution concentration should be monitored. If the concentration is ≥200 ng/µl, dilute it to 10 µg nucleic acid/50 µl.

10 µg of total RNA	x µl
10x TURBO™ DNase Buffer	5 µl
TURBO™ DNase (2 U)	1 µl
DEPC-treated H_2O	to bring volume up to 50 µl

The reaction was centrifuged briefly and then incubated at 37 ˚C for 30 min. To deactivate TURBO™ DNase, 0.1 volume of DNase inactivation reagent (Ambion ® DNA-free™ DNase Kit) was added to the reaction tube. The suspension was incubated at room temperature for 2 min, mixed occasionally and then centrifuged at 12,000 rpm for 2 min. The supernatant containing RNA was transferred to a new sterile microcentrifuge tube and stored at -80 ˚C.

2.8 Nucleic acid precipitation

Nucleic acid was precipitated by adding 1/10 vol of 3 M sodium acetate (pH 5.5) and 2.5 vols of cold absolute ethanol. The solution was immediately inverted and nucleic acids precipitated overnight at -20˚C. Precipitated nucleic acid was pelleted following centrifugation at 13,000 rpm for 25 min at 4˚C. The supernatant was discarded and the pellet was washed by adding iced-chilled 70% (v/v) ethanol followed by centrifugation at 13,000 rpm for 25 min. The supernatant was discarded and the pellet was dried briefly at room temperature for 5-10 min. Avoid drying the pellet for more than 10 min since overdried pellets are very difficult to redissolve. Dried pellets were resuspended by adding 20-30 µl sterile DEPC- H_2O and allowed to dissolve on ice for 30 min. In this step, the whole surface of the tube should be washed

thoroughly since some nucleic acid may attach to the wall. Dissolved nucleic acid solutions were stored at -80 ˚C. Quality of nucleic acid was checked by OD_{260} and OD_{280} measurement using a Nanodrop 2000C spectrophotometer (Thermo Fischer) and also by electrophoresis in agarose gels.

2.9 Agarose gel electrophoresis

Molecular biology grade agarose powder (Melford) at a final concentration of 1.0-2.0% (w/v) was dissolved by heating in a microwave for 1-3 min in 1x Tris acetate EDTA (TAE) buffer. To prepare a 50x TAE concentrate dissolve 242 g of Trizma base (Sigma) and 100 ml of 0.5 M EDTA (pH 8.0) in 800 ml distilled water and 57.1 ml of glacial acetic acid make up to 1L and autoclave. Ethidium bromide (10 mg/ml) (Sigma) or SYBR safe DNA gel stain (Invitrogen™) was added to a final concentration of 50 µg/100 ml after the melted solution cooled to 55-60 ˚C. The agarose was poured into a casting tray and left to set for 45 min. Before loading, the samples were mixed with 5x gel loading buffer (0.2% bromophenol blue (w/v) and 50% glycerol in distilled water), alongside a DNA size marker (Hyperladder™ I, BIOLINE). Electrophoresis was performed applying a voltage of between 30-80 V in 1x TAE buffer for 1-2 h and then visualised under a UV transilluminator (Bio-Rad) or blue-light transilluminator (Syngene).

2.10 Recovery of the nucleic acid samples from agarose gels

To extract dsRNA from agarose gels the BIO-101 RNaid kit (Anachem) was used. Samples of dsRNA were subjected to gel electrophoresis at 67 V for 2 to 3 h so that individual dsRNA species could be distinctly separated. Gels were then exposed to UV irradiation for very short time to reduce degradation. The dsRNA bands of interest were excised with a sharp scalpel blade and kept in separate sterile Eppendorf tubes. Individual dsRNA samples were then extracted following the supplied protocol.

To extract DNA, the MinElute® Gel Extraction kit (QIAGEN) was used to cleanup DNA fragments excised from gels.

2.11 Protein analysis by SDS-PAGE

Purified AfuTmV-1 proteins were separated in 12% SDS-polyacrylamide gels. The stacking and separation gels were prepared using fresh reagents as described in Table 3 and polymerized for 2 h.

Samples (5 µl) were mixed with an equal volume of 2X Laemmli buffer containing β-mercaptoethanol (Sigma) and then denatured at 99 ˚C for 5 min. The denatured samples were immediately snap cool on the ice and then loaded into each well. Page Ruler^{TM} plus prestained protein ladder (10 µl, #26619, Thermo Scientific, USA) was used as a protein marker. Fractionation of the protein samples was performed in 1x running buffer (10x running buffer: 30.3 g Tris, 144 glycine, 10 g SDS and dH_2O up to 1 L) for 80 min at 180V (BioRad) and the gels were stained by gently shake in 0.1% (w/v) Coomassie brilliant blue R-250 (CBB, BioRad, USA) for 2 h to overnight until the protein bands were visualized. Gels were destained by rinsing in a large volume of deionized water, 2-3 times for 5 min each. Gels were stored in deionized water.

Table 3. Recipe for stacking and resolving gels.

	Stacking gel	**Resolving gel**		
	4%	**7.5%**	**12%**	**X%**
30% Acrylamide/bis (Sigma)	1.98 ml	3.75 ml	6.0 ml	0.33x X ml
0.5M Tris-HCl, pH 6.8	3.78 ml	–	–	–
1.5M Tris-HCl, pH 8.8	–	3.75 ml	3.75 ml	3.75 ml
10% SDS	150 µl	150 µl	150 µl	150 µl
dH_2O	9 ml	7.28 ml	5.03 ml	11.03-(0.33x X)ml
TEMED*	15 µl	7.5 µl	7.5 µl	7.5 µl
10% APS **	75 µl	75 µl	75 µl	75 µl
Total volume	15 ml	15 ml	15 ml	15 ml

* TEMED: N,N,N',N'-tetramethylethylene diamine (Sigma)

** APS: Ammonium persulfate (Sigma), prepare fresh solution in water

2.12 In-gel protein digestion

In-gel protein digestion was performed to identify the separated proteins using the protocol provided with the Pierce^{TM} trypsin protease (Thermo Scientific, USA) with modification. The gel slices containing stained proteins were cut into 1x1 mm pieces, de-stained using 200 µl of 100 mM ammonium bicarbonate/50% ACN and then incubated at 37 °C for 30 min. Destaining buffer was removed and the de-staining step repeated 3 times until the stain was completely removed. Gels were shrunk by adding 50 µl ACN and incubation for 15 min at room temperature. ACN was removed and gels were air-dired for 10 min. Pierce^{TM} trypsin protease (50 µl) which cleaves at lysine-arginine residues was added to the samples, which were then digested at 37° C for 8-24 h. Gel pieces were extracted 3 times by adding 50 µl of 50%

ACN/0.1% TFA solution, incubated at 37° C for 5-15 min and then evaporated using a vacuum concentrator.

Partial identification of the protein sequences was carried out at the Department of Life Sciences, Imperial College London. Sequences of the tryptic digests were analyzed by Q-TOF mass spectrometry (MS/MS) using reverse-phase HPLC system. The molecular weights of each peptide fragment were compared to the protein sequences predicted from the nucleotide sequences of ORFs 1-4 of AfuTmV-1 using the MASCOT programme (Matrix Science, Ltd, London, UK).

---

3. Construction of cDNA library from dsRNA by random-PCR (rPCR)

Th procedure adapted from Froussard (1992) was used to construct a library of cDNA clones from the uncharacterised dsRNAs present in the Af293 isolate of *A. fumigatus*. This method allows the production of a cDNA library from small amounts of dsRNA uses a 26 nucleotide primer containing a random hexamer at its 3'-terminus for cDNA synthesis, Froussard-FOR and a PCR amplification step which utilises the same primer in combination with a complementary reverse primer, Froussard REV**.**

Froussard-FOR	5'- GCCGGAGCTGTGCAGAATTCNNNNNN
Froussard-REV	5'- GCCGGAGCTGTGCAGAATTC

Viral dsRNA (8 µl, 0.5 µg) was mixed with 2 µl of Froussard-FOR primer (100 µM) and the mixture denatured by the addition of 2 µl methyl mercuric hydroxide (CH_3HgOH; 100 mM. SERVA), following incubation at room temperature for 20 min and snap cooling on ice for 1 min.

First strand cDNA was synthesized as follows; 12 µl of denatured dsRNA was added to the first strand cDNA reaction mixture which was pre-heated at 50-55 ˚C for 1 min. The reaction mixture was prepared by adding the following components in a nuclease-free microcentrifuge tube, mixed by pipetting gently:

4 µl	5x first-strand buffer (Invitrogen)
1 µl	0.1 M DTT (Invitrogen)
1 µl	RNasin® Ribonuclease Inhibitor (Promega)
1 µl	10 mM dNTP mix (10 mM each dATP, dGTP, dCTP and dTTP; Promega)

Then 1 µl of SuperScript™ III reverse transcriptase (200 U/µl; Invitrogen) was added to the mixture above containing denatured dsRNA. The reaction was incubated at 50˚C for 1 h and then heated at 70 ˚C for 15 min to inactivate the enzyme. In other cDNA synthesis experiments that used either a random oligonucleotide primer or a sequence-specific primer the temperature of incubation for cDNA synthesis was 50˚C or 55˚C respectively.

The first strand cDNA reaction was size-fractionated using a Nanosep 30K column (Millipore). Sterile distilled H_2O was added in the reaction to make a volume of 500 µl solution and then centrifuged at 5000 rpm for 6 min, resulting in the recovery of *ca*. 100 µl of cDNA solution. Recovered cDNA was then subjected to PCR amplification.

PCR amplification with GoTaq® DNA polymerase

PCR amplification of cDNAs was performed in a 50 µl reaction mixture in thin-walled 0.5 ml PCR Eppendorf tubes and contained the following components:-

5 µl	cDNA (from first-strand reaction)
29.6 µl	DEPC-treated H_2O
10 µl	5x GoTaq® Reaction Buffer (Promega)
2 µl	10 mM dNTPs mix (Promega)
2 µl	Froussard-REV primer (100 µM)
1 µl	RNase H (NEB)
0.4 µl	GoTaq® DNA polymerase (5 U/µl; Promega)

Tubes were then placed in a thermal cycler (MJ Research Thermal cycler Model PTC 220 DYAD) and the following thermal cycling PCR amplification regime employed:-

Initial denaturation	95 ˚C	2 min	1 cycle
Denaturation	95 ˚C	1 min	30 cycles
Annealing	60 ˚C	1 min	
Extension	72 ˚C	3 min	
Final extension	72 ˚C	5 min	1 cycle
Hold	4 ˚C		1 cycle

* Annealing temperature should be optimized for each primer set based on the primer melting temperature (*Tm*). To determine the suitable annealing temperature for each primer pair, oligonucleotide primers >20 nt in length typically anneal for 10-30 sec at 3 ˚C above the *Tm* of

the lower *Tm* primer. However for primers <20 nt in length, an annealing temperature equivalent to the *Tm* of the lower primer should be used.

PCR products were purified using the QIAquick® PCR Purification kit (QIAGEN) and then separated by gel electrophoresis and visualized with SYBR safe DNA gel stain.

---

4. Genome walking strategy

4.1 Primer design

Sequence specific oligonucleotide primers were designed based on known sequences which were obtained from the random-primed PCR experiment using Froussard primers and/or the random hexamer. The oligonucleotide primers were generally 20-40 nt in length and ideally have a GC content of 40-60 %. Web based tools for primer designs were used to design or analyse primers such as;

1) Primer3 (http://frodo.wi.mit.edu/primer3) was used to analyze specific target regions and forward and reverse primer sequences.

2) OligoAnalyzer 3.1 (http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/default.aspx) was used to double check the primers identified by Primer3.

3) NEB Tm Calculator (www.neb.com/TmCalculator) was used to determine the annealing temperature when using Phusion DNA polymerase.

4.2 Synthesis of cDNA from dsRNA by genome walking

The PCR-based genome walking method has been commonly used to identify flanking genomic fragments from a known sequence. The procedure for dsRNA denaturation, cDNA synthesis and PCR amplification of the genome walking strategy is similar to the random-PCR method (rPCR). However a sequence specific primer designed from the known sequences of each dsRNA was used to synthesize first strand cDNA instead of using the Froussard primer. For the PCR amplification step, 1 µl of random hexamer (Promega; 100 mM) and 1 µl of sequence specific primer (100 mM) were added to the reaction mix. List of the primers used for the PCR amplification of each dsRNA was shown in Table 4.

Table 4. Specific oligonucleotide primers used for PCR amplification of AfuTmV-1 and AfV-F dsRNAs.

dsRNA components	Primers
AfuTmV-1 dsRNA1	**5'AfuNK-dsRNA1**; 5'-CACGTTTCCTGTAACCCACC-3'
	3'AfuNK-dsRNA1 (2); 5'-CCAGGTTGGGTTGACACC-3'
	3'AfuNK-dsRNA1(3); 5'-CCTAAAGCAGCCAAGATCGT-3'
	3'AfuNK-dsRNA1(4); 5'-AACTCCCTGTTCGCCAAACT-3'
	3'AfuNK-dsRNA1(5); 5'-CTCCACACCCCACGAGAC-3'
	GDNfor-dsRNA1; 5'-CGAGGAGACCTCACTTACGC-3'
	GDNrev-dsRNA1; 5'-CCCAAGCTTGACAGTGACCT-3'
AfuTmV-1 dsRNA2	**UCV gap FOR2**; 5'-ATGTGCGGGAACCAGGACGTCGT-3'
	UCV gap REV2; 5'-CGAACAGTGTATTGAGGGTGTC-3'
	15-RLM-UP; 5'-TGGTCCTGATTCTCGTAGCC-3'
	UCV2-RLM-UP; 5'-GACCTCATCCACCCCGTAAT-3'
	13-RLM-DWN; 5'-AGATCCGCCTACGCTCATAA-3'
AfuTmV-1 dsRNA3	**5'AFUNK-DSRNA3**; 5'-GTACAGGGGCTGGTCATCA-3'
	3'AFUNK-DSRNA3; 5'-CATCATGGCGGAGCAACTA-3'
AfuTmV-1 dsRNA4	**5'AfuNK-dsRNA4-2;** 5'-GTGTGGAGAAGGTCCAGGAG-3'
	3'AfuNK-dsRNA4-2; 5'-GTATCGCTCGCGTCAAGG-3'
	3'AfuNK-dsRNA4-3; 5'-GTAGGGCATCCACGCCTAAG-3'
AfV-F dsRNA4	**14-REV-UP**; 5'-CCTCTCACCCATGAACGATC-3'
	AFV-4-uni; 5'-CTGGGTCATGTAATCTCCTTTGT-3'

4.3 RNA linker-mediated rapid amplification of cDNA ends (RLM-RACE); Adapted from Coutts and Livieratos (2003).

RLM-RACE is a PCR-based technique which facilitates the cloning of full length cDNAs and is often used to specifically amplify 5'- or 3'-termini of dsRNAs. An oligonucleotide primer termed, LIG-Rev [5'-kinated-PO_4; 3'-OH-blocked]; 5'GATCCAACTAGTTCTAGAGCGG [3AC7]) was ligated to the 3' end of dsRNA using T4 RNA ligase (NEB). The dsRNA (7 µl, 0.5 µg) was denatured together with the LIG-Rev primer (1 µl: 100 pmol) by heating the mixture at 90 ˚C for 2 min and then snap-cooling on ice.

The ligation reaction mixture was prepared by adding 10 µl T4 RNA buffer (NEB), 1 µl ATP (10 mM; Promega) and 62.5 µl dH_2O (If the buffer contains ATP, add 63.5 µl dH_2O). This mixture was incubated at 37 ˚C for 10 min and then cooled at room temperature for 2 min prior to the addition of 10 µl DMSO, 2.5 µl RNase Inhibitor (Promega), 1 µl T4 DNA ligase (Promega; 1-3U/µl) and 5.0 µl T4 RNA ligase (NEB; 10U/µl). The reactants were then pelleted to the base of the Eppendorf reaction tube and the denatured dsRNA from above added prior to incubation at 17 ˚C for 15 h.

Following this incubation of the mixture, 400 µl dH_2O was added to the reaction to make a final volume of 500 µl. The mixture was size-fractionated and concentrated using a Nanosep 30K column (Millipore) and then centrifuged at 5000 rpm for 6 min to eliminate non-ligated dsRNA, resulting in the recovery of *ca*. 100 µl of ligated-dsRNA solution. To fill in gaps, 100 µl of ligated-dsRNA was added to 40 µl of 5x GoTaq® Reaction Buffer, followed by the addition of 4.8 µl of 10 mM dNTPs mix, 1 µl of GoTaq® DNA polymerase and 54.2 µl dH_2O to make a total volume of 200 µl. The mixture was incubated at 68 ˚C for 3 h and the ligated dsRNA precipitated overnight (at -20 ˚C) following addition of 0.1 vol 3 M sodium acetate and 2.5 vols cold absolute ethanol. The ligated dsRNA was recovered by centrifugation at 13,000 rpm for 25 min at 4˚C. The supernatant was discarded and the pellet was washed by adding iced-chilled 70% ethanol followed by centrifugation at 13,000 rpm for 25 min. The supernatant was discarded and the pellet was dried under a lamp for 30 min prior to resuspension in 8 µl sterile DEPC-treated water on ice for 30 min.

The ligated dsRNA (8 µl) was denatured by adding 2 µl CH_3HgOH (100 mM) together with 2 µl of a LIG-For oligonucleotide primer (5'- CCGCTCTAGAACTAGTTGGATC-3'; 100 µM). The mixture was incubated at room temperature for 20 min and then rapidly cooled on ice for 1 min.

First strand cDNA was then synthesized; using 12 µl of denatured dsRNA as template which was added to the reaction mixture below which was pre-heated at 50-55 ˚C for 1 min. The reaction mixture was prepared by mixing the following components:-

10	µl	5x First-strand buffer (Invitrogen)
4	µl	0.1 M DTT (Invitrogen)
1.25	µl	RNasin® Ribonuclease Inhibitor (Promega)
2.4	µl	10 mM dNTPs (10 mM each dATP, dGTP, dCTP and dTTP; Promega)
19.35	µl	DEPC-H_2O

SuperScript™ III reverse transcriptase (1 µl of 200 u/µl) was added to the mixture containing denaturated dsRNA and the whole reaction was incubated at 50˚C for random hexamer or degenerate primers, or 55˚C for sequence specific primers for 1 h and then 70 ˚C for 15 min to inactivate the enzyme. First strand cDNA reaction mixture was size-fractionated using a Nanosep 30K column as before prior to PCR amplification.

PCR was performed in a 100 µl reaction mixtures in thin-walled 0.5 ml PCR Eppendorf tubes and contained the following components:-

50 µl	cDNA (from first-strand reaction)
25.1 µl	DEPC-H_2O
20 µl	5x GoTaq® Reaction Buffer (Promega)
2.4 µl	10 mM dNTPs mix (Promega)
1 µl	Lig-For primer (100 µM)
1 µl	Sequence specific primer (100 µM)
1 µl	RNase H (NEB)
0.5 µl	GoTaq® DNA polymerase (5 units/ µl; Promega)

Tubes were then placed in the thermal cycler and the following thermal cycling PCR amplification regime employed, specifically for GoTaq® DNA polymerase:-

Initial denaturation	95 ˚C	2 min	1 cycle
Denaturation	95 ˚C	1 min	30 cycles
Annealing	60 ˚C	1 min	
Extension	72 ˚C	3 min	
Final extension	72 ˚C	5 min	1 cycle
Hold	4 ˚C	forever	1 cycle

---

5. Cloning techniques

5.1 Ligation

Purified PCR amplicons were directly ligated into pGEM®-T Easy (Promega, Madison, WI, USA) according to the manufacturer's instruction. Ligation reactions were set up as described below;

5 µl	2X Rapid Ligation Buffer (vortex before each use)
1 µl	pGEM®-T Easy Vector (50 ng)
3 µl	purified PCR product
1 µl	T4 DNA Ligase (3 Weiss units/µl)
10 µl	total volume

To produce the maximum number of transformants, the ligation reaction was incubated overnight at 4°C.

5.2 Preparation of *Escherichia coli* JM 109 competent cells

A 1 µl of *E. coli* JM109 High Efficiency Competent Cells (≥1 × 10^8 cfu/µg DNA) were grown overnight in 5 ml 2X LB broth at 37 °C with shaking (150 rpm). After overnight incubation, 500 µl of the culture was inoculated into 50 ml of 1X LB broth, incubated again at 37 °C with shaking for 3 h until the optical density (OD) at 600 nm reached 0.4-0.6. The cells were transferred to a pre-cooled 50 ml centrifuge tube and then chilled on ice for 10 min, followed by centrifugation at 3000 rpm for 10 min at 4 °C. The supernatant was discarded and the pellets were resuspended in 5 ml pre-cooled TSS (transformation storage solution), and then 150 µl of cell suspension was aliquoted into pre-cooled Eppendorf tubes, put immediately in a liquid nitrogen container and then stored at -80 °C.

5.3 Transformation of plasmids into *E. coli*

An aliquot of JM109 competent cells was removed from storage at -80 °C and placed on ice until thawed (5 min). The *E. coli* cells were mixed gently by flicking the tube, avoiding pipetting since the competent cells are very fragile. The ligation reaction was briefly centrifuged to collect the contents and a 40 µl of the competent cells was carefully transferred into a sterile microcentrifuge tube containing 5 µl of the ligation mixture on ice. The mixture was gently flicked and placed on ice for 20 min. The cells were heat-shocked for 45 sec in a water bath at exactly 42 °C (do not shake), and then immediately put on ice for 2 min. Then 900 µl of recovery SOC medium (10 ml LB medium, 50 µl 2 M $MgCl_2$, 200 µl 2 M glucose) was added to the cells prior to incubation for 1.5 h at 37 °C with 150 rpm. After incubation the cells were briefly centrifuged at 13,000 rpm for 30-60 sec and 800 µl of the supernatant was discarded. The remaining 200 µl loose pellet of cells was resuspended and then spread onto duplicate LB plates containing 100 µg/ ml ampicillin/ 200 mg/ml IPTG/ X-Gal plates for selection. The plates were incubated overnight (16-24 h) at 37 °C. After the overnight incubation, plates were stored at 4 °C to enhance blue/white selection efficiency of the transformants.

5.4 Screening of the transformants

Transformant clones which present as white colonies, were collected using sterile toothpicks. Each clone was transferred into 5 ml LB medium containing 7 µl ampicillin (0.05 g/ml). The cultures were incubated overnight (12-16 h) at 37 °C with shaking (150 rpm). Plasmid DNA was isolated and purified using the QIA prep Spin Miniprep kit (QIAGEN). Up to 20 µg of plasmid DNA was purified from 5 ml overnight cultures of *E. coli* in LB medium.

5.5 Restriction enzyme digestion

Single digestion

DNA plasmids purified using the QIA prep Spin Miniprep kit (QIAGEN) were digested with restriction enzyme *Eco*RI (NEB) as shown below. The reactants were combined in Eppendorf tubes in the order given:

11 µl	sterile deionized distilled H_20
2 µl	10X *Eco*RI enzyme buffer (buffer H)
6 µl	plasmid DNA
1 µl	*Eco*RI (20 units/ µl)
20 µl	total volume

The mixture was incubated at 37 °C for 2 h and then the restricted DNA and unrestricted recombinant plasmid DNA examined by electrophoresis in agarose gels.

Double digestion

The double-digest reaction was set up as followed:-

9.8 µl	sterile deionized distilled H_20
2 µl	10X enzyme buffer
6 µl	plasmid DNA
0.2 µl	BSA
1 µl	enzyme 1 (20 units/ µl)
1 µl	enzyme 2 (20 units/ µl)
20 µl	total volume

The mixture was incubated at 37 °C for 2 h or according to the manufacturer's protocol depending on the enzyme used.

Sequential digestion was performed when two incompatible restriction enzyme buffers were used or different incubation temperatures were required.

5.6 Sequencing and phylogenetic analyses

Sequencing reactions were prepared by adding 8.46 µl of the template DNA (miniprep-purified plasmid DNA, pGEM®-T Easy) and 1.54 µl of universal forward primer (Sigma). Plasmids containing inserts >600 bp were also sequenced with the universal reverse primer in separate reactions.

Universal primers	M-13 forward	5’ -GTAAAACGACGGCCAGT-3’
	M-13 reverse	5’ -AACAGCTATGACCATG -3’

All sequence samples were sent to the MRC CSC Genomics Core Laboratory for sequencing. Sequence data were analyzed using the bioinformatic tools and programs listed below;

BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi
MAFFT alignment	http://mafft.cbrc.jp/alignment/server
SIAS	http://imed.med.ucm.es/Tools/sias.html
CAP3	http://doua.prabi.fr/software/cap3
Protein tools	http://bio.lundberg.gu.se/edu/translat.html
Protein tools	http://web.expasy.org/translate
PROSITE	http://prosite.expasy.org/
Conserve domain	http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml
Pfam	http://pfam.sanger.ac.uk/
CLUSTALW	http://www.ebi.ac.uk/Tools/msa/clustalw2/
CLUSTALX (2.0)	www.clustal.org/clustal2/
ORF finder	www.ncbi.nlm.nih.gov/gorf/gorf.html
MFOLD	http://mfold.rna.albany.edu/?q=mfold
CentroidFold	http://www.ncrna.org/centroidfold/
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/
SignalIP 4.1	http://www.cbs.dtu.dk/services/SignalP/
FSFinder	http://wilab.inha.ac.kr/fsfinder2/

---

6. Northern blot hybridisation of the mycovirus, AfuTmV-1 dsRNAs

6.1 *In vitro* transcription-labeling of RNA with digoxigenin

The DIG Northern Starter Kit (Roche) was used to generate DIG labelled, single stranded RNA probes of defined length by *in vitro* transcription of template DNA in the presence of digoxigenin-UTP, using SP6, T7 or T3 RNA polymerases.

6.1.1 DNA template preparation

Representative plasmids of each dsRNA were retransformed and then isolated. Plasmid DNA was linearised using a restriction site in the polylinker, downstream of the cloned insert to be transcribed (multiple cloning site of the SP6 or T7 promoter region) and the length of the sequence to be transcribed should be in the range of 200 to 1000 bp. A restriction enzyme which leaves a 5' overhang was selected to avoid transcription of undesirable sequences using the NEB cutter V2.0 program (http://tools.neb.com/NEBcutter2). DNA plasmids and conditions used for restriction digests are shown in Table 5.

Table 5. DNA plasmids and conditions used for restriction digest for probe preparation.

Plasmid name	dsRNA	Insert size (bp)	Restriction enzyme	Digestion temp (˚C)	RNA polymerase for *in vitro* transcription
Frous1kb(1)	1	400	*Nde*I	37	T7
	1	400	*Nco*I	37	SP6
BamF	2	800	*Sal*I	37	T7
	2	800	*Sac*II	37	SP6
Race3a	3	730	*Nde*I	37	T7
	3	730	*Nco*I	37	SP6
BamR11F	4	900	*Sal*I	37	T7
	4	900	*Nco*I	37	SP6

The reactants were combined in the tube and a single digestion performed as described above. The linearised plasmid DNA was purified using the QIAquick® PCR Purification kit (QIAGEN) and then examined by Nanodrop reading and electrophoresis in agarose gels

6.1.2 *In vitro* transcription using the MAXI Script®T7 Kit

The MAXIScript®T7 Kit (Ambion®) was used to *in vitro* synthesise the unlabeled RNA probe *via* the T7 promoter. The transcription reaction was assembled at room temperature and all the components were added in order as shown below:-

To 20 µl	Nuclease-free water
1 µg	Linearized plasmid DNA template
2 µl	10x Transcription buffer
1 µl	10 mM ATP
1 µl	10 mM CTP
1 µl	10 mM GTP
1 µl	10 mM UTP
2 µl	T7/SP6 Enzyme mix
20 µl	total volume

The mixture was pipetted gently, microfuged briefly to collect the reaction mixture at the bottom of the tube and then incubated at 37 ˚C for 1 h. To remove the template DNA, TURBO DNase (1µl) was added to the reaction, incubated at 37 ˚C for 15 min and then a 0.1 µl of 0.5 M EDTA was added to the tube to stop the reaction. The transcript was then examined by Nanodrop reading and electrophoresis in agarose gels.

6.1.3 DIG-RNA labeling

The labeling procedure is designed for 1 µg of DNA template. RNA transcripts were labelled in an *in vitro* transcription reaction with digoxigenin-11-UTP using a labelling mixture and an optimized transcription buffer shown below:-

10 µl	1 µg linearised plasmid DNA
4 µl	5x labeling mix (vial 1a)
4 µl	5x transcription buffer (vial 1b)
2 µl	T7/SP6 RNA polymerase
20 µl	total volume

The reaction tube was centrifuged briefly and then incubated at 42 ˚C for 1 h. After the incubation, DNase I was added to remove template DNA following incubation for 15 min at 37 ˚C. The reaction was stopped by adding 2 µl 0.2 M EDTA (ethylene diamino-tetracetic acid, pH

8.0). The labelled-transcript was divided into 5 µl aliquot and then checked by Nanodrop reading.

6.2 DIG northern blot hybridisation

6.2.1 Separation of dsRNA samples on an agarose gel

Standard protocol for gel electrophoresis was used for dsRNA separation as described in Sambrook *et al*. (1989). A 1% (w/v) agarose gel was prepared in 1x Tris acetate EDTA (TAE) buffer and then SYBR safe DNA gel stain (Invitrogen™) was added to a final concentration of 50 µg/100 ml. The agarose was poured into a casting tray and left to set for 45 min. Before loading, the samples were mixed with 5x gel loading buffer, alongside a DNA size marker (Hyperladder™ I, BIOLINE). Electrophoresis was performed applying a voltage of 60 V in 1x TAE buffer for 5 h until the dsRNA was well-separated and then visualised under a blue-light transilluminator (Syngene).

6.2.2 Transferring dsRNA to a membrane and fixation

Agarose gel was denaturated by soaking in 0.25 N HCl for 20 min. The gel was placed on a new tray and then immersed in 100 mM NaOH for 30 min, followed by neutralization twice in 100 mM Tris-HCl (pH 8.0) for 20 min.

A blot transfer sandwich was set up by placing a fiber pad, which had been soaked with 1x TAE buffer, onto a blot cassette with the gray side down. A piece of pre-wetted chromatography paper (3 MM CHR, Whatman) was then placed on the fiber pad, followed by the equilibrated gel. A sterile pipette was used to roll over the gel to remove all air bubbles that formed between the gel and paper. A piece of positively charged Amersham Hybond™-N membrane (GE Healthcare) was carefully placed directly on top of the gel and then the blot assembly was completed by adding a soaked sheet of Whatman paper and fibre pad on the top of the blot stack. The cassette was closed firmly and placed in a Trans-Blot® electrophoresis transfer tank (BIO-RAD) filled with 1xTAE transfer buffer. A stir bar was added to maintain even buffer temperature and ion distribution in the tank. A power supply (BIO-RAD) was connected to the tank and the transfer achieved following overnight electrophoresis at 15V, 0.2 A at 4 ˚C.

Upon completion of electrophoresis the blotting sandwich was disassembled. The membrane was placed on Whatman 3MM paper soaked with 2xSSC and then the RNA was fixed to the membrane by UV-cross linking for 25 min at the energy level of 2000x100 µJ/cm^3 (UVP CL-1000 Ultraviolet Crosslinker).

6.2.3 Hybridisation with RNA probes

DIG Easy Hyb (10 ml) was prewarmed to hybridization temperature at 68 ˚C for 20 min in a UVP HC-3000 HybriCycler™ hybridization oven. The membrane was placed in a Falcon tube containing pre-warmed DIG Easy Hyb and then pre-hybridised at 68 ˚C for 30 min with gentle agitation in the cycler. DIG-labeled RNA probe (5µl) was denatured by boiling at 95 ˚C for 5 min and rapidly cooled in ice. The denatured probe (100 ng/ml) was then added to a prewarmed DIG Easy Hyb (4 ml) and mixed well but avoiding foaming. The pre-hybridisation solution in the tube containing membrane was poured off and then probe/hybridisation mixture was added to the membrane. The hybridisation tube was incubated for overnight at 68 ˚C with gentle agitation. After hybridisation, the solution was poured off and then membrane was washed in an RNase-free plastic container. Low stringency buffer (2xSSC, 0.1% SDS) was used to wash the membrane twice for 5 min at room temperature in the HybriCycler, followed by a high stringency wash with a prewarmed 0.1xSSC, 0.1% SDS solution at 68 ˚C for 15 min, twice in the HybriCycler.

6.2.4 Immunological detection

After hybridization and stringency washing the membrane was rinsed briefly for 5 min in a 50 ml washing buffer on a rotary shaker at room temperature. The buffer was poured off and then membrane was incubated in 100 ml blocking solution for 30 min, followed by incubating in 50 ml antibody solution (containing Anti-digoxigenin-AP, Fab-fragments, conjugated to alkaline phosphatase) for 30 min on the rotary shaker. The membrane was washed twice in 100 ml washing buffer for 15 min, equilibrated for 5 min in 100 ml detection buffer on the rotary shaker and then placed on a development folder with RNA transfer side facing up. Approximately 1 ml of CDP-Star solution (containing chemiluminescent substrate for alkaline phosphatase) was quickly applied onto the membrane until evenly soaked. The second sheet of the folder was immediately covered on the membrane to spread the substrate evenly and without air bubbles over the membrane, and then the edge of the folder was sealed. Membranes were not allowed to dry at any stage since it would result in dark background. Enzymatic dephosphorylation of CDP-Star leads to a light emission at a maximum wavelength of 465 nm. The membrane was exposed to an imaging device developed by phosphorimaging (FujiFilm LAS-3000 Imaging System) for 2-10 min at -25 ˚C.

---

7. Curing of *A. fumigatus* Af293 from AfuTmV-1 infection and effects of virus on fungal growth and morphology

7.1 Curing of *A. fumigatus* Af293 NK isolate with cycloheximide

Viral dsRNA was eradicated from the fungus by treatment with cycloheximide. Cycloheximide is a protein synthesis inhibitor which blocks translational initiation and elongation during protein synthesis. It has been used to eradicate mycoviral infection from several different fungi (Robinson and Deacon, 2002; Dalzoto *et al*., 2006). The procedure used in this study was adapted from Van Diepeningen *et al*. (2006) with modifications. Sequential hyphal tip isolation was performed to eliminate the virus titre. The cultures were point-inoculated onto ACM agar plates supplemented with a range of cycloheximide concentration (75, 100, 125 and 150 mM) and incubated at 37 ˚C for 5 days. The cultures were passaged for 5 generations and then recovered on normal ACM agar. Using an infected wild type as a control, colonies were selected based on morphological or pigmentation differences as potentially cured isolates.

The potentially cured state of of three sub-cultures (NK105, 106 and 125) were verified by gel electrophoresis of the extracted virus dsRNAs and specific RT-PCR amplification of a fragment of AfuTmV-1 dsRNA2. Virus dsRNA samples were prepared using the RNeasy® Plant mini kit. Northern blot hybridization was also performed to confirm the eradication of all four genomic AfuTmV-1.

7.2 Effects of dsRNA on growth of *A. fumigatus* Af293 NK isolate

7.2.1 Radial growth analysis

To determine the effect of AfuTmV-1 infection on *A. fumigatus* Af293 isolate, the virus-free isolates (NK105, 106 and 125) and AfuTmV-1 infected Af293 were investigated for mycelial growth rate on ACM and *Aspergillus* Minimal Medium (AMM) agar. Spore suspensions were prepared and suspended in sterile distilled water. Spores was counted and their numbers adjusted using a haemocytometer (Boeco, Germany) to obtain a concentration of 100 spores/ml. Spore suspensions (10 µl) of each isolate was pipetted onto the centre of the plate and incubated at 37 ˚C. Colony diameters were measured every 12 h over a period of 7 days. Colony growth rate and the morphology of the cultures were observed and compared with those of cured and infected strains. All experiments were performed in triplicate. Student's t test was used to analyze the data for differences in growth rate.

7.2.2 Biomass production in liquid media

Individual plates for each isolate were grown and conidia were harvested in a small quantity of water. An average of 10^6 spores was used as inoculum for 60 ml cultures of ACM and AMM broth. Mycelial weight was measured by harvesting the mycelia for each individual isolate from 3 replicate flasks by filtration through pre-weighted sterile Miracloth. Samples were taken every 24 h over a period of 7 days. Mycelial mats were dried at 37 ˚C for 3 days. Growth curves were plotted for each isolate.

---

8. *In vitro* replication of AfuTmV-1; adapted from Nogawa *et al*. (1996).

AfuTmV-1 VLP fraction and virus-free fraction were purified through the sucrose gradient centrifugation and CsCl gradient centrifugation as described in 2.4. The reaction mixture was prepared by mixing 100 µl of AfuTmV-1 with 200 µl of replication reaction mixture (100 mM Tris-HC1 pH 7.5, 10 mM magnesium acetate (SIGMA), 1 mM each of ATP, CTP, GTP and 0.5 mM DIG-labeled-11-UTP (RNA labeling mix, Roche, Germany), 200 units/ml RNase inhibitor (Promega), 50 µg/ml actinomycin D (Sigma), and 5 mM β-mercaptoethanol (Sigma). The mixture was incubated at 37 °C (optimal temperature for *A. fumigatus* growth) at 5, 10, 15, 30 min and overnight (16 hr). Total RNA was extracted using phenol/chloroform and precipitated with 0.1V of sodium acetate and 2.5V of cold ethanol at -20 °C overnight. Then total RNAs were precipitated with 0.1V of 4M LiCl and 2.5V of cold ethanol at -20 °C for 3 h to separate the synthesized ssRNAs from dsRNAs. The suspension was centrifuged at 13000xg for 15 min at 4°C. The pellet was washed with 70% cold ethanol (-20 °C) and then centrifuged at 13000xg for 5 min at 4°C. Two RNA fractions were electrophoresed in 1% agarose gel using TAE buffer. The gel was blotted onto a nylon membrane and then subjected to northern blot hybridization procedure according to the manufacturere's protocol. The hybridization signals were detected by enzymatic immunoassay.

Dot blot hybridization analysis

DIG-labelled ssRNAs synthesized during *in vitro* replication was also detected using dot blot hybridization. The labelled ssRNAs were precipitated using LiCl and then used as hybridization probes. Unlabeled ssRNAs corresponding to positive and negative strands of AfuTmV-1 dsRNAs (dsRNA1-4) were synthesized from four plasmids (pGDN9, pRNA2, p5'RNA3 and p5'RNA4) using MAXIScript®T7 Kit (Ambion®) *via* T7 and SP6 promoter, respectively. Approximately 500

ng of each ssRNA was blotted onto a nylon membrane (Amersham Hybond™-N membrane (GE healthcare)). The membrane was then subjected to northern blot hybridization procedure according to the manufacturere's protocol. The hybridization signals were detected by enzymatic immunoassay.

---

9. *In vitro* translation of AfuTmV-1 dsRNAs

Double stranded RNAs from the PEG-purified virus were extracted and subsequently treated with DNase and S1 nuclease prior to *in vitro* translation. Total dsRNAs were denatured at 95 ˚C for 5 min and immediately cooled on ice. Individual dsRNA segment was size-fractionated as previously described.

The translation was performed in the TNT® Coupled Reticulocyte Lysate System (Promega, USA) by adding 1 µg of dsRNA to 25 µl of the translation mixture (see table 6) incorporating with [^{35}S]methionine (PerkinElmer Cat.# NEG709A) and incubated at 30 ˚C for 2 h. An aliquot of 5 µl were taken every 30 min and mixed with 5 µl of 2 X SDS gel-loading buffers prior to heating at 100 ˚C for 5 min. Aliquots of 5 µl of denatured translation mixture were loaded in a NuPAGE® Novex® 4-12% Bis-Tris protein gel (Invitrogen) and electrophoresed in MOPS SDS running buffer for 45 min at 190 V (Bolt® mini gel tank). Precision plus protein™ all blue standard (BIO-RAD) was used as marker proteins. After electrophoresis the gel was fixed in fixing solution and agitated slowly on an orbital shaker (by dehydrating the gel in 1X dimethylsulphoxide (DMSO) for 20 min twice and soaked in DMSO containing 22% (w/v) 2,5-diphenyloxazole (PPO) overnight). The gel was rehydrated by immersion in milliQ water for a total of 30 min (10 min at a time) on a shaker and then rinsed with water several times. The washed gel was placed on two wet sheets of Whatman® 3MM filter paper, covered with plastic wrap and dried at 80 ˚C for 2 h using a vacuum gel dryer. The dried gel was exposed to Kodak® Biomax MR Film in the dark for 3 h to 10 days at room temperature.

Table 6. TNT® Reticulocyte Lysate reactions.

Component	Standard reaction
TNT® Rabbit Reticulocyte Lysate	12.5 µl
TNT® reaction buffer	1 µl
TNT® RNA polymerase (T7)	0.5 µl
Amino acid mixture, minus methionine (1mM)	0.5 µl

[^{35}S]methionine (>1000Ci/mmol at 10 mCi/ml)	1	µl
RNasin® ribonuclease inhibitor (40u/ml)	0.5	µl
DNA or RNA template(s) (1µg/µl)	1	µl
Nuclease-free water to a final volume of	25	µl

---

10. Transfection assay

10.1 Protoplast preparation

Protoplasts from *A. fumigatus* Af293 virus-free isolate (NK125) were prepared according to the protocol described by Szewczyk *et al*. (2006) with modifications.

Conidia suspension was prepared by harvesting a 3-day old *A. fumigatus* Af293 virus-free culture using 0.05% Tween 80. ACM liquid medium (50 ml) was inoculated with 1x10^8 conidia/ml and incubated on a rotary shaker (180 rpm) at 37 °C overnight (12-16 h). Conidia germination was checked under a microscopy by examining 10 µl of the overnight culture. Germinated spores were harvested by centrifugation at 3000 rpm for 8 min. During centrifugation 1.28 g of Novozyme234 (Novo Industries) or 0.4 g of lysing enzymes from *Trichoderma harzianum* (Glucanex®, SIGMA, L-1412) to replace Novozyme234 was added to 20 ml of osmotic medium (OM, pH 5.8). The enzyme solution was mixed vigorously and then sterilised by filtration (Nalgene™, Thermo Scientific). This solution was always prepared less than 30 min prior to use. Fungal tissue was completely resuspended in enzyme solution and then incubated with slow agitation of 100 rpm at 30 °C for 3 h (Note that higher temperature may result in lysis of the protoplasts). The digestion mixture was gently transferred to a Falcon tube and then overlayed with 5 ml of protoplast trapping buffer by leaning the tube at 45 degree angle and slowly adding 1 ml of the buffer down the side of the tube at a time without disrupting the OM buffer layer. The layer was centrifuged at 5000 rpm for 15 min using the sealed swinging bucket rotor. The top cloudy layer containing protoplasts was removed and put into a 50 ml centrifuge bottles (Polycarbonate with cap, 29x104 mm, #357002, Beckman) on ice. Cold STC solution (15 ml) was gently added to 5 ml of protoplast solution on ice and then centrifuged at 6000 rpm for 8 min at 4 °C (Beckman Coulter Avanti-J-26xP, rotor ID JA 25.50 fixed rotor). The supernatant was poured off carefully, leaving the whitish strip of protoplasts along the bottom/side of the tube. The protoplast pellets were resuspended in 0.5 ml of cold STC on ice and then centrifuged at 13,000 RPM for 30 sec. The protoplasts were resuspended in 120 µl of cold STC, yielding *ca*. 10^7 protoplasts/ml.

10.2 Transfection assay

PEG-purified AfuTmV-1 VLPs, LiCl-purified dsRNAs and Proteinase K+S1 treated-dsRNA samples were prepared as described previously and used to individually transfect *A. fumigatus* Af293 virus-free isolate (NK125) protoplasts. For the Proteinase K treated-sample, dsRNAs were treated with the enzyme as shown below prior to S1 nuclease treatment;

Proteinase K treatment:-

All potentially contaminating proteins in AfuTmV-1 dsRNA samples were removed following treatment with Proteinase K. The dsRNAs (10 mg) mixed with 1% SDS were denaturated by heating at 95 ˚C for 15 min and then allowed to cool. A 15 mg/ml (100x) concentrated stock of Proteinase K (from *Tritirachium album*, SIGMA) was prepared by dissolving in 50 mM Tris-HCl (pH 8.0) and 1 mM $CaCl_2$. For protein digestion, a 0.01 volume of Proteinase K stock was added to the reaction to a final concentration of 0.15 mg/ml. The reaction was then incubated at 37 ˚C for at least 1-2 h and dsRNAs isolated following phenol/chloroform extraction and ethanol precipitation

Approximately 30 µg (10 µl) of VLPs or dsRNAs were added to 100 µl of protoplasts and vortexed 7-8 times for 1 sec to mix the VLPs/dsRNAs with the protoplasts throroughly, achieving a final concentration of 0.2 µg/µl RNA+protoplast solution. An aliquot of 20 µl of protoplasts without RNA was used as a control. Freshly filtered 60% PEG 4000 solution (50 µl) at room temperature was added to the mixture and the whole vortexed for 4-5 times for 1 sec. The mixture was incubated on ice for 50 min and then transferred into 2 ml microcentrifuge tube. Then further 60% PEG 4000 solution (1 ml, room temperature) was added and the whole mixed well by pipetting 10 times, followed by incubation at room temperature for 25 min (Note that PEG solution should be shaken and filter steriled immediately before use). Then STC solution (4 ml) was added to the protoplast transfection mixture, facilitating plating. A solution of 200 µl was plated gently onto the SMM regeneration and ACM plates. Plates were sealed using parafilm and incubated right-side up overnight at 30 ˚C and then inverted the next day and kept at 37 ˚C for 1-2 days after transfectant colonies formed. Colonies were harvested and transferred to fresh ACM medium and incubated at 37 °C for 7 days on a rotary shaker (150 rpm). During cultivation, cultures were pelleted and re-inoculated into a new ACM medium every 7 days for two more generations.

The presence of the virus was investigated by purification of the transfected fungal mass using RNeasy® Plant Mini Kit (QIAGEN) or LiCl extraction procedure and then visualized the dsRNA

bands on 1% gels. RT-PCR amplification of viral dsRNA using a gene specific primer pair (GAP2 primers) designed based on the dsRNA2 sequence was also performed to confirm successful virus infection following transfection.

---

11. Identification of *SINEs* in *A. fumigatus* Af293 genome

11.1 Sequence dataset

Aspergillus fumigatus Af293 reference genome sequences (sequenced by Nierman *et al*., 2005) were available through the NCBI database. The accession numbers of sequence resources are shown in Table 7. The information of *A. fumigatus* Af293 genome was also retrieved and analysed from other sources available online such as The Central *Aspergillus* Resource (CADRE; www.cadre-genomes.org.uk, Mabey *et al*., 2004), BROAD INSTITUTE *Aspergillus* comparative database (http://fungicyc.broadinstitute.org/AF2/organism-summary?object=AF2) and the *Aspergillus* Genome Database (*Asp*GD; http://www.aspgd.org, Arnaud *et al*., 2010).

Table 7. Accession numbers of sequence resources on NCBI database.

Name	RefSeq	GenBank	Chromosome length (bp)
Chromosome 1	NC_007194.1	CM000169.1	4,918,979
Chromosome 2	NC_007195.1	CM000170.1	4,844,372
Chromosome 3	NC_007195.1	CM000171.1	4,079,167
Chromosome 4	NC_007197.1	CM000172.1	3,927,224
Chromosome 5	NC_007198.1	CM000173.1	3,948,411
Chromosome 6	NC_007199.1	CM000174.1	3,778,736
Chromosome 7	NC_007200.1	CM000175.1	2,058,334
Chromosome 8	NC_007201.1	CM000176.1	1,833,124

11.2 Search strategy and computational analyses

CENSOR software developed by the Genetic Information Research Institute (GIRI; http://www.girinst.org/censor; Kohany *et al*., 2006) was used to interrogate the *A. fumigatus* Af293 genomic sequences in order to screen for putative *SINEs* against a reference collection of repeats. The homologous portions of the query and reference sequences were masked and reports classifying all found repeats were generated. RepeatMasker server (www.repeatmasker.org; Smit *et al*., 2013-2015), which screens DNA sequences for interspersed repeats, was also used to screen the Af293 genomic sequences against reference sequences on the Repbase libraries (Jurka, 2000). The reported sequences which were

classified as non-long terminal repeat (non-LTR); *SINEs*, *SINE2*/tRNA and *SINE3*/5S, were selected as *SINE* candidates and then subjected to further analysis.

To identify the tRNA-related *SINEs*, the genomic sequences were submitted to the tRNAScan-SE program (http://lowelab.ucsc.edu/tRNAscan-SE; Lawe and Eddy, 1997) to search for the presence of basic tRNA features such as RNA polymerase III promoter sequences and tRNA cloverleaf secondary structures using relaxed parameters and EuFindtRNA algorithm. Subsequently, the sequences predicted from the programme were resubmitted to tRNAScan-SE using default settings. Sequences yielding positive reads were discarded as they are considered as true tRNAs while the remaining sequences yielding negative reads which showed no support for 5'tRNA-related region were considered for further analyses. The tRNA cloverleaf secondary structures and tRNA origin of the remaining sequences were also identified from the programme. In addition, Genomic tRNA Database against the *A. fumigatus* tRNAs (GtRNAdb; http://gtrnadb.ucsc.edu/Aspe_fumi) was also used to separate tRNAs from the candidate *SINEs*.

All predicted *SINE* sequences retrieved from Censor, RepeatMasker scan and tRNAScan-SE were submitted to BLAST analysis using sequence databases accessible from the National Center for Biotechnology Information (NCBI) server (www.ncbi.nlm.nih.gov/BLAST) and the *A. fumigatus* Genome Map Viewer (http://www.ncbi.nlm.nih.gov/mapview) to expand the boundaries of the masked sequences to 1000 bp upstream and 2000 bp downstream of both termini.

The sequences showing predicted tRNA-related structures were manually inspected for the presence of degenerate RNA polymerase III promoter A and B box sequences, extended 3' tRNA-unrelated sequence before an oligothymidine tract and also target site duplication sequences flanking the *SINEs*. An additional check for the presence of the RNA polymerase III promoter sequence was performed by multiple sequence alignment of the predicted *SINEs* to the *A. fumigatus* Af293 tRNA gene sequences obtained from the Genomic tRNA Database. To identify the 5S rRNA-related *SINEs*, sequences classified as *SINE3*/5S using Censor and RepeatMasker were aligned against the *A. fumigatus* 5S rRNAs and searched for the RNA polymerase III promoter.

Characterization, classification, distribution and location of *SINEs* were performed by comparative analysis of the sequences. Multiple sequence alignments were performed using Clustal Omega and MAFFT alignment and manual edition of the sequences. Phylogenetic

analyses were performed using the neighbor-joining (NJ) method in MAFFT alignment. Sequence divergence was calculated using DIVEN (Deng *et a*l., 2010; http://indra.mullins.microbiol.washington.edu/DIVEIN/).

11.3 PCR amplification of *A. fumigatus SINEs*

Sequence specific oligonucleotide primers were designed based on the sequences of candidate *SINEs* to cover A and B box regions, yielding approximately 100-150 bp amplicons. A list of the oligonucleotide primers used in this study is shown in the Table 8.

Table 8. Oligonucleotide primers used for PCR amplification of *A. fumigatus* Af293 *SINEs*.

AfuSINEs	**Primers**	**Primer sequences**	**Product size (bp)**
AfuSINE3-1a	SINE3F SINE3R1a	5'-GACCATAGGGTGTGGAAAAC -3' 5'-CCAGCACTCATACGAAAAGCA-3'	132
AfuSINE3-3a	SINE3F SINE3R3a	5'-GACCATAGGGTGTGGAAAAC-3' 5'-CCCTATGGTTTAGCGTCTCTCA-3'	133
AfuSINE3-3c	SINE3F SINE3R3c	5'-GACCATAGGGTGTGGAAAAC-3' 5'-CCCGTATCACCAGCACTCAT-3'	144
AfuSINE3-4a	SINE3F SINE3R4a	5'-GACCATAGGGTGTGGAAAAC-3' 5'-CCCTATGGTTCAGTGTCTCTCA-3'	133
AfuSINE3-5c	SINE3F SINE3R5c	5'-GACCATAGGGTGTGGAAAAC-3' 5'-ATCTGTAAAGTAGCAGGCCTTT-3'	112
AfuSINE2-1a	SINE2F1a SINE2R1a	5'-TAGTGGTAAGCGCTCCGAG-3' 5'-GGTAAATGTTGTGCCAGCGA-3'	102
AfuSINE2-3a	SINE2F3a SINE2R3a	5'-CTCTGGGGTCGGTCGTTAAG-3' 5'-GTCGGCTTTCAACATACTCACA-3'	120
AfuSINE2-4a	SINE2F4a SINE2R4a	5'-CCTGGGCAGAAATATGATGGG-3' 5'-CCCAAAGGCACCAAATTCCT-3'	110
AfuSINE2-7a	SINE2F7a SINE2R7a	5'-GCGGAATGGAAAGGTTCGAA-3' 5'-GAATAGAATGCGGCCCACTG-3'	139
AfuSINE2-5d	SINE2F5d SINE2R5d	5'-GAGGATGGTCGGAAAATCGC-3' 5'-ACCTAATACTCTCAGCTCTCCC-3'	101
AfuSINE2-3c	SINE2F3c SINE2R3c	5'-GTTGTCTCGGCACAGATGG-3' 5'-TCACTGGGTCTTTTGGCTAGA-3'	134
AfuSINE2-4c	SINE2F4c SINE2R4c	5'-GGGCAGAATCAACAACGTGT-3' 5'-TTGGTAAGTGTTTCGGTGGC-3'	123
AfuSINE2-7e	SINE2F7e SINE2R7e	5'-ACTTATATAGCAGAGTGGTA-3' 5'-TCCTATAGCTTCCCTTAAAG-3'	123
β-tubulin	TUB-For TUB-Rev	5'-CGACGCGAGACTGTGTTAAG-3' 5'-TAGTTGGTGCTCAAGGATGG-3'	440 (PCR) 210 (RT-PCR)

A. fumigatus Af293 genomic DNA was prepared using the DNeasy® Plant Mini Kit (QIAGEN) with mini protocol provided by the manufacturer. PCR amplification was performed in a 50 µl reaction mixture containing 5x GoTaq® Reaction Buffer, 10 mM dNTPs mix, forward and reverse primers (100 µM),RNase H and GoTaq® DNA polymerase (5 U/µl).

High stringency of PCR conditions was employed as followed:-

Initial denaturation	94 ˚C	5 min	1 cycle
Denaturation	94 ˚C	30 sec	35 cycles
Annealing	55 or 60 ˚C	30 sec	
Extension	72 ˚C	30 sec	
Final extension	72 ˚C	5 min	1 cycle
Hold	4 ˚C		1 cycle

PCR products were purified using the QIAquick® PCR Purification kit (QIAGEN) and then separated by gel electrophoresis and visualized with SYBR safe DNA gel stain.

11.4 RT-PCR amplification of *A. fumigatus SINEs*

A. fumigatus Af293 total RNA was prepared using the RNeasy® Plant Mini Kit (QIAGEN) with a mini protocol provided by the manufacturer. Total RNA was then purified to eliminate the genomic DNA using TURBO™ DNase (Ambion) and DNase inactivation reagent (Ambion ® DNA-free™ DNase Kit). The same reactions and PCR condition were used for RT-PCR amplification of *A. fumigatus* Af293 total RNA with the same primer pairs (Table 8).

The cDNA for RT-PCR of *SINEs* and β-tubulin transcripts was synthesized using 10 µl of total RNA (*ca.* 50 ng/µl) using random hexamers and Superscript III reverse transcriptase (Invitrogen, UK) following the manufacturer's protocol. The reaction was prepared as following:-

10 µl	total RNA
0.5 µl	sequence specific primer (500 µg/ml)
1 µl	10 mM dNTPs mix (Promega)
1.5 µl	DEPC-H_2O

The reaction was incubated at 65 ˚C for 10 min and then snap cooled on ice for 2 min. Then the reaction mix for cDNA synthesis was added as following:-

4	µl	5x First-strand buffer (Invitrogen)
1	µl	0.1 M DTT (Invitrogen)
1	µl	RNasin® Ribonuclease Inhibitor (Promega)

SuperScript™ III reverse transcriptase (1 µl of 200 U/µl) was added to the mixture. The reaction tube was briefly centrifuged to collect the contents and then the whole reaction was incubated at 25˚C for 5 min, followed by 50 ˚C for 60 min (the reaction should be subjected to the temperature of 55˚C if a sequence specific primer was used), followed by the incubation at 70 ˚C for 15 min to inactivate the enzyme. The quantity of cDNA was check by Nanodrop reading (ssDNA-33).

RT-PCR was performed in a 50 µl reaction mixtures in thin-walled 0.5 ml PCR Eppendorf tubes and contained the following components:-

5	µl	cDNA (from first-strand reaction)
29.6	µl	DEPC-H_2O
10	µl	5x GoTaq® Reaction Buffer (Promega)
2	µl	10 mM dNTPs mix (Promega)
1	µl	Sequence specific-Forward primer (100 µM)
1	µl	Sequence specific-Reverse primer (100 µM)
1	µl	RNase H (NEB)
0.4	µl	GoTaq® DNA polymerase (5 units/ µl; Promega)

Tubes were then placed in the thermal cycler and the same PCR condition regime used for *SINE* amplification from genomic DNA was employed. PCR products were purified using the QIAquick® PCR Purification kit (QIAGEN) and then separated by gel electrophoresis and visualised as before.

11.5 Southern blot hybridization of *A. fumigatus SINEs*

Target DNA was prepared from the *A. fumigatus* Af293 strain. Aliquots of 10 µg of gDNA were individually digested with 1 µl of restriction enzymes (*Hind*III (New England Biolabs)) which does not cut *SINE* elements. Digestion reactions were incubated at 37 °C overnight.

The gDNA (10 ng) was also used as DNA template for probe preparation labeled with PCR DIG probe synthesis kit (Roche, UK) according to the manufacturer's instructions. Gene specific primers were designed to generate the PCR probes under the PCR cycling conditions as described in Table 9.

Table 9. Gene specific primers, reaction set-up and PCR conditions for PCR probe synthesis.

Primers used for PCR probe synthesis:

Primers	Primer sequences	Product	Product size (bp)
See Table 8			
AfuFKS-F AfuFKS-R	5'-GCCTGGTAGTGAAGCTGAGCGT-3' 5'-CGGTGAATGTAGGCATGTTGTCC-3'	1,3-beta-glucan synthase /Afu6g12400	101

PCR reaction set-up:

Component	DIG labeled probe (µl)	Unlabeled DNA control (µl)	Labeled kit control (µl)	Final conc.
dH_2O	28.25	28.25	29.25	-
10xPCR buffer with $MgCl_2$	5	5	5	1x
PCR DIG labeling mix	5	-	5	0.2mM
dNTP stock solution	-	5	-	0.2 mM
Upstream primer	5	5	5	1µM each
Downstream primer	5	5	-	-
Enzyme mix	0.75	0.75	0.75	2.6U
Template DNA	1	1	5	10 ng
Final volume	**50**	**50**	**50**	

PCR cycling conditions: (SINE_SB)

	Cycles	Time	Temp (°C)
Initial denaturation	1	2 min	95
Denaturation	1-10	10 sec	95
Annealing		30 sec	60
Elongation		2 min	72
Denaturation	11-30	10 sec	95
Annealing		30 sec	60
Elongation		2 min	72
Final elongation	1	7 min	72
Cooling		forever	4

Restriction fragments of gDNA (5µg/lane) were electrophoresed on 1% w/v agarose gel containing SYBR safe DNA gel stain in 1XTAE. Electrophoresis was performed applying a voltage of 60 V for 5 h until the DNA was well-separated and then visualised under a blue-light

transilluminator. Gels were denatured by soaking in 0.25 N HCl for 10 min, then immersed 0.5M NaOH + 1.5 M NaCl for 15 min twice, followed by neutralization in 0.5 M Tris-HCl (pH 7.0) + 1.5 M NaCl for 15 min twice. DNA was transferred onto a positively charged Amersham Hybond^{TM}-N membrane (GE healthcare) by running the gel cassette overnight with a voltage setting of 15V, 0.2 A at 4 ˚C. Upon completion of the run, the membrane was fixed by UV-cross linking for 25 min at the energy level of 2000x100 μJ/cm^3.

DIG Easy Hyb (10 ml) was prewarmed to hybridization temperature at 50 ˚C for 20 min. Then, the membrane was placed in pre-warmed DIG Easy Hyb and pre-hybridised at 42 ˚C for 30 min. DIG-labeled PCR probe (10μl) was denatured by boiling at 95 ˚C for 5 min, rapidly cooled in ice and then added to a prewarmed DIG Easy Hyb (5 ml). The hybridisation tube was incubated overnight at 42 ˚C with gentle agitation. After hybridisation, the membrane was washed using low stringency buffer (2xSSC, 0.1% SDS) twice for 5 min at room temperature, followed by a high stringency wash with a prewarmed 0.1xSSC, 0.1% SDS solution at 68 ˚C for 15 min twice. After hybridization, low stringency and stringency washes, an immunological detection was performed as described previously in northern blot hybridization using reagents provided from the manufacturer (DIG Northern Starter Kit, Roche). The membrane was exposed to a phosphorimager (FujiFilm LAS-3000 Imaging System) for 20-30 min at -25 ˚C.

11.6 Detection of small RNA molecules homologous to *A. fumigatus SINEs*

To generate probes for small RNA hybridization, *A. fumigatus SINEs* and β-tubulin DNA were amplified using oligonucleotide primers shown in Table 8. Then, purified PCR products (QIAquick® PCR Purification kit, QIAGEN) were cloned into pGEM®-T Easy (Promega, Madison, WI, USA) according to the manufacturer's instruction. Insert orientation and integrity of plasmid pSINEs (for *SINEs*) and pTUB (for β-tubulin) was confirmed by DNA sequencing. Sense and antisense *SINE*-specific, and β-tubulin RNA probes were synthesized with T7 and SP6 RNA polymerases using DIG Northern Starter Kit (Roche). Prior to hybridization to small RNAs, probes were cleaved to an average length of 50 bp by alkaline hydrolysis as described by Kreuze *et al*. (2005).

11.6.1 Isolation of small RNA (Low molecular weight (LMW) RNAs)

A protocol for small RNA isolation was taken from Lu *et al*. (2007) with some modifications. Small RNAs were isolated as follow;

First, 3 g of frozen mycelia was ground to fine powder under liquid nitrogen. Then total RNAs were isolated using Trizol reagent as indicated in the manufacturer's protocol. In this case, 40 ml of Trizol was used for 3 g of mycelia, yielding *ca.* 500 µg of total RNA. Total RNAs were dissolved in DEPC-treat H_2O to a concentration of 1 µg/µl (Note: *ca*. 300-400 µg of total RNAs were used for next steps). Then, mRNA and rRNA (high molecular weight; HMW RNAs) were precipitated by adding 50% PEG (MW=8000) to a final concentration of 10% and 5M NaCl to a final concentration of 0.5M. The solution was mixed and quenched on ice for 1 h, then centrifuged at maximum speed at 4 °C for 10 min to pellet HMW RNAs. Supernatant fractions containing LMW RNAs was transferred to a new microcentrifuge tube and precipitated with 2.5 vols of 100% EtOH. The solution was mixed and kept at -20 °C for at least 2 h, then centrifuged for at least 30 min at 4 °C to pellet LMW RNAs. The supernatant was removed carefully. Then the pellet was washed with 80% EtOH, air-dried and dissolved in DEPC-treat H_2O (Note: 10 µl DEPC-treat H_2O was used to resuspend LMW RNAs from 100 µg of total RNAs).

11.6.2 Preparation of small RNAs by polyacrylamide gel electrophoresis

For the detection of *SINE* homologous small RNAs, 100 µg of the small RNA fraction was separated in 15% (w/v) polyacrylamide Tris-borate-EDTA-urea gels (Bio-Rad, Sweden). Polyacrylamide gel was prepared as followed;

Gel mix: 9.6 g urea, 7.5 ml 40% acrylamide/bis-acrylamide stock solution, 2 ml 10xTBE and DEPC-treated H_2O to 20 ml. The solution was mixed using stirrer and simultaneously warmed at 37 ˚C to dissolve urea and then cooled to room temperature. Freshly prepared solution of 10% ammonium persulfate (APS; 120 µl) was added to the gel solution and mixed well. Then, 9.2 µl of TEMED was added and mixed by swirling. The solution was poured into the glass plates and allowed to polymerize for 45 min at room temperature. After removing the comb, wells were rinsed with 1xTBE. The gel was pre-run for 15-30 min at 200 V and the wells were washed again using 1xTBE.

For denaturing polyacrylamide-urea PAGE of small gel (25 ml) with different concentration can be prepared as followed;

Components	6%	10%	15%
10x TBE Urea 40% Acrylamide ddH_2O TEMED 10% APS	2.5 ml 10.5 g 3.75 ml 11.25 ml 25 µl 125 µl	2.5 ml 10.5 g 6.25 ml 8.75 ml 25 µl 125 µl	2.5 ml 10.5 g 9.375 ml 5.625 ml 25 µl 125 µl
Total volume	25 ml	25 ml	25 ml

10X TBE buffer: 108 g Tris base
55 g Boric acid
40 ml 0.5M EDTA (pH 8.0)
Add H_20 to 1 Liter

For sample loading, 10 µl (*ca*. 100 µg) of LMW RNAs was added with 10 µl of 2x loading dye. The RNA solution was mixed by vortexing and heated at 65 °C for 5-10 min and immediately kept on ice. A 3 µl aliquot of heat denatured 10-100 bp Low Molecular Weight Marker (Affymetrix) mixed with 3 µl of 2x loading dye was loaded as a marker. The gel was run in 1xTBE at 200 V for 1-1.5 h (the first dye can go until 2cm from the bottom) and stained with 1xTBE/SYBR Gold RNA gel stain (Invitrogen^{TM}) (1 µg/ml) for 5-10 min.

For northern blot hybridization of small RNA, nucleic acids were transferred to nylon membranes. The RNA was fixed to the membrane by 1-ethyl-3-(3-dimethylaminopropryl carbodiimide); EDC-cross linking for 2 h at 60°C (Pall *et al*., 2007). Then membranes were prehybridized and hybridized as described previously. Hybridized membranes were exposed to a phosphorimager as before.

11.7 Secondary structure prediction of *A. fumigatus SINE* transcripts

Secondary structures of *SINEs* were predicted using Mfold (http://mfold.rna.albany.edu; Zuker 2003) with default settings. The folding temperature was set to 37 °C according to the temperature used for growth of the *A. fumigatus* Af293 cultures. The RNA structure showing the lowest free energy was accepted as the most likely structure for *SINEs*.

---

12. Silencing vector construction

12.1 Construction of silencing vectors based on AfuTmV-1 dsRNA mycovirus

AfuTmV-1 dsRNA2 was selected to construct mycovirus-based silencing vectors. Initially, a full length cDNA of dsRNA2 was generated by RLM-RACE amplification using internal oligonucleotide primers based on a unique *Bam*HI sequence in dsRNA2 (BamF and BamR primers) in combination with Lig-For and Lig-Rev primers, yielding two PCR fragments which are 1195 and 1130 bp in size and designated as amplicons A and B, respectively. Amplicons A and B were re-amplified using internal BamF and BamR primers in combination with respectively 5- and 3'- primers (RNA2FS and RNA2RS) to incorporate T7 RNA promoter sequences and restriction enzyme sites (*Hind*III, *Xba*I, *Eco*RI and *Kpn*I) on both ends for subsequent *in vitro* transcription. Primers used for construction of the vectors are shown in Table 10. PCR amplification was performed in a 50 µl reaction mixtures using proof reading Phusion® High-fidelity DNA polymerase (NEB). The reaction mixture and PCR conditions were set as follows;

5 µl	cDNA (from first-strand reaction)
28.5 µl	DEPC-H_2O
10 µl	5x Phusion HF Buffer (NEB)
1 µl	10 mM dNTPs mix (Promega)
2.5 µl	Forward primer (10 µM)
2.5 µl	Reverse primer (10 µM)
0.5 µl	Phusion® DNA polymerase (5 units/ µl; NEB)
50 µl	reaction mixture

Tubes were then placed in the thermal cycler and the following thermal cycling PCR amplification regime employed, specifically for Phusion® DNA polymerase:-

Initial denaturation	98 ˚C	30 sec	1 cycle
Denaturation	98 ˚C	10 sec	35 cycles
Annealing	53-72 ˚C	30 sec	
Extension	72 ˚C	30 sec	
Final extension	72 ˚C	10 min	1 cycle
Hold	4 ˚C	forever	1 cycle

PCR generated amplicons A and B were purified using the QIAquick PCR purification kit (QIAGEN) and respectively restricted with *Hind*III+*Bam*HI and *Bam*HI+*Kpn*I and then ligated in a three-way ligation into pUC18 vector (Fermentas) doubly-restricted with *Hind*III and *Kpn*I, yielding pAfuTmV-1-RNA2 vector. Representative recombinant plasmids were digested with restriction enzymes to confirm successful cloning of the insert in the correct orientation. Plasmids were also sequenced at both ends and the central region to confirm that it contained the cDNA sequence of dsRNA2 including the reconstructed *Bam*HI site and the terminal T7 RNA promoter sequences.

Table 10. Primers used for construction of silencing vectors based on AfuTmV-1 dsRNA2.

Primers	Primer sequences	Primer pairs	Annealing temp (°C)	Product size (bp)
T7	5´- **TAATACGACTCACTATAGGG**- 3´			
RNA2FS	5'CCCAAGCTTTCTAGA**TAATACGACTCACTATAGGG**TGAATAAAAGATTTGTTCTG -3'	RNA2FS+ BAMR	52	1195
RNA2RS	5'CCCGAATTCGGTACC**TAATACGACTCACTATAGGG**CCCCCAGGCCCGGCGGGGGG -3'	RNA2RS+ BAMF	70	1130
BAMF	5'- CTCAACAAGGATCCCATGCTCG -3'			
BAMR	5'- CGAGCATGGGATCCTTGTTGAG -3'			
ALB-NgoF	5'CCCCGCCGGCGCGGTTGGATCGAAAGGT3'	ALB-NgoF+ALB-NgoR	65	550
ALB-NgoR	5'CCCCGCCGGCGTTGCATTGGGAGCGGCC3'			

A truncated version of the full-length clone was constructed using two restriction sites for *Ngo*MIV (NEB) which deleted an internal fragment, leaving an internal sequence and the 5'- and 3'-untranslated regions. The resulting plasmid was designated as pAfuTmV-1-NgoMIV construct. This plasmid was used as a prototype vector for silencing vector construction.

Plasmid vector, pAfuTmV-1-NgoMIV, was subsequently modified to express the transcripts of the engineered cDNA of dsRNA2 containing *ALB1/PKSP* fragment in both sense and anti-sense orientations. Initially, the targeting *ALB1/PKSP* fragment was amplified from *A. fumigatus* Af293 genomic DNA using primers containing *Ngo*MIV sequence as described in Table 10 (Note the that size of the targeting gene to be silenced can be 300-1500 bp which locates on the exon region). The pAfuTmV-1-NgoMIV plasmid and *ALB1/PKSP* PCR products were then digested with *Ngo*MIV, purified and ligated, resulting in a pAfuTmV-1-ALB1 construct.

Each recombinant plasmid were introduced into *E. coli* JM 109 competent cells and grown on LB plates containing 100 µg/ ml ampicillin/ 200 mg/ml IPTG/ X-Gal plates for blue/white colony selection. Vector integrity and insert orientation were confirmed by sequencing analysis at the modification sites and insert junctions of each construct.

12.2 Construction of silencing vectors based on a *SINE* element using fusion PCR

Fusion PCR allows for the construction of silencing vectors using transcriptional fusion of a target gene to a *SINE* element. Initial PCRs were separated for amplification of *SINE* and *ALB1/PKSP* sequences from *A. fumigatus* Af293 genomic DNA using primers listed in Table 11, whereas the second PCR was employed to fuse the two fragments into a single molecule using nested primers listed in table 12.

Table 11. Primers used for construction of silencing vectors based on *AfuSINE2-1a*.

Constructs	Target sequences	Orientation	Primers	Primer sequences
pAF1-ALB1	*AfuSINE2-1a*	Sense	1-SINE1aPmeI-F:	CCCGTTTAAACCAAGTGTACATAGAGC
			2-SINE1b-R:	GAGACCTTTCGATCCAACCGCCCTCTATGTACACTTAGTAG
	ALB1/FKSP	Sense	3-SINE1c-F:	CTACTAAGTGTACATAGAGGGCGGTTGGATCGAAAGGTCTC
			4-ALB1aPmeI-R:	GGGGTTTAAACGTTGCATTGGGAGCGGCC
pAF2-ALB1	*ALB1/FKSP*	Sense	5-ALB1bPmeI-F:	CCCGTTTAAACGCGGTTGGATCGAAAGGT
			6-SINE1d-R:	ACCGGCTCTATGTACACTTGGTTGCATTGGGAGCGGCCG
	AfuSINE2-1a	Sense	7-SINE1e-F:	CGGCCGCTCCCAATGCAACCAAGTGTACATAGAGCCGGT
			8-SINE1fPmeI-R:	GGGGTTTAAACCCTCTATGTACACTTA
pAF3-ALB1	*ALB1/FKSP*	Sense	5-ALB1bPmeI-F:	CCCGTTTAAACGCGGTTGGATCGAAAGGT
			9-ALB1d-R:	CAGCTCCTCGCCCTTGCTCACCATGTTGCATTGGGAGCGGCC
	GFP	Sense	10-GFP1a-F:	GGCCGCTCCCAATGCAACATGGTGAGCAAGGGCGAGGAGCTG
			11-GFP1b-R:	GGGATCGATAAGTCGTGCTGCTTCATG
	PSKP/1BLA	Anti-sense	12-ALB1eClaI-F:	CCCATCGATCCCGTTGCATTGGGAGCGGCC
			13-ALB1fPmeI-R:	GGGGTTTAAACGCGGTTGGATCGAAAGGT
		NOTE:	*Pme*I site	= underlined
			*Cla*I site	= underlined and italic
			AfuSINE2-1a sequence	= shaded in a grey box
			ALB1/FKSP sequence	= shaded in a pink box

GFP sequence = shaded in a light green box

Table 12. Primers and annealing temperatures used for PCR amplification of each construct.

Constructs	Primer pairs	Annealing temp (˚C)	PCR products	Product size (bp)
pAF1-ALB1	1-SINE1aPmeI-F 2-SINE1b-R	30	*AfuSINE2-1a* = 1A	383
	3-SINE1c-F 4-ALB1aPmeI-R	69	*ALB1/FKSP* = 1B	531
	1-SINE1aPmeI-F 4-ALB1aPmeI-R	30	*AfuSINE2-1a*+ *ALB1/FKSP* = 1C	894
pAF2-ALB1	5-ALB1bPmeI-F 6-SINE1d-R	65	*ALB1/FKSP* = 2A	526
	7-SINE1e-F 8-SINE1fPmeI-R	38	*AfuSINE2-1a* = 2B	401
	5-ALB1bPmeI-F 8-SINE1fPmeI-R	38	*ALB1/FKSP*+*AfuSINE2-1a* = 2C	894
pAF3-ALB1	5-ALB1bPmeI-F 9-ALB1d-R	65	*ALB1/FKSP* = 3A	535
	10-GFP1a-F 11-GFP1b-R	60	GFP = 3B	288
	5-ALB1bPmeI-F 11-GFP1b-R	45	*ALB1/FKSP*+GFP = 3C	771
	12-ALB1eClaI-F 13-ALB1fPmeI-R	65	*PSKP/1BLA* = 3D	519

To synthesize individual and fusion PCR products, PCR amplification of the *A. fumigatus* Af293 genomic DNA was performed in a 50 µl reaction mixture using GoTaq® DNA polymerase (5 U/µl; Promega). The reaction mixture was set as described in the random PCR (rPCR) protocol (see section 3). For PCR conditions, each fragment was amplified using the PCR conditions as described in the PCR amplification using Phusion High-fidelity DNA polymerase (see section 12.1) with different annealing temperature for each primer pair.

Detailed strategy to construct each *SINE*-based silencing plasmid is as followed;

Construct of pAF1-ALB1

A portion of *AfuSINE2-1a* was amplified from *A. fumigatus* af293 genomic DNA using primers 1-SINE1aPmeI-F and 1-SINE1aPmeI-F to add a *Pme*I restriction site at the 5' end and a portion of *ALB1* at the 3' end. A portion of the coding sequence of *ALB1* from nucleotides 704 to 1203 was PCR amplified using oligonucleotide primers 3-SINE1c-F and 4-ALB1aPmeI-R to add a portion of the *AfuSINE2-1a* sequence at the 5' end and a *Pme*I restriction site at the 3' end. The two PCR products were then amplified using primers 1-SINE1aPmeI-F and 4-ALB1aPmeI-R to amplify an 894 bp fragment. The resulting product was digested with *Pme*I and cloned using standard methods (Sambrook *et al*. 1989) into the corresponding site in pAF1-ALB1, yielding an *AfuSINE2-1a-ALB1/PKSP* transcriptional fusion in the sense orientation (*ᵖgpdA::AfuSINE1::ALB1 /PKSP:: his2Aᵗ* cassette).

Construct of pAF2-ALB1

A portion of the coding sequence of *ALB1* from nucleotides 704 to 1203 was PCR amplified from *A. fumigatus* af293 genomic DNA using oligonucleotide primers 5-ALB1bPmeI-F and 6-SINE1d-R to add a *Pme*I restriction site at the 5' end and a portion of the *AfuSINE2-1a* sequence at the 3' end. A segment of *AfuSINE2-1a* was amplified using primers 7-SINE1e-F and 8-SINE1fPmeI-R to add a portion of *ALB1* at the 5' end and a *Pme*I restriction site at the 3' end. The two PCR products were then amplified using primers 5-ALB1bPmeI-F and 8-SINE1fPmeI-R to amplify an 894 bp fragment. The resulting product was digested with *Pme*I and cloned into the corresponding site in pAF2-ALB1, yielding an *ALB1/PKSP*-afuSINE1 transcriptional fusion in the sense orientation (*ᵖgpdA::ALB1 /PKSP::AfuSINE1:: his2Aᵗ* cassette).

Construct of pAF3-ALB1

This construct was designed as an inverted repeat of 500 bp of the *ALB1/PKSP* gene coding sequence separated by a 250 bp spacer fragment of green fluorescent protein (GFP) sequence into the background vector. A portion of the coding sequence of *ALB1* from nucleotides 704 to 1203 was PCR amplified from *A. fumigatus* af293 genomic DNA using oligonucleotide primers 5-ALB1bPmeI-F and 9-ALB1d-R to add a *Pme*I restriction site at the 5' end and a portion of GFP sequence at the 3' end (product A). A portion of the GFP coding sequence was PCR amplified from plasmid pEAQ-HT-GFP (Sainsbury *et al*., 2009; Peyret and Lomonossoff, 2013) using primers 10-GFP1a-F and 11-GFP1b-R to add a portion of *ALB1* sequence at the 5' end and an *Cla*I restriction site at the 3' end (product B). PCR products A and B were then used as a template for fusion PCR with primers 5-ALB1bPmeI-F and 11-GFP1b-R to amplify a 771 bp

fragment (Product C). This was then digested with *Pme*I and *Cla*I and ligated into the digested plasmid pSK379. The antisense *ALB1* was PCR amplified using primers 12-ALB1eClaI-F and 13-ALB1fPmeI-R to incorporate a *Cla*I site at the 5' end and a *Pme*I restriction site at the 3' end and ligated into the appropriate sites of plasmid containing the *ALB1::GFP* sense segment, yielding pAF3-ALB1 (*pgpdA::ALB1/PKSP::GFP::PSKP/1BLA::his2A*^{t} cassette).

12.3 Transfection and transformation assays for gene silencing study

Protoplasts from *A. fumigatus* Af293 (AfuNK virus-infected wild type isolate) were prepared as described previously.

Transfection using in vitro transcribed dsRNAs (recombinant mycovirus)

Double-stranded RNAs were synthesized from the pAfuTmV-1-NgoMIV and pAfuTmV-1-ALB1 vectors which were derived from the mycovirus-based vector pAfuTmV-1-RNA2, using MAXI Script®T7 and MEGAscript RNAi Kit (Ambion). The transcription reaction was assembled at room temperature as followed.

To 20	µl	Nuclease-free water
1-2	µg	Linearized DNA template with opposing T7
2	µl	10x T7 Transcription buffer
2	µl	10 mM ATP
2	µl	10 mM CTP
2	µl	10 mM GTP
2	µl	10 mM UTP
2	µl	T7 Enzyme mix

The mixture was microcentrifuged briefly and then incubated at 37 ˚C for 16 h. The complementary RNAs were annealed by incubation at 75 ˚C for 5 min and then cooled to room temperature for 5 hr. DNA and ssRNA were removed by treatment with DNase I and RNase (MEGAscript RNAi Kit, Ambion) using the condition given below.

20	µl	dsRNA
21	µl	Nuclease-free water
5	µl	10X digestion buffer
2	µl	DNase I
2	µl	RNase

The digestion was incubated at 37 ˚C for 1 h. Purification of dsRNA was performed using phenol:chloroform extraction (Sevag extraction) as described previously. Additionally, nucleotides, salts and remaining proteins were removed by purification using the column supplied with the kit. The dsRNA pellet was resuspended in 15 µl nuclease-free water and then the integrity of dsRNA was examined by Nanodrop reading and electrophoresis in 1.0% agarose gels.

Approximately 30 µg of *in vitro* synthesized dsRNAs was used to inoculate *A. fumigatus* Af293 protoplasts. Transfection was performed using the same protocol as described previously (see section 10.1 and 10.2). The dsRNAs were co-transformed with 1 µg of pUC18 which was used for an ampicilin resistance gene as a selective marker. Protoplasts transformed with pUC18 were used as a control for the transformation procedure. Additionally in the experiments using GFP assay, dsRNA transcription control provided with the Megascript RNAi kit (Ambion) was used as a negative control. A solution of 200 µl and serial dilution (10^3-10^6) of the solution were plated gently onto the SMM or ACM selection plates. A 0.1 mg/ml ampicillin and IPTG/X-gal solution was added in the media when required. Plates were sealed using a parafilm and incubated right-side up for overnight at 30 ˚C and then inverted the next day and kept at 37 ˚C after formation of transfectant colonies. Single fungal colonies were randomly collected from the selection plates and transferred to fresh ACM plates, following incubation at 37 °C for 3 days.

Transformation using silencing vectors derived from the pSK379 plasmid

For transformation *A. fumigatus* Af293 protoplasts were incubated with 30 µg of supercoiled integrative silencing plasmid vectors (pAF1-ALB1, pAF2-ALB1 or pAF3-ALB1) containing a pyrithiamine (PT) resistance gene (*ptrA*) as a selectable marker. Transformation was conducted using the same protocol as described previously (see section 10.1 and 10.2).

Following transformation 200 µl of protoplasts and serial dilutions thereof (10^2-10^5) were plated onto SMM selection plates without pyrithiamine in order to allow the protoplasts to regerate cell walls. Plates were sealed using parafilm and incubated at 30 ˚C overnight. For selection of transformants, top ACM agar (0.7%) containing 0.05 or 0.1 µg/ml pyrithiamine was overlaid onto the transformant plates and then incubated right-side up overnight at 37 °C. Colonies successfully growing on the surface of the overlay were transferred to pyrithiamine free ACM agar plates for 3-4 days and also reselected on the SMM agar containing pyrithiamine and finally transferred to ACM liquid medium.

The genomic DNA of the transformants was extracted using the DNeasy® Plant Mini Kit (QIAGEN) as described earlier and then used for PCR amplification to confirm the presence of the insertions.

Southern blot hybridization was also performed to verify the presence of the integrative plasmid DNA using DIG-labelled PCR product as a probe. Primers (ptrA-F 5,-TCCTCGGACGAAACTGAACA-3' and ptrA-R 5'-AATCTTCAGGTCCGGACGAG-3', annealing temp = 51 °C, product size= 564 bp) were designed to amplify a fragment of the *ptrA* gene to confirm plasmid integration. Extracted genomic DNA was digested with *Hind*III for 16 h at 37 °C. Digested DNA (20 µl) was electrophoresed on 1% agarose gel for 6 h denaturated and transferred on to Hybond^{TM}-N^{+}- membrane (GE healthcare Bio-Sciences, Germany). Southern blot hybridization and detection of DNA-DNA hybrids were performed using the DIG labeling system and CDP-star (Roche Applied Science, Germany) as described previously.

The integrity of the insert and plasmid construct were also checked by PCR amplification of the promoter::insert::terminator cassette using primers specific to each plasmid construct as follows:-

1) pAF1-ALB1 and pAF2-ALB1
pAF1-F: 5'-TTGGACGGCGTACTAGATCC-3’ and pAF1-R: 5'-CGAGCTTTCCCACTTCATCG-3',
annealing temp= 52 °C, amplicon size= 1028 bp

2) pAF3-ALB1 pAF3-F: 5'-TCGAAGATAGACATGGCCCG-3’ and pAF3-R: 5'-CGAGCTTTCCCACTTCATCG-3', annealing temp= 52°C, amplicon = 1359 bp

Chapter 3

Results and Discussion (Part I)

(Development of RLM-RACE Technology)

Chapter 3: Results and Discussion (Part I)

(Development of RLM-RACE Technology)

1. Characterisation of Aspergillus foetidus virus-fast dsRNA4 (AfV-F dsRNA4)

It has been previously found that the filamentous fungus, *Aspergillus foetidus*, harbours isometric virus particles containing dsRNAs (Buck and Ratti, 1975; Jamal, 2010). The virions showed different mobility rates on gel, thus they were designated as Aspergillus foetidus virus-slow (AfV-S) and Aspergillus foetidus virus-fast (AfV-F) components. Sodium dodecyl sulphate poly acrylamide gel electrophoresis (SDS-PAGE) analysis revealed that these viral dsRNAs were separately encapsidated in different particles (Buck and Ratti, 1975). AfV-S consists of two viruses, AfV-S1 and -S2, and the genomic sequences of both have been analysed and illustrated that the former is a *Victorivirus* in the *Totiviridae* family and the latter is an unclassified mycovirus distantly related to mammalian caliciviruses and plant comoviruses (Kozlakidis *et al*., 2013).

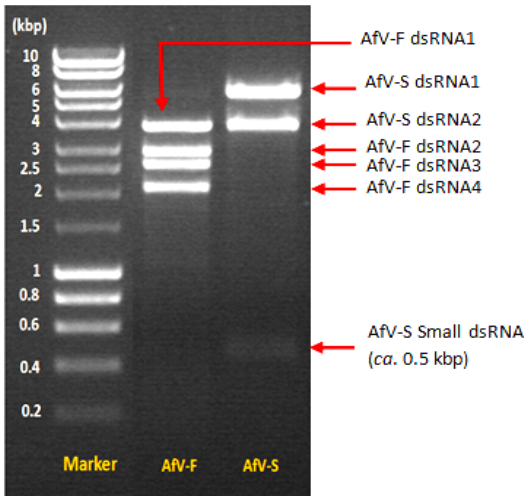


Fig. 7 Agarose gel electrophoresis of Aspergillus foetidus dsRNAs showing RNA profile of *Aspergillus foetidus* virus-fast (AfV-F) and Aspergillus foetidus virus-slow (AfV-S) with their approximate sizes. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

At the beginning of the project, the sequences of three of the four genomic components of AfV-F (dsRNAs1-3) were known but the sequence of the smallest component, dsRNA4 was incomplete. In order to develop technology in preparation of sequencing the 5'- and 3'-termini of the AfuTmV-1 dsRNAs, attempts were made to use and refine the RLM RACE

procedure to complete the sequence of the AfV-F dsRNA4 component which appeared to be *ca*. 2 kbp in length (Fig. 7).

1.1 Construction of a cDNA library of AfV-F dsRNA4 by rPCR method

Various methods and priming strategies were employed to make cDNA clones of AfV-F dsRNA4. Initially random-primed PCR (rPCR) was used with gel-purified AfV-F dsRNA4 as source material as described. Following the PCR amplification step, amplicons of 200 to 600 bp in length appeared as a smear on agarose gels (Fig. 8a). These amplicons were cloned and 23 recombinant clones were sequenced.

Following analysis of the sequence data a sequence specific primer (14-REV-UP primer; 5'-CCTCTCACCCATGAACGATC-3') designed from the upstream region of the longest clone was used in combination with a random hexamer to synthesize the additional cDNAs using a single-primer, genome-walking RT-PCR protocol. This reaction yielded four differently sized amplicons (*ca*. 150, 250, 350 and 500 bp) which were individually gel purified, cloned and sequenced (Fig. 8b).

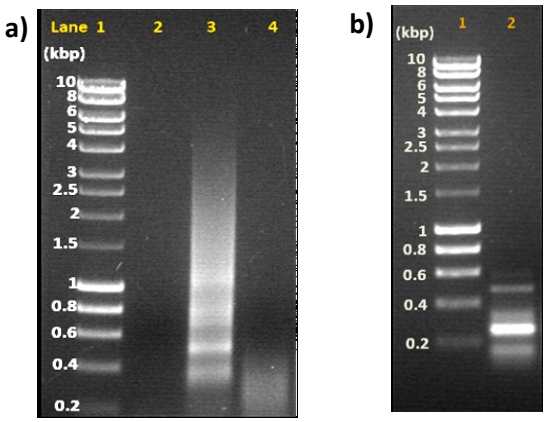


Fig. 8 Agarose gel electrophoresis showing the PCR products generated by random-primed PCR (rPCR) with Froussard primers (a); Lane 1, Marker; Lane 2, negative control; Lane 3, positive control using kanamycin RNA; Lane 4, PCR product from dsRNA4. PCR products generated from 14-REV-UP primer in combination with random hexamer using AfV-F dsRNA4 as a template (b). Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

1.2 Determination of the 5'-and 3'- terminal seqeunces of AfV-F dsRNA4 by RLM RACE

RLM-RACE was employed to determine the 5'-and 3'- terminal sequences of AfV-F dsRNA4 using sequence specific primers for PCR amplification.

Sequences of the sequence specific oligonucleotide primers used *viz*. 14-REV-UP primer; 5'-CCTCTCACCCATGAACGATC-3', and AFV-4-uni; 5'-CTGGGTCATGTAATCTCCTTTGT-3' were based on previously known sequence information and data obtained above. Using the 14-REV-UP primer, three differently sized 5'- specific amplicons *ca*. < 200, 600 and 1000 bp were obtained (Fig. 9a) whereas amplification with AFV-4-uni yielded a smear of 3'-specific products with a prominent band of *ca.* 200 bp (Fig. 9b). The 1 kbp and 200 bp products from each of the above two reactions were individually gel purified, cloned and sequenced.

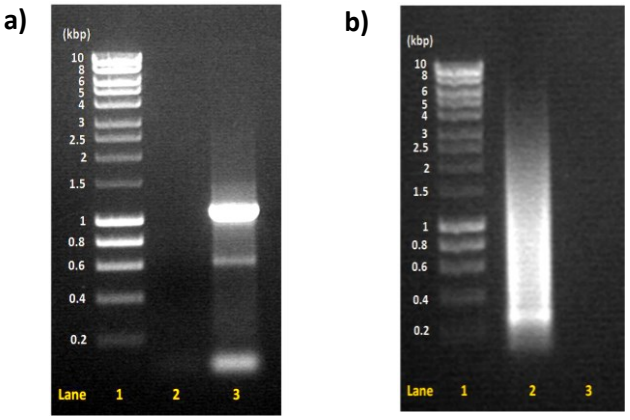


Fig. 9 Agarose gel electrophoresis showing the PCR products generated by RLM-RACE. The amplicons produced from RLM-RACE amplification of the 5'-terminus using AfV-F dsRNA4 as a template (a); Lane 1, Marker; Lane 2, negative control; Lane 3, PCR products; the amplicons produced from RLM-RACE amplification of the 3'-terminus using AfV-F dsRNA4 as a template (b); Lane 1, Marker; Lane 2, PCR products; Lane 3, negative control. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

1.3 The genomic sequence of AfV-F dsRNA4

The genetic organization of AfV-F dsRNA4 is shown in Fig. 10a. All clones were aligned and assembled to obtain the complete sequence of the fourth component. The sequence analysis showed that AfV-F dsRNA4 comprises 2005 bp with 5' and 3' UTRs (of 50 and 168 bp in length, respectively; Fig. 10b). These UTRs flank a single ORF which encodes a putative protein of 580 amino acid residues. The sequence is found to contain a stretch of 3'-poly (A) tails 44 nt in length on the coding strand. The full-length sequence of AfV-F dsRNA4 was deposited in the GenBank database under the accession number HE647818.

a) Genome organisation of AfV-F dsRNA4

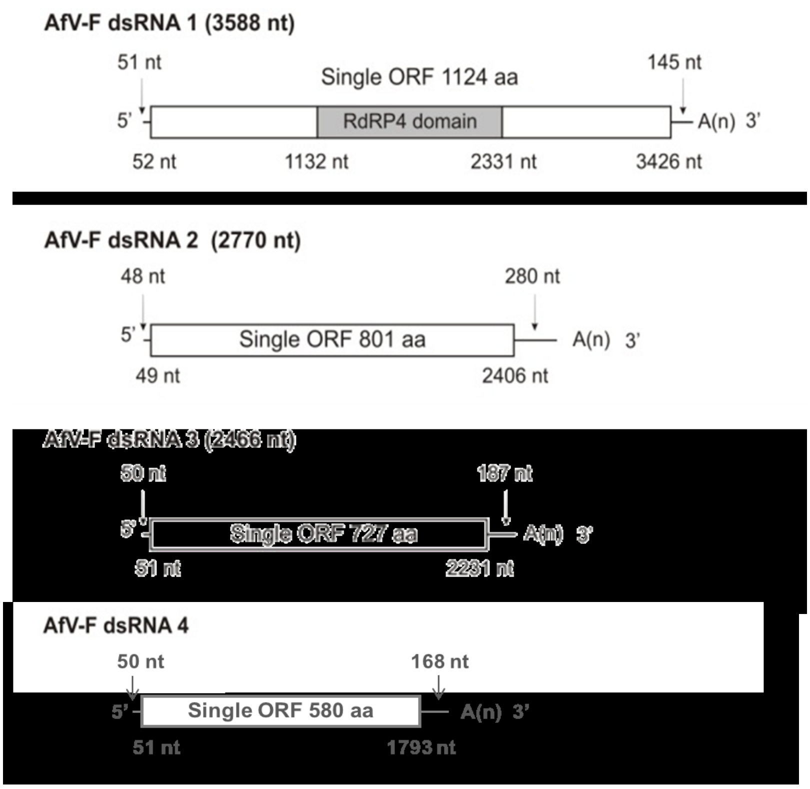
b) Complete sequence of AfV-F dsRNA4 (2005 nt)

AGCUGACAGUCCGGUAAGGAACUCAUUCGAGUACCCUUCCGCUGUACGUGAUGCUUGGUGUGAAUGUGCGAGCGGGUGUCGACAUCCG
UCGGCCCCCACGCCCCCCUGAGCCUGGAGGUCAUGAUUUGACAAUGGUGAUUGAGAGUGAGUUUAGCCGUGCAGCCCGCGGUUUUGCU
UAUACUGUCAGAUUCUCGCGUGGAUUUUCGGAUGAUGCCUGGUUGCCUAAGGCAGAACUCCGCGAGUAUAAGCAGACGCUCGCUGGUC
GUGAUGUGAAGGUUGAUGUCUGUGAGGGAGGUUUCUAUCUGGCCGCUGGAAAGAUAAUGAACCUCAUCCACGCCUUCAUUAGCGACUG
GGAGGGUGGCUGCAUUGGUGAAGCUCAGGUUGAUGUCUCGCUUCCUGGUCCGUCUGACACGCCCCCUAUAACCUUCCGCACACAGGUG
UCCACCCGUACUGUAGCUCCUUUGAAGUACAGUUAUGUGUUCGCCGAGCGCCUAAUUCCGGAGGUUAUUGCCGAGGAUGAUACGCCGG
CGCCUGCUGCGUUGUGCGCUGCGUGUGAUCAUGAGGAAGCCAUCGCCACUGACCUCUAUGUGGCUGGUCCUGUGACUGACACACUUCG
CGUUGAGUGGUUCGUCGAUGUCCUUGAGAAUGCUGGCCUUGGUAGCUCUUUUGCUGCUAUCGGUAACUUCUCUCAUGAGGCGUCGGAA
GUGCUCGCGUCGUAUGGGUGUCGCCGGCUGCCUGAGUGGCUGUGGGGUGAGCAGUUCGAUCCGCCGCCUGAUAUACCGCGCUUUAUCU
UGGUUGGUGACCGUCCCCGGGAGGGCUGGCUAAGUACAAAUCGAUUCUUAUCGAUCCUUGCCUCUCGCGGUUACCACGCCGCCGCCUG
GGCACAGCAUCUCAAGGGGCUUGAGGGGGAGUUCUUCUUCUCUUGCCCACCGACUGGUGUCUAUGGUGGAUGGCUGAUAGAGAGUGCU
GUGGAUGAACGGGGGCCGUUGUGCUACGGCAAAAUUGAGUGUGUCGCGGGGUUGGACCCUACAUUCUCUCUCUUUUGCGAUCGAUCGU
UCAUGGGUGAGAGGUUUCCUUUCCGCAUAGCGAAGGAAUUUGGCCUGAGCGCUGUUGCGCUAACUCAGGGUUGUACCGUCACUGUGCC
UCUUGAUCAGACGGUAUCCUGGCUCUGUGUUCUGGCUUAUGCGGUGGGGGCCUGGUGCGUUCAGGAUGUGCGUCCUAUGCUCGAUCUC
GAGCGGUACUCUAGCGCCGUGGGCAUCGCCCAGAGCUUGCGCCGUAUCGUGCAUCCGCCGGGGCGCUUCGACAUAUGGGAUGAUAGCG
GUGCCGGUAUGGCCUUUGACAACUUCUUCCAGCUGCCACCGCUCCUAACGGAUGUGCUUGACCGUCGCCUCGCGCUCCUGAUAAGUUA
UACGCGUGAGGGUGCUCCUAUCCUUUGGGGCGAGCGAUGGGAUCCGCCGGGAAAGCGCCUUUACCAUAUGUACGACGAUCUGCGUGAG
AAGGUGUCGCCGGAGGAGUUAGAUUGGGAGCCUGGGAUAGUGCUUGAGCGCCGACAUCUGUCGGUACUCUGCUCAGCGCGCAUUUUCC
GCUUGCUGGCGUAUCUUGUGCGUGGUUAUCAUGCACUUCGCAGGCCUCUCAGUAUGCGCACUGUUCAAUUGAUUUGUUGGGUUUGCUC
GCGCUUGGAUGACCCACUGACCAUUGUGCGUUUCCUGUGGUACCCCUUCUCGGGCCUGGCACGCCAUGAUUUGGAUUUCGUAUUGCCC
CUCGUGGAAUAUGCACCUUGCUGGGUCAUGUAAUCUCCUUUGUCAUACCUAUAAGGGAGGCCAGAUCCCUGUUUUGCAAUGGCUGGCU
AGCCCCGGAGUGCCCCGUAAUGCCCCAGCAGUGUCCGGGAAUAAGGCCGAGCUGUGGAAAAUAGUCAUAUGGGGGAUCCCUGAGGGGU
GAGCAUCAAGGGCCACAUUUUUGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 10 Diagrammatic representation of genome organisation of AfV-F dsRNA4 (a); and its complete sequence (b). Start and stop codons were shaded in pink and red, respectively.

The complete genome organization of Aspergillus foetidus virus-fast (AfV-F) and Aspergillus foetidus virus-slow (AfV-S) are shown in Fig. 11.

a)



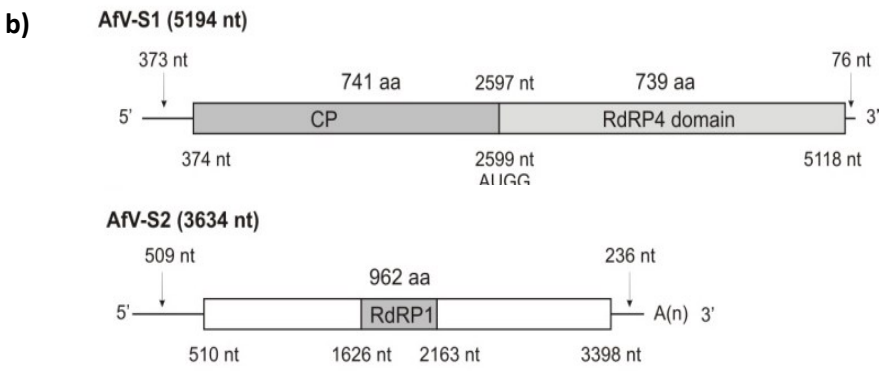


Fig. 11 Diagrammatic representation of the genome organization of Aspergillus foetidus virus-fast (AfV-F) dsRNAs (a); and Aspergillus foetidus virus-slow (AfV-S) dsRNAs (b).

The sequences of 5'- and 3'-UTRs of the four AfV-F dsRNAs were aligned using MAFFT multiple alignment program (Katoh *et al*., 2002). The alignment showed strictly conserved GCUGACAGUCCG- and -AUCUCCUUUGU sequences at the 5'- and 3'- termini, respectively, among those four components (Fig. 12a and 12b). It been suggested that the presence of conserved terminal sequences is a feature of ssRNA and dsRNA viruses with multipartite genomes and it is important in transcription, replication cycle and packaging of viral RNA (Wei *et al*., 2004). Additionally, the presence of poly(A) tail at the 3'-termini of AfV-F dsRNAs might be involved in stability and translatability of the RNA sequences by exhibiting a protective function against ribonecleases and also serving as template for initiation of translation as described in picornavirus (Bergamini *et al*., 2000).

a) Multiple alignment of nucleotide sequences of the 5'-termini of the four AfV-F dsRNAs

```
dsRNA1          UAGCUGACAGUCCG-GUUUGGAACUCACUCGAGUACCCAUCCGCUGUGCGUG
dsRNA2          -GGCUGACAGUCCGUGUUUGGAUCUCAUUCGAGUACCAUCCGGCUGUAC
dsRNA3          -GGCUGACAGUCCG-UAGUGGAGUAGAUUCAACUACCCACACGCUGUACGAG
dsRNA4          -AGCUGACAGUCCG-GUAAGGAACUCAUUCGAGUACCCUUCCGCUGUACGUG
                   ************ :: :*** :::*:**:*:****:  ::*****:*: :
```

b) Multiple alignment of nucleotide sequences of the 3'-termini of the four AfV-F dsRNAs

```
dsRNA1          AUCUCCUUUGUUCAAUCAUUUAUCCACUGAGUGCGAGGCUACAUCCCUG---AUUCGUCA
dsRNA2          AUCUCCUUUGUCUAU-UCCUUAU-------AUGCGAUGCUGGAUCCCCA--CCUACGGUG
dsRNA3          AUCUCCUUUGUCUAUAUCACUAU-------UCGCGAGGCUGGAUCCCCACGCUUGCCGUG
dsRNA4          AUCUCCUUUGUC---AUACCUAU-------AAGGGAGGCCAGAUCCCUG----UUUUGCA
                 *********** :   :   ***         *:**:**: :*****      * : :

dsRNA1          AUGGGUAGCUAGC----------------ACGCCUGAG----------------------
dsRNA2          AUGGCCAGCUAGCAACGGAGCGUGCCGU-ACACGCCAGCAGUGUCCGGGAAUAAGGCCGA
```

```
dsRNA3          AUGGCCAGCUAGCUACGGAGCGUGCCGU-ACACGCCAGCAGUGUCCAGGAAUAAGGCUGA
dsRNA4          AUGGCUGGCUAGCCCCGGAGUGCCCCGUAAUGCCCCAGCAGUGUCCGGGAAUAAGGCCGA
                ****: :******   :::: :  :::: *: * ::**::::::::: :::::::::: ::

dsRNA1          --------------------GUCAGAUCCCUGGUAGGUGAGUACUAAUGGCUGACUAGGG
dsRNA2          GCUGUGGGUUAAAGUCAUAUGGUGGAUCCCUGAUGGGUGAGCAUCAAUGGCCACAUG---
dsRNA3          GCUGUGGUUUAAAGUCAUAUGGCGGAUCCCUGGUGGGUGAGCACCAAUGGCCGCAUG---
dsRNA4          GCUGUGGAAAAUAGUCAUAUGGGGGAUCCCUGAGGGGUGAGCAUCAAGGGCCACAUU---
                 :::::::   : :::::::::*: :******** ::*******:* :**:***: ::*

dsRNA1          UUCGUUCAUAACCCA(N)
dsRNA2          -------UCUGACCA(N)
dsRNA3          -------UCUGGC-A(N)
dsRNA4          -------UUUGCC-A(N)
                       : :: * *
```

Fig. 12 Comparative alignment of the 5-UTRs (a) and 3-UTRs (b) of AfV-F dsRNAs using the MAFFT program (Katoh *et al.,* 2002). Asterisks signify identical nucleotides at the indicated positions and colons specify that three out of four nucleotides are identical at the indicated positions.

1.4 Characterisation of the defective RNA of A. foetidus virus-slow (AfV-S)

Examination of the AfV-S dsRNA profile reveals the presence of a small dsRNA *ca.* 500 bp in size (Fig. 7). Unsuccessful attempts were made to generate cDNA clones of this species using the rPCR procedure. The reasons for this failure are unknown and new approaches to obtain clones will be made in the future.

Chapter 3

Results and Discussion (Part II)

(Characterisation of *A. fumigatus* Af293 Mycovirus)

A novel mycovirus from the human pathogenic fungus *Aspergillus fumigatus* contains four unique double-stranded RNAs as its genome and is infectious as dsRNA

We report the discovery and characterization of a double-stranded RNA (dsRNA) mycovirus isolated from the human pathogenic fungus *Aspergillus fumigatus* - Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1) that reveals several unique features not found previously in positive-strand RNA viruses including the fact that it represents the first dsRNA that is not only infectious as a purified entity but also as a naked dsRNA. The AfuTmV-1 genome consists of four dsRNAs the largest of which encodes an RNA dependent RNA polymerase (RdRP) containing a unique GDNQ motif normally characteristic of negative-strand RNA viruses, and in some positive-strand RNA viruses. The third largest dsRNA encodes an S-adenosyl methionine-dependent methyltransferase (SAM) capping enzyme and the smallest dsRNA a PA rich protein which apparently coats but does not encapsidate the viral genome. A combination of a capping enzyme with a picorna-like RdRP in the AfuTmV-1 genome is a striking case of chimerism and the first example to our knowledge of such a phenomenon. The replication cycle of AfuTmV-1 appears to be intermediate between dsRNA and positive-strand RNA viruses, as well as between *bona fide* viruses and capsid-less RNA elements.

The unique sequences of the four AfuTmV-1 dsRNAs and the proposed roles of the proteins predicted from those elements in virus replication, including the combination of a capping enzyme and the RdRP on separate dsRNAs strengthen the suggestion for the assignation of a new virus family where AfuTmV-1 would represent the prototype member.

Mycovirus | infectious dsRNA | capping enzyme | virus evolution.

Chapter 3: Results and Discussion (Part II)

(Characterisation of *A. fumigatus* Af293 Mycovirus)

2. Characterisation of Aspergillus fumigatus tatramycovirus (AfuTmV-1)

Prior to exploiting mycoviruses as a potential vector for silencing genes in *A. fumigatus*, the genomic sequences of the viruses to be used in the study need to be determined. Presently the complete genomic sequences of two of the viruses found in a large collection of *A. fumigatus* environmental and clinical isolates (Bhatti *et al*., 2012) are known, namely Aspergillus fumigatus chrysovirus (AfuCV: Jamal *et al*., 2010) and Aspergillus fumigatus partitivirus-1 (AfuPV-1; Bhatti *et al*., 2011a). However both of these viruses are encapsidated and their dsRNA genomes may not be suitable substrates for incipient silencing by the fungal silencing machinery. This concept was demonstrated clearly when the levels of siRNA accumulation of viral origin in a fungus infected with an encapsidated dsRNA totivirus were significantly smaller by a factor of 150-fold as compared to those found in a fungus infected with an unencapsidated hypovirus (Himeno *et al*., 2010). These observations suggest that dsRNA mycoviruses could avoid antiviral silencing defence mechanisms of their hosts by retaining their genomes and RNA replication enzymes in virions. Thus it may be wholly advantageous to investigate the Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1) in detail, which appears to be unencapsidated and is present in low concentration in the fungal host. With this in mind, it is imperative to clone and sequence the genome of AfuTmV-1 with the long term goal of exploiting elements of its dsRNA genome in viral vectors for fungal gene silencing.

Aspergillus fumigatus clinical isolate Af293 was originally obtained from Professor Nancy Keller, Department of Medical Microbiology and Immunology and the Department of Bacteriology at University of Wisconsin, Madison, Wisconsin, USA through Dr Elaine Bignell, CMI, Imperial College, London. The Af293 isolate contains AfuTmV-1 which has a dsRNA profile comprising four segments ranging in size from 1.2 to 2.4 kbp (Fig. 13) that could be resolved in 1% agarose gel. These viral dsRNAs were designated as dsRNA1 to dsRNA4 with increasing order of mobility. Interestingly, it can be seen from Fig. 13 that dsRNA2 shows the highest accumulation as compared to the other three dsRNAs. In addition, four dsRNAs of AfuTmV-1

have always been found together, suggesting that they likely represent the genome of a single virus.

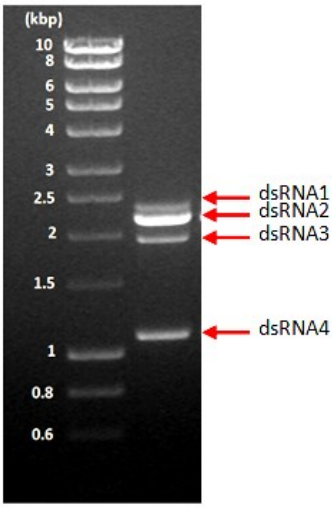


Fig. 13 Agarose gel electrophoresis of Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1) dsRNAs. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

2.1 Sensitivity of AfuTmV-1 and dsRNA to DNase, S1 and RNaseIII

The chemical identity of AfuTmV-1 genomic material was confirmed by resistance to DNase I, S1 nuclease and RNase III. To verify that AfuTmV-1 contains dsRNA as its genome nucleic acids from Af293 mycelia were extracted using the LiCl extraction method which fractionates and separates ssRNA and dsRNA.

Phenol/chloroform-treated dsRNA extracts were subsequently digested with DNase I (digests ssDNA and dsDNA), S1 nuclease (digests ssDNA and ssRNA) and RNase III (digests dsRNA, and converts long dsRNA into siRNA). The digests were analysed by 1% agarose gel electrophoresis showed that the elements were resistant to DNase I and S1 nuclease but sensitive to RNase III, suggesting that these elements are dsRNA molecules (Fig. 14a).

The unencapsidated nature of the virus was confirmed using RNase III enzyme treatment of virus-like particles (VLPs) and AfuTmV-1 dsRNA. Equal concentrations (*c.a.* 15 µg/ 100 µl) of purified VLPs and AfuTmV-1 dsRNA were digested with RNase III for 20 min. Control treatments with buffer and incubation only and no RNase III had no effect on the dsRNA profile of neither VLPs nor dsRNA (Fig. 14b). However following RNase III treatment of both VLPs and AfuTmV-1 dsRNA no dsRNAs were visible. This result suggesst that the AfuTmV-1 dsRNAs are not conventionally encapsidated in virus particles. Encapsidated Aspergillus

fumigatus partitivirus-1 (AfuPV-1) was also used as a negative control. Purified AfuPV-1 treated with RNase III showed clear dsRNA bands after phonol/chloroform extraction whereas the purified AfuPV-1 dsRNAs treated with RNase III showed no visible dsRNA bands on the gel (data not shown).

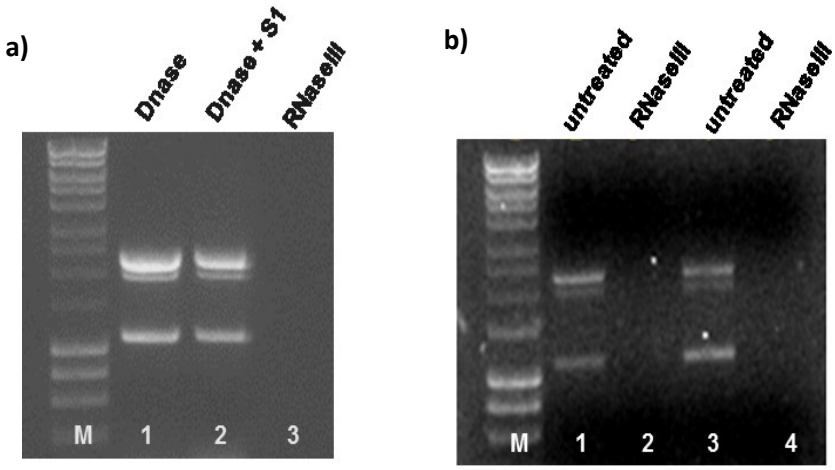


Fig. 14 Sensitivity of the AfuTmV-1 VLPs and dsRNA to enzymes. AfuTmV-1 dsRNAs were digested with DNase, Lane 1; DNase+S1, Lane 2 and RNase III, Lane 3 (a). RNase III digestion of virus-like particles and dsRNA extracts (b). Crude virus-like particle extract, Lane 1; crude virus-like particle extract digested with RNase III, Lane 2; phenol/chloroform-treated dsRNA, Lane 3; and phenol/chloroform-treated dsRNA digested with RNase III, Lane 4. Hyperladder 1 (Lane M; 10 kbp; Bioline) was used as a marker.

2.2 Construction of cDNA clone library of AfuTmV-1 dsRNAs by rPCR

Unfractionated AfuTmV-1 dsRNAs were used as a template for rPCR amplification using the Froussard primers as described. The PCR amplicons generated in the procedure appeared as a smear of products 200-1000 bp in size on agarose gels stained with ethidium bromide (Fig. 15).

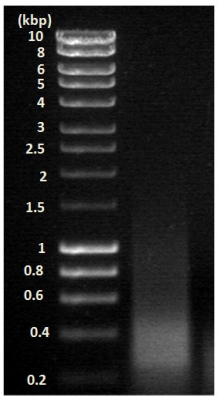


Fig. 15 Agarose gel electrophoresis showing the PCR products generated by random-primed PCR method (rPCR) using AfuTmV-1 dsRNAs as a template. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

PCR amplicons were ligated into the pGEM-T Easy vector and the ligations were transformed into *E. coli* competent cells as before. Putative recombinants were screened as previously and 19 of those containing inserts were sequenced. Insert sequences (200-500 bp in size) were analysed using the BLASTX protein search tool to identify homology with reference protein sequences in the GenBank database. One clone was predicted to share homology with a putative methyl transferase gene and the sequence of a dsRNA element found in the flax rust fungus *Melampsora lini* (Dickinson *et al*., 1993; Accession number X64371).

Two contiguous sequences were obtained from the sequence assembly of AfuTmV-1 cDNA clones generated from rPCR and it was possible to assemble these 2 contiguous sequences - contig 1 and contig 2, 892 and 864 bp long respectively (Fig. 16). The two contiguous sequences were screened with BLAST programme to interrogate the GenBank protein database and both were found to share significant homology with a protein found in a fungus from the genus *Alternaria* sp. FA0703, nominated as the complete mRNA of a dsRNA segment (Clone AltR1; Accession number FJ595830). Two further complete accessed dsRNA sequences from this fungus are also deposited in the GenBank database. Contig1 and contig2 were respectively positioned 5'- and 3'-terminal on the homologous *Alternaria* mRNA sequence (Fig. 17) and did not overlap.

(a) Contiguous sequence 1

>contig1 (seq15+9+14) (892nt)

UGUGCCACCUCUCCCCCAACCCCUCCGCGAUCACCGAUUACGGGGUGGAUGAGGUCAAGAGCGUGGUGUCAUGGCUCGUCUCCUCGCC
CGGGAUGUUCUACGGCGACAUUGUUGCCGCCACUGGAAUCACCCCUGACAUCCUCUCUACUUUCCUUCCCGCGCUCGAUGAUCAAGCC
UGCGCCGACGCCGCCUCUCUUGCCAACGCCAUGGCGACCCCCGAUGCCGAGUGGGACCGCGAGCACCGCAACCGUUUCACGUCCGCAC
UCGCGCACCACUCCGCCGACGCUGCUAGGCUCGCCGCCCAGUCGUCGGCUGAGUUCGCUGCGGUCGACGCCAUGGAAGAGGCUGGCUA
CGAGAAUCAGGACCACGCUCGCACCGCGGUCAACGAGCUCAACCGCGUCUAUGGCAAUACCUACCGGGUCUUCACUCGUCAACACGCC
CGCUCCCGCGUGGCCAUCGCCGCCUCUUUCAAGAUCGAUGACCACGGCACCCUCUUCCACGUCCUCACCCCCUGGUGCCUCUCCCGAA
CAUCGGCCAUGAUCUACGUCGCUGUCAUCGCCAGGCAUGUGAGGCGUCGCCACAUCGAGGAGCGCGCCUCCAGAUUGGCACGCACUGC
CCUGACCUACGCCGCUCGCGCCGGCUUAGAGGCUGUACGCGCCAUCAUAGAGCCCGCCAUUGCCAUAGUGACAAGUUUCCGUUACGAC
CCUGCCAAGAUCGCCUUCGUGGACGAUUUUGACCGGGUCAGGGAGUCCAUCACUACGAGGAGUCCUCCAACGCUUGUCGCCUUCGCUU
GCGUCCAUAUGUGCGGGAACCAGGACGUCGUUUCCUUGGCGGCGCGACUCCGCAUGGAGUCGCUACCACAACUCCAUGAGACCCCGCC
AGCCGUCGCCGA

(b) Contiguous sequence 2

>contig2 (seq13+25) (864nt)

UGCGCUCGAAGGGGGUCGACACCCUCAAUACACUGUUCGUCAUCGAAUGGGGCGGUGAAUUUGACUAUCCUACCGUCAUGGCUGCCCU
CGCCGUUGCCAGGUUGGAUGUCUGUGUCGACAUUGGUCCCUCGGGCAUUGACCUACCCGGCGCCGACGCCAAGGCCGAGGAUGGCACC
GAAGAGUACAACUACUCUUUGCUGCUCGCCUCAGCCGAGGAGCGGCGCAUGCCUCGUAUGCCCAGGGUGCCCUAUCCGUCCGACGAGC
CCCUUGAGGCCAAACUGCGGCGCGUCGCGCAGUUCUACUCUCUCUCGGGCUCUGCCAAUAUAGCAUACGUCAGCGGGGGGGUCACCCA
AAACGAGGGGGUUCCUGUCAACAUCUGUAUCGACACCGCCAACAGGCUUAAUGCCAUACGCAAUGUCACCCCGGACGUUAACGUCGUU
UAUGCCUGCGCCGAAGCUCUCCUCCCCCCCCCUCUGCGGGCACGGACUCGAGCUCGGGCCCGACGACUACCUUCAGUGCGCGUCCCUGG
UCGAUGAGGAUUGCCCCGCCUGCAAUGCUCACUACCGCGCCACCGCCGUCCUCUCUGCUGCCUGCGCUGACAGACAGGUGCGCAUCGU
UAAGCCAAGAUCCGCCUACGCUCAUAACGGGCACCUGUCCUUGGAGCUCGUCCCGGGCCACCCUACAUGCGUGGACGACACUCUCCUC
ACCGUCGACUCCAUGGUCGCGGCCAAUGUCCUCCGCAACCACGAUUGGGGCGAUGAGCCCGUUACUCCACCAGCCGGCGGCAACCCGA
CAUCGCCUACCCUCGCCUGCGUCACAGCGGAGAACAUGCGCACUGUGUACUCCGCUCUCAAUAACGGCGCCA

Fig. 16 Two contiguous sequences obtained from the sequence assembly of AfuTmV-1 cDNA clones using the CAP3 sequence assembly program (Huang and Madan, 1999). Contig 1 was assembled from clones 9, 14 and 15 resulting in an 892 bp fragment (a); Contig 2 was assembled from clones 13 and 25 resulting in an 864 bp fragment (b).

(a) Contiguous sequence 1

```
>gb|ACL80752.1|  hypothetical protein [Alternaria sp. FA0703] Length=696 Score =  145
bits (367),  Expect = 1e-36 Identities = 116/280 (41%), Positives = 165/280 (59%), Gaps
= 1/280 (0%) Frame = +3

Query  6    HLSPNPSAITDYGVDEVKSVVSWLVSSPGMFYGDIVAATGITPDILSTFLPALDDQACAD  185
            HL+ +P +       + V  V++W+ ++PG F+GD++++ GITP   +  L  L D+  A+
Sbjct  27   HLTKSPVSHVLVDANLVSEVLAWMAAAPGPFFGDLISSAGITPADCAQLLEDLTDKQAAE  86

Query  186  AASLANAMATPDAEWDREHRNRFTsalahhsadaarlaaqssaefaavdaMEEAGYENQD  365
            AA++ NAM   D EWDR  R+++T+A   H  DAA++AAQSSAEFAA+DA+  AGY   D
Sbjct  87   AATMVNAMTIADDEWDRGTRDQYTTATVLHERDAAKIAAQSSAEFAALDALTAAGYAGVD  146

Query  366  HARTAVNELNRVYGNTYRVFTRQHARSRVAIAASFKIDDHGTLFHVLTPWCLSRTSAMIY  545
             A  A++ + R + + Y+V  RQ ARSR  + A F   D G +  V+TPWCLSR +A  Y
Sbjct  147  DAVGALSVMRRTFKHNYKVHVRQFARSRHCVVARFVRSD-GAVRIVITPWCLSRRNAQAY  205

Query  546  vaviarhvrrrhieerasrlartaltyaaraGLEAVRAIIEPAIAIVTSFRYDPAKIAFV  725
             A+      R  +E R    ART  T+A R  L+AVR II PA+A++T + YD   +AFV
Sbjct  206  AALCMADWHRHMVEMRTINRARTVATFAERTSLQAVRDIIAPALALLTRYSYDATNMAFV  265

Query  726  DDFDRVRESITTRSPPTLVAFACVHMCGNQDVVSLAARLR  845
            D   R  + +  R+P   VAFA V   G  D     AR R
Sbjct  266  DGSGRPVQKVGARAPYVAVAFAAVFSSGAGDATLGLARAR  305
```

(b) Contiguous sequence 2

```
>gb|ACL80752.1| hypothetical protein [Alternaria sp. FA0703] Length=696 Score =  186
bits (472),  Expect = 2e-51 Identities = 108/289 (37%), Positives = 169/289 (58%), Gaps
= 6/289 (2%) Frame = +2

Query  2    SFLRSKGVDTLNTLFVIEWGGEFDYPTVMAALAVARLDVCVDIGPSGIDLPGADAKAEDG  181
            + L  KG    +TL+V+EWGG      V+ A A A++D+ +D+  SG+D+PG D  A+D
Sbjct  390  ALLAMKGHHKDSTLYVVEWGGRVSAHAVLGAAATAKIDIALDVAGSGVDVPGIDVYADDD  449

Query  182  TEEYNYSLLLASAeerrmprmprVPYPSDEPLEAKLRRVAQFYSLSGSAN---IAYVSGG  352
               ++Y L LASA  R +PRMP VPY +  PL  KL  V  + S+ G A+   +AYV+GG
Sbjct  450  DPPHHYQLYLASARRRMLPRMPSVPYRAGTPLIEKLSLV--YESVGGRASDYKLAYVNGG  507

Query  353  VTQNEGVPVNICIDTANRLNAIRNVTPDVNVVYACAEALLPPLCGHGLELGPDDYLQCAS  532
            ++Q    PV++C D+  R+ A+      + + +   E LLPP C H +++ PD+YL
Sbjct  508  MSQLAETPVSVCTDSNARMAAMDGARLTLPIPFYSTEVLLPPACHHIVDMDPDEYLDDFG  567

Query  533  LVDEDCPACNAHYRANAVLSAACADRQVRIVKPRSAYAHNGHLSLELVPGHPTCVDDTLL  712
              + +C  C AHYRA + +  A     VR+ KPRS +AHN    LE + G  + ++D+L
Sbjct  568  TGNPNCFQCEAHYRALSAVGDALDRPGVRLTKPRSVFAHNALFGLEWI-GTDSAMEDSLE  626


Query  713  TVDSMVAANVLRNHDWGDEPVTPPAGGNPTSPTLACVTAENMRTVYSAL  859
            T+D+ +A N LRN+++ + P  P  G + +SP L  +  +++R VY+ L
Sbjct  627  TLDAAIACNALRNYEYPNPPPDPTPGKDLSSPDLLELMQDSVRQVYTLL  675
```

(c) Positions of AfuTmV-1 contiguous sequences on the homologous *Alternaria* mRNA sequence

```
MADLTRLRTILLEHGSSVVPVGSAVYHLTKSPVSHVLVDANLVSEVLAWMAAAPGPFFGDLISSAGITPADCAQLLED
LTDKQAAEAATMVNAMTIADDEWDRGTRDQYTTATVLHERDAAKIAAQSSAEFAALDALTAAGYAGVDDAVGALSVMR
RTFKHNYKVHVRQFARSRHCVVARFVRSDGAVRIVITPWCLSRRNAQAYAALCMADWHRHMVEMRTINRARTVATFAE
RTSLQAVRDIIAPALALLTRYSYDATNMAFVDGSGRPVQKVGARAPYVAVAFAAVFSSGAGDATLGLARARKAAADTL
TDAVKLKTATDYLEMADSPFAGLALFSRALAEYRDKRHASPWGDRVNKGAMLGYVARERDKLTRNTAACINRIEEIAA
LLAMKGHHKDSTLYVVEWGGRVSAHAVLGAAATAKIDIALDVAGSGVDVPGIDVYADDDDPPHHYQLYLASARRRMLP
RMPSVPYRAGTPLIEKLSLVYESVGGRASDYKLAYVNGGMSQLAETPVSVCTDSNARMAAMDGARLTLPIPFYSTEVL
LPPACHHIVDMDPDEYLDDFGTGNPNCFQCEAHYRALSAVGDALDRPGVRLTKPRSVFAHNALFGLEWIGTDSAMEDS
LETLDAAIACNALRNYEYPNPPPDPTPGKDLSSPDLLELMQDSVRQVYTLLAGPFAGPLPSADAQSVAASVA
```

Fig. 17 The putative proteins predicted to be encoded by the sequences of contig1 and contig2 and their similarities to the *Alternaria* sp. FA0703 clone Altr1 hypothetical protein are shown in (a) and (b) are highlighted in (c) in red and green respectively.

In order to identify the sequence located in the putative gap between the contig1 and contig2 sequences, two oligonucleotide primers were designed for RT-PCR. Namely a forward primer based on the sequence of contig1, UCV gap FOR2 (5'-ATGTGCGGGAACCAGGACGTCGT-3') and a reverse primer based on the sequence of contig2, UCV gap REV2 (5'-CGAACAGTGT ATTGAGGGTGTC-3'). Following PCR amplification, the 340 bp amplicon (Fig. 18) was cloned and sequenced and the sequence of the nearly full length cDNA clone of the dsRNA was assembled.

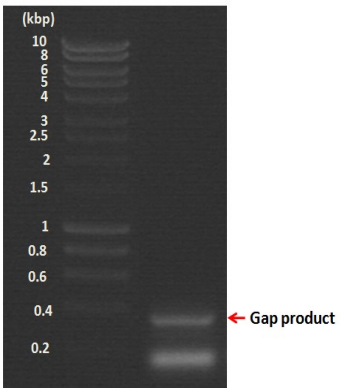


Fig. 18 Agarose gel electrophoresis showing the PCR products generated with gap primers using AfuTmV-1 dsRNAs as a template. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

Examination of this assembled sequence and a comparison of the putative polypeptide encoded by the sequence with the *Alternaria* sp. FA0703 clone Altr1 hypothetical protein revealed that both N- and C-terminal sequences of the protein were absent and that the sequence of the UTRs were yet to be determined. From the size of the assembled cDNA sequence it would appear that it represented clones of the AfuTmV-1 dsRNA2 element when the profile of the dsRNAs was examined on agarose gels (Fig. 13).

2.3 Determination of the 5'- and 3'-termini of dsRNA2 using RLM-RACE

Three sequence specific 5'- and 3'- specific oligonucleotide primers were designed from the assembled contig sequence of AfuTmV-1 dsRNA2 for RLM RACE viz. (15-RLM-UP; 5'-TGGTCCTGATTCTCGTAGCC-3'; UCV2-RLM-UP; 5'-GACCTCATCCACCCCGTAAT-3' and 13-RLM-DWN; 5'-AGATCCGCCTACGCTCATAA-3' respectively) using 114.5 ng/µl of dsRNA in the reactions. The RLM-RACE amplicons generated from the two 5'-end specific primers, 15-RLM-UP and UCV2-RLM-UP were *ca*. 500 and 200 bp in size, respectively, whereas the amplification

of 3'-terminal amplicon was *ca*. 400 bp in size (Fig. 19a). All three amplicons were cloned into pGEM-T Easy vector, sequenced and the sequences analysed as before.

Besides, RLM-RACE also generated totally fortuitously an apparently full length amplicon of another of the AfuTmV-1 dsRNA elements *ca*. 1.1 kbp in size (Fig. 19b). From the size of the amplicon it would appear that it represented clones of the AfuTmV-1 dsRNA4 element when the profile of the dsRNAs was examined on agarose gels (Fig. 13).

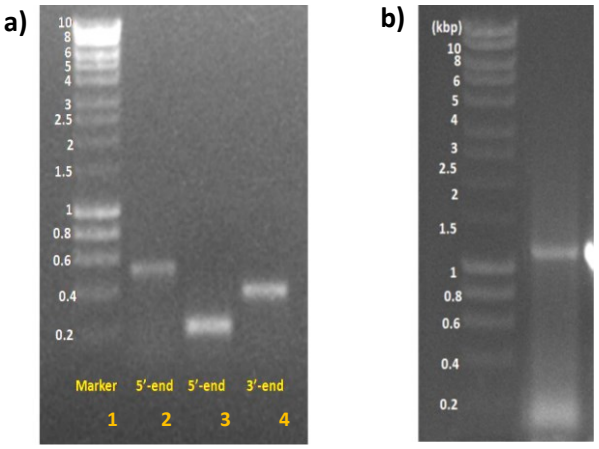


Fig. 19 Agarose gel electrophoresis showing the PCR products generated by RLM-RACE. RLM-RACE of AfuTmV-1 dsRNA2 (a); Lane1, Marker; Lane 2, the amplicons produced from RLM-RACE amplification of the 5'-terminus using 15-RLM-UP as a primer; Lane 3, the amplicons produced from RLM-RACE amplification of the 5'-terminus using UCV2-RLM-UP as a primer; Lane 4, the amplicons produced from RLM-RACE amplification of the 3'-terminus using 13-RLM-DWN as a primer; and RLM-RACE of AfuTmV-1 dsRNA 4 with a *ca*. 1.1 kbp PCR product (b); Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

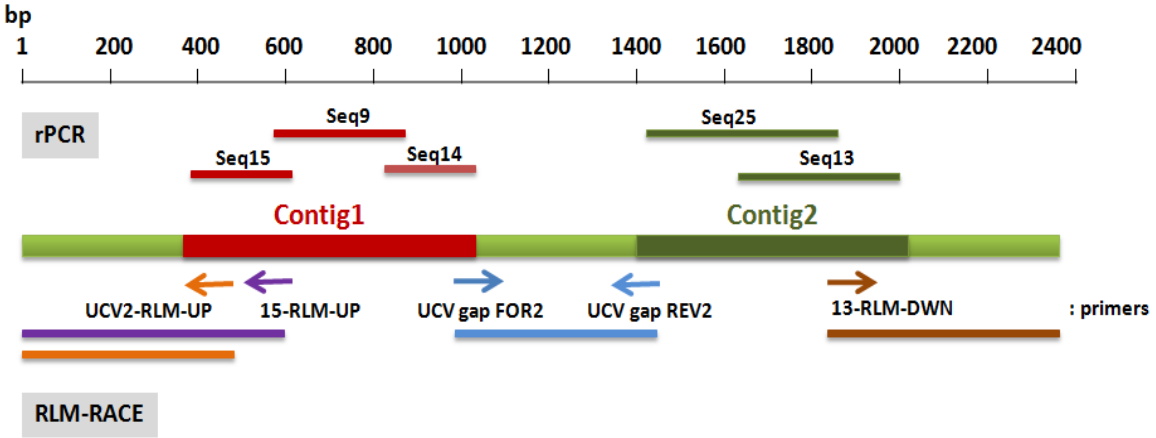
2.4 The genomic sequence of AfuTmV-1 dsRNA2

The contig sequences, gap-filling RT-PCR sequence and the 5'- and 3'-terminal sequences obtained from RLM-RACE were aligned and assembled to obtain the complete sequence of AfuTmV-1 dsRNA2. Completing the sequence of dsRNA2 revealed a basic structure of genetic organization that it comprises 2233 bp with 5' and 3' UTRs 70 and 72 bp in length, respectively. The 5'-sequence of AfuTmV-1 dsRNA2 contains two AUG codons and the upstream one is in a more favourable context to potentially initiate translation (Kozak, 1986; Lütcke *et al*., 1987). The UTRs flank a single ORF2 which encodes a putative protein of 696 amino acid residues on the positive (sense) strand (calculated molecular weight 75.76 kDa). GC content calculation revealed that the genome contains 62% GC (Fig. 20). No ribosomal frameshifting signal (shifty heptanucleotide motif GGGAAAC; X-XXY-YYZ where X is A, C, G or U, Y is A or U and Z is A, C or

U) was predicted from the sequence using FSFinder algorithm (Moon *et al.*, 2004). A schematic representation of the genetic organization is shown in Figs. 20a and 20b.

Following a search of the public databases the deduced amino acid sequence of the AfuTmV-1 dsRNA2 ORF using PROSITE, Pfam protein database, Conserved Domain database and SUPERFAMILY1.75 revealed no match to any accessed protein on the database. However, the result of PSI-BLAST search and BLASTX showed that AfuTmV-1 dsRNA2 shared respectively 39% amino acid sequence identity to a dsRNA encoded protein found in a fungus from the genus *Alternaria* sp. FA0703 (AA1; a 2233 bp dsRNA fragment, ACL80752.1, *E* value = $3e^{-156}$), a *ca.* 2.3 kbp mycovirus dsRNA, nominated B2, isolated from the flax rust, *Melampsora lini* (RZB2: Ren Zhang unpublished results; *E* value = $6e^{-109}$; 35% identical; 53% similar) and a hypothetical protein encoded from Cladosporium cladosporioides virus 1 (YP_009052471.1, identity = 33%, *E* value = $6e^{-83}$). The amino acid sequence was also searched against a MEROPS protease database (Rawlings *et al*., 2012). The result revealed most significant similarity to a family C26 unassigned proteases (cysteine catalytic type) from *Dehalogenimonas lykanthroporepellens* (MER451098, identity = 38.3%, *E* value = $3.1e^{+00}$). However, it cannot be inferred that protein encoded from ORF2 is associated to a cysteine protease since the *E* value of the sequence similarity is more than 1. In addition, the alignment of the predicted protease region in ORF2 against other virus-encoded cysteine proteases was not sufficient to make any conclusions. No signal peptide and trans-membrane helix were predicted using SignalP 4.0 (Petersen *et al*., 2011) and TMHMM (Krogh *et al*., 2001). Further examination of the protease activity would be interesting to confirm function of the protein encoded from dsRNA2.

a) Schematic representation of the strategies used for cDNA synthesis of AfuTmV-1 dsRNA2



b) Genome organisation of AfuTmV-1 dsRNA2

c) Complete sequence of AfuTmV-1 dsRNA2 (2233 nt)

UGAAUAAAAGAUUUGUUCUGGGCAAGCACGCUGCAACGUCGAGCAACUCCCAUAUAGGGGACGAUGAGACAUGGCGGACCUAACCCGU
CUUAGGAAUAUUGUUCUGUCGGGCGAUAUGAAAGUCGCCCUCACUGCAGCAAUUCUGUACCACCUCUCCCCCAACCCCUCCGCGAUCA
CCGAUUACGGGGUGGAUGAGGUCAAGAGCGUGGUGUCAUGGCUCGUCUCCUCGCCCGGGAUGUUCUACGGCGACAUUGUUGCCGCCAC
UGGAAUCACCCCUGACAUCCUCUCUACUUUCCUUCCCGCGCUCGAUGAUCAAGCCUGCGCCGACGCCGCCUCUCUUGCCAACGCCAUG
GCGACCCCCGAUGCCGAGUGGGACCGCGAGCACCGCAACCGUUUCACGUCCGCACUCGCGCACCACUCCGCCGACGCUGCUAGGCUCG
CCGCCCAGUCGUCGGCUGAGUUCGCUGCGGUCGACGCCAUGGAAGAGGCUGGCUACGAGAAUCAGGACCACGCUCGCACCGCGGUCAA
CGAGCUCAACCGCGUCUAUGGCAAUACCUACCGGGUCUUCACUCGUCAACACGCCCGCUCCCGCGUGGCCAUCGCCGCCUCUUUCAAG
AUCGAUGACCACGGCACCCUCUUCCACGUCCUCACCCCCUGGUGCCUCUCCCGAACAUCGGCCAUGAUCUACGUCGCUGUCAUCGCCA
GGCAUGUGAGGCGUCGCCACAUCGAGGAGCGCGCCUCCAGAUUGGCACGCACUGCCCUGACCUACGCCGCUCGCGCCGGCUUAGAGGC
UGUACGCGCCAUCAUAGAGCCCGCCAUUGCCAUAGUGACAAGUUUCCGUUACGACCCUGCCAAGAUCGCCUUCGUGGACGAUUUUGAC
CGGGUCAGGGAGUCCAUCACUACGAGGAGUCCUCCAACGCUUGUCGCCUUCGCUUGCGUCCAUAUGUGCGGGAACCAGGACGUCGUUU
CCUUGGCGGCGCGACUCCGCAUGGAGUCGCUACCACAACUCCAUGAGACCCCGCCAGCCGUCGCCGAGUACACCGACCACGUCGACGA
CGAACAGUCCGCCCUCCUCAUGUUCUCUCGCGUCCUCAAGAGACUGAGACAUGAGCGCGACAGGGGGGUGCACGGCCCCCGCCUCAAC
AAGGAUCCCAUGCUCGGCUACAUAGCCCGGGGAGUACGCCGUUCCGAGAGGAACACCAUGGCGUGCAUCAACAGGGUUGAGGAGGCCG
UGUCCUUUCUGCGCUCGAAGGGGGUCGACACCCUCAAUACACUGUUCGUCAUCGAAUGGGGCGGUGAAUUUGACUAUCCUACCGUCAU
GGCUGCCCUCGCCGUUGCCAGGUUGGAUGUCUGUGUCGACAUUGGUCCCUCGGGCAUUGACCUACCCGGCGCCGACGCCAAGGCCGAG
GAUGGCACCGAAGAGUACAACUACUCUUUGCUGCUCGCCUCAGCCGAGGAGCGGCGCAUGCCUCGUAUGCCCAGGGUGCCCUAUCCGU
CCGACGAGCCCCUUGAGGCCAAACUGCGGCGCGUCGCGCAGUUCUACUCUCUCUCGGGCUCUGCCAAUAUAGCAUACGUCAGCGGGGG
GGUCACCCAAAACGAGGGGGUUCCUGUCAACAUCUGUAUCGACACCGCCAACAGGCUUAAUGCCAUACGCAAUGUCACCCCGGACGUU
AACGUCGUUUAUGCCUGCGCCGAAGCUCUCCUCCCCCCCCCUCUGCGGGCACGGACUCGAGCUCGGGCCCGACGACUACCUUCAGUGCG
CGUCCCUGGUCGAUGAGGAUUGCCCCGCCUGCAAUGCUCACUACCGCGCCACCGCCGUCCUCUCUGCUGCCUGCGCUGACAGACAGGU
GCGCAUCGUUAAGCCAAGAUCCGCCUACGCUCAUAACGGGCACCUGUCCUUGGAGCUCGUCCCGGGCCACCCUACAUGCGUGGACGAC
ACUCUCCUCACCGUCGACUCCAUGGUCGCGGCCAAUGUCCUCCGCAACCACGAUUGGGGCGAUGAGCCCGUUACUCCACCAGCCGGCG
GCAACCCGACAUCGCCUACCCUCGCCCGCGUCACAGCGGAGAACAUGCGCACUGUGUACUCCGCUCUCAAUAACGGCGCCGUGGGGCC
CUUCGACUAUGACGACAUGGCGUCAAUUGCCGCUUCAAUAGCGCCGUACACUGCGCUGCAUCGUGCUUGACCAGCGUCCGACGGGGAG
GGGCUUAAUUGCCCCCCCCGCCGGGCCUGGGGG

d) Amino acid sequence encoded from ORF2 (75.76 kDa)

MADLTRLRNIVLSGDMKVALTAAILYHLSPNPSAITDYGVDEVKSVVSWLVSSPGMFYGDIVAATGITPDILSTFLPALDDQACADAA
SLANAMATPDAEWDREHRNRFTSALAHHSADAARLAAQSSAEFAAVDAMEEAGYENQDHARTAVNELNRVYGNTYRVFTRQHARSRVA
IAASFKIDDHGTLFHVLTPWCLSRTSAMIYVAVIARHVRRRHIEERASRLARTALTYAARAGLEAVRAIIEPAIAIVTSFRYDPAKIA
FVDDFDRVRESITTRSPPTLVAFACVHMCGNQDVVSLAARLRMESLPQLHETPPAVAEYTDHVDDEQSALLMFSRVLKRLRHERDRGV
HGPRLNKDPMLGYIARGVRRSERNTMACINRVEEAVSFLRSKGVDTLNTLFVIEWGGEFDYPTVMAALAVARLDVCVDIGPSGIDLPG
ADAKAEDGTEEYNYSLLLASAEERRMPRMPRVPYPSDEPLEAKLRRVAQFYSLSGSANIAYVSGGVTQNEGVPVNICIDTANRLNAIR
NVTPDVNVVYACAEALLPPLCGHGLELGPDDYLQCASLVDEDCPACNAHYRATAVLSAACADRQVRIVKPRSAYAHNGHLSLELVPGH
PTCVDDTLLTVDSMVAANVLRNHDWGDEPVTPPAGGNPTSPTLARVTAENMRTVYSALNNGAVGPFDYDDMASIAASIAP

hypothetical protein [*Alternaria* sp. FA0703] Sequence ID: gb|ACL80752.1|

```
Query  71    MADLTRLRNIVLSGDMKVALTAAILYHLSPNPSAITDYGVDEVKSVVSWLVSSPGMFYGD  250
             MADLTRLR I+L     V    + +YHL+ +P +      + V  V++W+ ++PG F+GD
Sbjct  1     MADLTRLRTILLEHGSSVVPVGSAVYHLTKSPVSHVLVDANLVSEVLAWMAAAPGPFFGD  60

Query  251   IVAATGITPDILSTFLPALDDQACADAASLANAMATPDAEWDREHRNRFTsalahhsada  430
             ++++ GITP   +  LL D+  A+AA++ NAM   D EWDR  R+++T+A   H  DA
Sbjct  61    LISSAGITPADCAQLLEDLTDKQAAEAATMVNAMTIADDEWDRGTRDQYTTATVLHERDA  120

Query  431   arlaaqssaefaavdaMEEAGYENQDHARTAVNELNRVYGNTYRVFTRQHARSRVAIAAS  610
             A++AAQSSAEFAA+DA+  AGY   D A  A++ + R + + Y+V  RQ ARSR  + A
Sbjct  121   AKIAAQSSAEFAALDALTAAGYAGVDDAVGALSVMRRTFKHNYKVHVRQFARSRHCVVAR  180

Query  611   FKIDDHGTLFHVLTPWCLSRTSAMIYvaviarhvrrrhieerasrlartaltyaaraGLE  790
             F   D G +  V+TPWCLSR +A  Y A+      R  +E R    ART  T+A R  L+
Sbjct  181   FVRSD-GAVRIVITPWCLSRRNAQAYAALCMADWHRHMVEMRTINRARTVATFAERTSLQ  239

Query  791   AVRAIIEPAIAIVTSFRYDPAKIAFVDDFDRVRESITTRSPPTLVAFACVHMCGNQDVVS  970
             AVR II PA+A++T + YD   +AFVD   R  + +  R+P   VAFA V   G  D
Sbjct  240   AVRDIIAPALALLTRYSYDATNMAFVDGSGRPVQKVGARAPYVAVAFAAVFSSGAGDATL  299
```

```
Query  971   LAARLRMESLPQLHETPP--AVAEYTDHVDDEQSALLMFSRVLKRLRHERDRGVHGPRLN  1144
               AR R  +   L +        +Y +  D   + L +FSR L   R +R     G R+N
Sbjct  300   GLARARKAAADTLTDAVKLKTATDYLEMADSPFAGLALFSRALAEYRDKRHASPWGDRVN  359

Query  1145  KDPMLGYIARGVRRSERNTMACINRVEEAVSFLRSKGVDTLNTLFVIEWGGEFDYPTVMA  1324
              K  MLGY+AR   +  RNT ACINR+EE  + L  KG    +TL+V+EWGG      V+
Sbjct  360   KGAMLGYVARERDKLTRNTAACINRIEEIAALLAMKGHHKDSTLYVVEWGGRVSAHAVLG  419

Query  1325  ALAVARLDVCVDIGPSGIDLPGADAKAEDGTEEYNYSLLLASAeerrmprmprVPYPSDE  1504
              A A A++D+ +D+  SG+D+PG D  A+D    ++Y L LASA  R +PRMP VPY +
Sbjct  420   AAATAKIDIALDVAGSGVDVPGIDVYADDDDPPHHYQLYLASARRRMLPRMPSVPYRAGT  479

Query  1505  PLEAKLRRVAQFYSLSGSAN---IAYVSGGVTQNEGVPVNICIDTANRLNAIRNVTPDVN  1675
              PL  KL  V  + S+ G A+   +AYV+GG++Q    PV++C D+  R+ A+      +
Sbjct  480   PLIEKLSLV--YESVGGRASDYKLAYVNGGMSQLAETPVSVCTDSNARMAAMDGARLTLP  537

Query  1676  VVYACAEALLPPLCGHGLELGPDDYLQCASLVDEDCPACNAHYRATAVLSAACADRQVRI  1855
               + +   E LLPP C H +++ PD+YL      + +C  C AHYRA + +  A     VR+
Sbjct  538   IPFYSTEVLLPPACHHIVDMDPDEYLDDFGTGNPNCFQCEAHYRALSAVGDALDRPGVRL  597

Query  1856  VKPRSAYAHNGHLSLELVPGHPTCVDDTLLTVDSMVAANVLRNHDWGDEPVTPPAGGNPT  2035
               KPRS +AHN    LE + G  + ++D+L T+D+ +A N LRN+++ + P  P  G + +
Sbjct  598   TKPRSVFAHNALFGLEWI-GTDSAMEDSLETLDAAIACNALRNYEYPNPPPDPTPGKDLS  656

Query  2036  SPTLARVTAENMRTVYSALNNGAVGPF  2116
              SP L  +  +++R VY+ L     GP
Sbjct  657   SPDLLELMQDSVRQVYTLLAGPFAGPL  683
```

Fig. 20 Schematic representation of the strategies used for cDNA synthesis of AfuTmV-1 dsRNA2. A series of overlapping cDNA clones were generated by rPCR, RT-PCR and RLM-RACE using sequence specific primers (a). Diagrammatic representation of genome organization of AfuTmV-1 dsRNA2 (b); The AfuTmV-1 complete sequence (c); Amino acid sequence encoded from ORFs and the alignment against *Alternaria* sp. FA0703 hypothetical protein (d). Start and stop codons were shaded in pink and red, respectively. CAA repeats were underlined. Zinc finger-like motif was highlighted in gray.

Interestingly a zinc finger-like motif was detected in three putative proteins of the AfuTmV-1, AA1 and RZB2 (excluding C. cladosporioides virus 1) as shown in the alignment (Fig. 21). It has been described that proteins with zinc finger domains are associated with binding activity to DNA, ssRNA, dsRNA or protein and exhibit regulatory function important for virus replication and infectivity as same as reported in retroviruses such as HIV (Finerty and Bass, 1997; Boukhvalova *et al*., 2010). However, the significance of the zinc finger-like motif detected in AfuTmV-1 is still obscure. Thus, no function can be certainly ascribed to the putative protein encoded by AfuTmV-1 dsRNA2 thus far.

```
AA1                 RLTLPIPFYSTEVLLPPACHHIVDMDPDEYLDDFGTGNPNCFQCEAHYRALSAVGDALDR
AfuTmV-1 dsRNA 2    TPDVNVVYACAEALLPPLCGHGLELGPDDYLQCASLVDEDCPACNAHYRANAVLSAACAD
RZB2PEP             IPKERVQFITSEFIIPMPCEVHLEEDPEQFLSSFGRIEQECNDCLRYHVAQMSLLSFTGD
                              * ::*  *        :         . :*  *  .  :   :
AA1                 RLTLPIPFYSTEVLLPPACHHIVDMDPDEYLDDFGTGNPNCFQCEAHYRALSAVGDALDR
AfuTmV-1 dsRNA 2    TPDVNVVYACAEALLPPLCGHGLELGPDDYLQCASLVDEDCPACNAHYRANAVLSAACAD
RZB2PEP             IPKERVQFITSEFIIPMPCEVHLEEDPEQFLSSFGRIEQECNDCLRYHVAQMSLLSFTGD
                              * ::*  *        :         . :*  *  .  :   :
```

Fig. 21 Amino acid sequence alignment of zinc finger-like motifs of regions on proteins of unknown function putatively encoded by AfuTmV-1 dsRNA2, a 2233 bp dsRNA isolated from *Alternaria* sp. FA0703 (AA1; accession number ACL80752, and a *ca.* 2.3 kbp mycovirus dsRNA, nominated B2, isolated from the

flax rust, *Melampsora lini* (RZB2:Ren Zhang unpublished results). Asterisks signify identical amino acid residues, colons signify highly conserved residues and single dots signify less conserved but related residues. Zinc finger-like regions is shaded in grey.

2.5 Construction of a cDNA clone library of dsRNA1 by rPCR and RLM-RACE

Efforts have been made to complete the characterisation of the largest component of AfuTmV-1 namely dsRNA1, using random-primed PCR (rPCR) amplification and RLM-RACE procedure. Fractionated AfuTmV-1 dsRNA1 was used as a template for cDNA synthesis. The PCR amplicons generated appeared as a smear of products 200-400 bp in size on agarose gels stained with ethidium bromide. PCR amplicons were cloned and 17 putative recombinants containing inserts were sequenced. As a result, a 602 bp contiguous sequence was obtained from the sequence assembly of AfuTmV-1 dsRNA1 cDNA clones generated from rPCR (Fig. 22a).

To determine the terminal sequences of AfuTmV-1 dsRNA1, 3'-RLM-RACE was conducted in which a modified Lig primer was ligated to 3'-end of the strand. Sequence specific 5'- and 3'-specific oligonucleotide primers (5'AfuNK-dsRNA1; 5'-CACGTTTCCTGTAACCCACC-3' and 3'AfuNK-dsRNA1 (2); 5'-CCAGGTTGGGTTGACACC-3' and 3'AfuNK-dsRNA1 (4); 5'-AACTCCCTGTTCGCCAAACT-3') were designed from the assembled contig sequence of AfuTmV-1 dsRNA1 for RLM RACE. The RLM-RACE amplicons generated from the 5'-end specific primers was *ca*. 650 bp in size, whereas the 3'-terminal amplicon was *ca*. 400 bp and 550 bp in size (Fig. 22a and 22b). All amplicons were cloned into pGEM-T Easy vector. At least 6 RLM-RACE clones of each amplicon were sequenced.

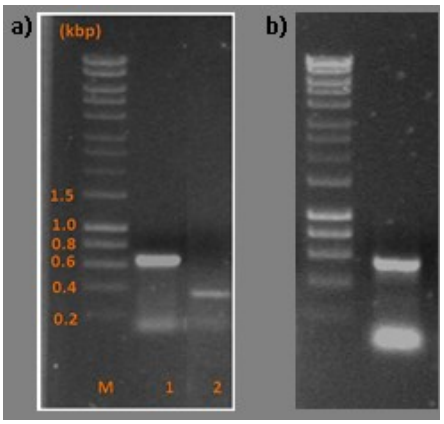


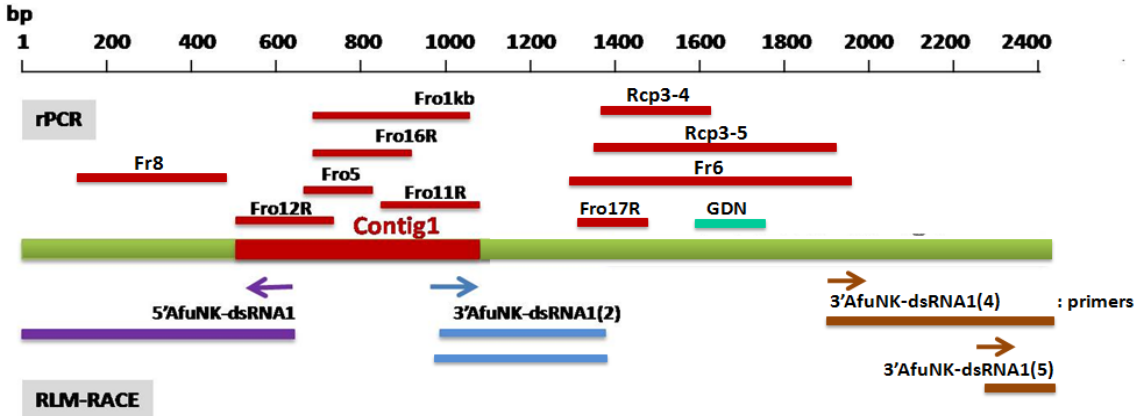
Fig. 22 Agarose gel electrophoresis showing the PCR products generated by RLM-RACE. RLM-RACE of AfuTmV-1 dsRNA1; Lane 1, the amplicons produced from RLM-RACE amplification of the 5'-terminus using 5'AfuNK-dsRNA1 as a primer; Lane 2, the amplicons produced from RLM-RACE amplification of the 3'-terminus using 3'AfuNK-dsRNA1(2) as a primer (a). The amplicon produced from RLM-RACE amplification of the 3'-terminus using 3'AfuNK-dsRNA1(4) as a primer (b). Hyperladder 1 (Lane M; 10 kbp; Bioline) was used as marker.

2.6 The genomic sequence of AfuTmV-1 dsRNA1

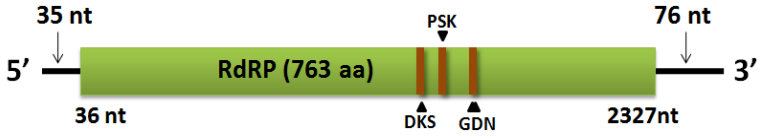
The contig sequences from rPCR and the 5'- and 3'-terminal sequences from RLM-RACE were aligned and assembled to obtain the sequence of AfuTmV-1 dsRNA1. A schematic representation of the complete genome of AfuTmV-1 dsRNA1 is shown in Fig. 23. The complete nucleotide sequence of dsRNA1 revealed that it is 2403 bp in size and contains a single deduced ORF1 from nucleotide position 36-2327, with the potential to encode a putative protein of 763 aa on the positive strand. The calculated molecular weight of the coding protein is 84.2 kDa and the genome contains 63% GC residues. No ribosomal frameshifting signal was predicted from the sequence.

Significant similarities were found once the deduced amino acid sequence of ORF1 was used to search against public database. Part of ORF1, amino acid position 411-553, aligned with PROSITE (Falquet *et al*., 2002) profile PS50507 (RdRP of positive ssRNA viruses catalytic domain, score = 8.748) which the sequences of all positive ssRNA viruses, two small families of dsRNA viruses (the *Partitiviridae* and the *Totoviridae*) and some eukaryotic RdRPs are known to belong to this RdRP class. Search against the Pfam protein database (Finn *et al*., 2008) identified domain Pfam00680 in the ORF1 (RdRP_1, *E* value = $1.2e^{-03}$). A search of Conserved Domain database (Marchler-Bauer *et al.*, 2013) identified the following domains at the nucleotide position 417-599 of the AfuTmV-1 dsRNA1, showing the alignment in the same region of the RdRP superfamily member cd01699 (RT_like, Reverse transcriptase, *E* value = $3.08e^{-13}$) and Pfam00680 (RdRP_1, *E* value = $2.47e^{-06}$). A PSI-BLAST and BLASTX search of GenBank nr database showed the best match to putative RNA-dependent RNA polymerase of Cladosporium cladosporioides virus 1 (YP_009052470.1, 46% identity, *E* value = 0.0) and putative RNA-dependent RNA polymerase of Walrus calicivirus (ssRNA positive-strand viruses, *Caliciviridae*, NP_786919.1, 29% identity, *E* value = $5e^{-06}$). No signal peptide and trans-membrane helix was predicted. Taken together, it can be concluded that AfuTmV-1 dsRNA1 encodes an RdRP typical of positive-sense ssRNA viruses of RdRp subgroup 1. Interestingly, the presence of the conserved RdRP motifs of the AfuTmV-1 showed a unique GDN RdRP motif instead of GDD motif as commonly found in viruses. However the function of the AfuTmV-1 RdRP is still unknown. It could be primarily assumed that the enzyme probably form a fusion protein complex with a structural protein encoded from the sequence or from other dsRNAs for replication.

a) Schematic representation of the strategies used for cDNA synthesis of AfuTmV-1 dsRNA1



b) Genome organisation of AfuTmV-1 dsRNA1



c) Complete sequence of AfuTmV-1 dsRNA1 (2403 nt)

CGAAUAAAAGGAAACUUCUCCGCUUGUGUGCUCCAAUGUCGGCCAAUCCCGAUAUAGGGGUCCCAGAGACCUUGAAACUCGUCGUCAG
CGGUGCGCCCAGCAACGCUCCCGCCCCUACCGCCUCCUACGUCUCCGACAUUGUCUCGUGGUACGCCCGCCGCAACCUCAAGGGUGGC
UUACUGCAGGUCGAUCACACCAUGUCUUUCGUCGUCUGGCCAACCUCCGCCCCCAUCUCCGUCCCGGUCCACGGUGUCACCACUGGUU
CCGAUCCCGUCGUCACGGCCCCACAAGCGCGCGCCCGCGCCUUUCGGCUCAAGUCCUAUGCUGAUGGGGACUUGUUGCGCGAACUCAG
GCAGUUCGAUGGUGCCAAAAACGCCAGCGAGCUCGGGCCCCGUGCCUUCGUCCGCGACUCUCACCGUCUCGACGCUGCUGUGGAAAAG
GCCUCCAAACUUACGGACCCGACGUUCCACAUCUCGCAGUACCAGCUCCCCCACCCCUACUCCUUUGGCGGCGGACCCCCCAACCCUG
AGAGGCCCCUCACCGCUCCUCUAAUCUCGGCGAUUAACAAGGUCUCCCAGCGCACCCGCGACCCGGUGGGUUACAGGAAACGUGCCAA
GGAGUCCAUUGACCUCGGCGACUUCACCACUCAUGACCCGGACACGCUCCACCCGCGCUUCCUCGAGUAUGUCCAUGAGCGCACGCGC
AGUGUCGACGGGCCCACCGAUGACGCCAUGCGUGCCGCCCAGACCGUCUUCGCCCGGCUUUGGCGAAGGAAGGGCUGCAAGGUCAAGG
CCAGGAGCCUCUCCGAUGCGCAGCCUGACAAUCUCCUUGCCAUCAUAAAGAAGGGGUCGCCGGGGGAAUAUCGCUCACUCGGCGCCGA
GGACCGACGUGACCCCCGCCUCAUCGCCACCAUGUCCUCCUCCCUCCUUCGGUACGCCAGCGCGGGGGUCCAGGUGGCGCGCGGGCGC
CCCCCCCCAGGUUGGGUUGACACCACGACCCAGGUCACGCUCACGUUCGGCAAGCGCGAGCCUAAAGCAGCCAAGAUCGUCGACGGAG
UUCGCCAAGCGCCCGUCCCGCGCUUUAUAUUCAACCUUUCACCGGUGAACUACGCCCUCGCCAGCUUCCUCCACUACGACAUUUCCCA
CUUCCUCAUGGACAACGACCCCACCCACGGACCUGGGUUCGGACCCGGCAGGGGCAGGGCGAGGAAAUUCAUGGACCUGGUUGAGAGG
GCUUUCGACGGGCGCUUCUCCACUCCCGACGGAGCGCGGCUCAUCAUGAGCGACAUUACCAAGUGGGAUGCCAACAUGUGCGAGGCGC
UCAUCAAGUACUCCAUUGACCUCCUCGAGGACGCCGUCGACAAGUCUGCAUUGUCCCCUGAGGGCCUCGCUACGCGGGGCCUGAUGUA
CAGGGUCGCCCGGCGGCAACUCCUUGAGAAGCUCGUCGAGCAUCCCGCGGGGUACUUCGUCAAGCUCUACGGGUGCAUGCCAAGCGGU
UCCUUCUACACCUCGCUUGUCAACACUACGGGGAACAACCUCUUGGUCAUAGGUCACGCCAUAGCACGGGCCGUCGAGGAGACCUCAC
UUACGCACCAUGGCGCAGCCGAGCUUCUCGCCGAUGCGGUCGAUGGCACUCUCAUCUCUUACGGAGACAACCAGCUCUUCUCUGAGCA
CCUCUUCUCAGUGCUGGGCCUCGCCUACGACCCUGAAAAGCAUGCCGAGUUCCUCGCUCGUUUCGGCAUGAAGCUUAAGGUGGAUGAG
ACUGAGGUCACUGUCAAGCUUGGGCGCGUGCGGUUCUGCUCCCGGUCCUUGGUCAGGACGCCACAUGGUCUUCUCAUUACGCGCAGCC
ACAACUCCCUGUUCGCCAAACUCGCCGGACGGCCCCGUCACGACCCAGUAGUCGACAAGCUCUACGUGCGCGCAAUGAUGGUCGACCA
CAUGGGCACAGACCCCAUUGUCUACGCCAUCCUGAAUGAGAUUGACCGGUCCCUCAAUGUCUCCCUCGAGGCCGCCGGUCUCACCGAC
GCCGCCAAGAAAGUCCUCGAGGACACGGCCCAGAGCAUGUUCGGGAACCGCGAACAGGACGCGCUGCUCGCCGUCUACAGGGCGCUCU
CCGAGACCGUCAUCGACCGGCGCGCCCUGCUCUCGCUCCACACCCCACGAGACGGCGACCACGACCCGGGACGCCUGCACACCUCGGU
CUCCACCGGCAUGCACCUCUUCACCGGGGAGCUGACGCCCGCCGCUCAGUGGGCUUAUGAAUGUACCGUUGAGAAGUGGUGUCAGUAC
CUCCACGACACUGAUCAGGAGGGCGUCAUGUUCGACUAAUCACCGCACACAAGCCGGCCCGGCGGGAGGGGGCGCGCGUUGUCGAUCC
GCGUGCCCCCCCCGCCGGGGCGGGGGG

d) Amino acid sequence encoded from ORF1 (84.2 kDa)

MSANPDIGVPETLKLVVSGAPSNAPAPTASYVSDIVSWYARRNLKGGLLQVDHTMSFVVWPTSAPISVPVHGVTTGSDPVVTAPQARA
RAFRLKSYADGDLLRELRQFDGAKNASELGPRAFVRDSHRLDAAVEKASKLTDPTFHISQYQLPHPYSFGGGPPNPERPLTAPLISAI
NKVSQRTRDPVGYRKRAKESIDLGDFTTHDPDTLHPRFLEYVHERTRSVDGPTDDAMRAAQTVFARLWRRKGCKVKARSLSDAQPDNL
LAIIKKGSPGEYRSLGAEDRRDPRLIATMSSSLLRYASAGVQVARGRPPPGWVDTTTQVTLTFGKREPKAAKIVDGVRQAPVPRFIFN
LSPVNYALASFLHYDISHFLMDNDPTHGPGFGPGRGRARKFMDLVERAFDGRFSTPDGARLIMSDITKWDANMCEALIKYSIDLLEDA
VDKSALSPEGLATRGLMYRVARRQLLEKLVEHPAGYFVKLYGCMPSGSFYTSLVNTTGNNLLVIGHAIARAVEETSLTHHGAAELLAD
AVDGTLISYGDNQLFSEHLFSVLGLAYDPEKHAEFLARFGMKLKVDETEVTVKLGRVRFCSRSLVRTPHGLLITRSHNSLFAKLAGRP
RHDPVVDKLYVRAMMVDHMGTDPIVYAILNEIDRSLNVSLEAAGLTDAAKKVLEDTAQSMFGNREQDALLAVYRALSETVIDRRALLS
LHTPRDGDHDPGRLHTSVSTGMHLFTGELTPAAQWAYECTVEKWCQYLHDTDQEGVMFD

e) Amino acid sequence similarity of AfuTmV-1 dsRNA1 and C. cladosporioides virus 1 RdRP

```
Query  81   VTAPQARARAFRLKSYA-------DGDLLRELRQFDGAKNASELGPRAFVRDSHRLDAAV  133
            V+AP    ++FR    A       + + L  LR+FD   +  +   R+F RDSHRLD AV
Sbjct  71   VSAPSPVPQSFRTTRAALTRIAGTEKEKLESLRRFDTPSD-EKTSRRSFTRDSHRLDDAV  129

Query  134  EKASKLTDPTFHISQYQLPHPYSFGGGPPNPERPLTAPLISAINKVSQRTRDPVGYRKRA  193
             KA+ L DP+F I   Q   P SFGGGPP  ERPL    +    ++++       YR+
Sbjct  130  LKAASLKDPSFRIYPLQRETPLSFGGGPPGEERPLKPTGLVTAGRMAKAAGTDELYRREM  189

Query  194  KESIDLGDFTTHDPDTLHPRFLEYVHERTRSVDGPTDDAMRAAQTVFARLWRRKGCKVKA  253
              ++ +GD TTHDPD+LHPRFLEYVHERTRSV    + A+  A TV   LWR +G + +A
Sbjct  190  ATALKMGDSTTHDPDSLHPRFLEYVHERTRSVPEKENKALAWAVTVMKELWRNQGIRAEA  249

Query  254  RSLSDAQPDNLLAIIKKGSPGEYRSLGAEDRRDPRLIATMSSSLLRYASAGVQVA-RGRP  312
            R +  A+PD L +++  GS GEY  +GA +R+DP ++  +S S+ RY  AG  +A RGR
Sbjct  250  RGIDSAEPDVLSSLMNPGSAGEYAEMGATNRKDPEVLELLSKSIKRYYKAGHAMATRGRR  309

Query  313  PPGWVDTTTQVTLTFGKREPKAAKIVDGVRQAPVPRFIFNLSPVNYALASFLHYDISHFL  372
             P W D T Q  ++FGK+E KAAK+V+GVR  PVPRFIFN SPVNYALA+FLH D+SH L
Sbjct  310  AP-WADFTQQPVMSFGKKEAKAAKLVNGVRTPPVPRFIFNPSPVNYALAAFLHGDLSHQL  368

Query  373  MDNDPTHGPGFGPGRGRARKFMDLVERAFDGRFSTPDGARL----IMSDITKWDANMCEA  428
               DPTHGPGFGPGRG+A KF+D V          P  A L    IMSDI KWDANM EA
Sbjct  369  QTRDPTHGPGFGPGRGKAWKFLDKVAAHL-----LPGKAELSCKAIMSDIAKWDANMSEA  423

Query  429  LIKYSIDLLEDAVDKSALSPEGLATRGLMYRVARRQLLEKLVEHPAGYFVKLYGCMPSGS  488
            L+  + DL+E  VDKS+L   G ATR +M  VA+RQL+ KL+EHP+GY ++L+GCMPSGS
Sbjct  424  LLSATFDLMESVVDKSSLDATGRATRRIMADVAKRQLMVKLIEHPSGYLLELFGCMPSGS  483

Query  489  FYTSLVNTTGNNLLVIGHAIARAVEETSLTHHGAAELLADAVDGTLISYGDNQLFSEHLF  548
            FYTS VNT GN+LL +       +E+       +   +A      L+SYGDNQL  + LF
Sbjct  484  FYTSCVNTIGNDLLALSLLGVTLMEQGVELSDVSPISVAQQASSDLVSYGDNQLIFDSLF  543

Query  549  SVLGLAYDPEKHAEFLARFGMKLKVDETEVTVKLGRVRFCSRSLVRTPHGLLITRSHNSL  608
            S  G++Y  E+H   LA FGMKLKVDET V+ +LG VRFCSR  + TPHGL I RSH S+
Sbjct  544  SRFGVSYSLERHEAHLAAFGMKLKVDETGVSSQLGDVRFCSRGALLTPHGLAIVRSHTSI  603

Query  609  FAKLAGRPRHDPVVDKLYVRAMMVDHMGTDPIVYAILNEIDRSLNVSLEAAGLTDAAKKV  668
            F K+ GR   DPV++KLYVRA+MVD +GTDPI+Y  L  ++ S++V  +A        K+
Sbjct  604  FHKIGGRAEVDPVINKLYVRALMVDLLGTDPILYHGLQHLEASIDVPGDARLSEHRLHKM  663

Query  669  LEDTAQSMFGNREQDALLAVYRAL-SETVIDRRALLSLHTPRDGDHDPGRLHTSVSTGMH  727
            +E  A+ ++GN   +++ +    L S +  DR  LLSL  P D      R  T ++ G
Sbjct  664  VEPFAKKLYGNSSPESVASFVSLLKSPSAPDRGVLLSLRLPADAPGALKRFGTGLTLGSG  723

Query  728  LFTGELTPAAQWAYECTVEKWCQYLHDTDQEGVMFD  763
              T  L    +W  + +   + +Y++DT Q  ++++
Sbjct  724  KATTGLDDVGRWLLDLSPADYWKYVNDTGQTSIIYN  759
```

Fig. 23 Schematic representation of the strategies used for cDNA synthesis of AfuTmV-1 dsRNA1. A series of overlapping cDNA clones were generated by rPCR, RT-PCR and RLM-RACE using sequence specific primers (a). Diagrammatic representation of the complete genome sequence of AfuTmV-1 dsRNA1 (b); and the sequence (c); Amino acid sequence encoded from ORF2 (d); and an alignment of the amino acid sequence similarity of AfuTmV-1 dsRNA1 and C. cladosporioides virus 1 RdRP (e). The putative start and stop codon of an ORF is shaded in pink and red, respectively.

2.7 Construction of a cDNA clone library of dsRNA3 by rPCR and RLM-RACE

As described previously random-primed PCR method (rPCR) was employed to make cDNA clones of the unfractionated AfuTmV-1 dsRNAs. Following the PCR amplification step amplicons, 200 to 1000 bp in length, appeared as a smear on agarose gels as shown in Fig. 15. Sequence analysis revealed that clones 9 and 12 generated from rPCR (*ca.* 830 and 900 bp in size) belong to parts of dsRNA3. RLM-RACE was also used to generate cDNA which yielded five

different sizes of clones (race5b, race3a, race2d, race2b and race2a) ranging in size from 500-750 bp. Attempts to generate a complete cDNA clone of dsRNA3 sequence were performed. However, RLM-RACE products obtained were smaller than the actual size of dsRNA3. It is possible that Lig-modified reverse primer and Lig forward primer preferentially attach to fragmented dsRNA to generate amplicons.

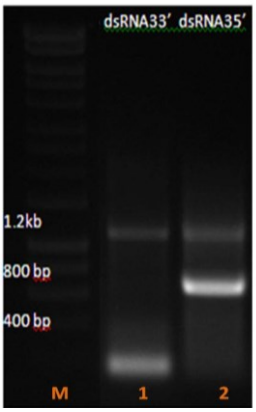


Fig. 24 Agarose gel electrophoresis showing the PCR products generated by RLM-RACE. The amplicons produced from RLM-RACE amplification of the 3'-terminus of AfuTmV-1 dsRNA3 using 3'AFUNK-DSRNA3 as a template (Lane 1); the amplicons produced from RLM-RACE amplification of the 5'-terminus using the 5'AFUNK-DSRNA3 as a template (Lane 2). Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

The sequences of the 5'-and 3'- termini of AfuTmV-1 dsRNA3 were extended and confirmed using RLM-RACE. The sequence specific oligonucleotide primers, 5'AFUNK-DSRNA3; 5'-GTACAGGGGCTGGTCATCA-3' and 3'AFUNK-DSRNA3; 5'-CATCATGGCGGAGCAACTA-3', were designed to amplify the 5'-and 3'- termini respectively. Using the 5'AFUNK-DSRNA3 primer, two differently sized 5'- specific amplicons *ca*. 750 and 1200 bp were obtained whereas amplification with 3'AFUNK-DSRNA3 primer yielded 3'-specific products of *ca.* 200 and 1200 bp (Fig. 24). All the products generated from each of the above two reactions were individually gel purified, cloned and sequenced.

2.8 The genomic sequence of AfuTmV-1 dsRNA3

The genetic organization of AfuTmV-1 dsRNA3 is shown in Figs.25a and 25b. All clones were aligned and assembled to obtain the complete sequence. The sequence shows two possible ORFs. The first AUG codon in ORF3 (nt 28-30) is located in an unfavourable translation initiation region. The second AUG codon of dsRNA3 (nt 52-54) is a much more favourable codon for translation initiation because the codon is surrounded by an A at -3 position and a G at +4 position (Kozak, 1986; Lutcke *et al*., 1987). The analysis showed that AfuTmV-1 dsRNA 3

comprises 1970 bp with 5' and 3' UTRs (of 51 and 74 bp in length, respectively). These UTRs flank a single ORF which encodes a putative protein of 614 amino acid residues (calculated molecular weight 66.66 kDa) which contained a high proportion of positively charged amino acid in an APC basic domain. The GC content calculation revealed that the genome contains 64% GC residues. No ribosomal frameshifting signal was predicted from the sequence.

PSI-BLAST and BLASTX search of GenBank nr database and SUPERFAMILY showed that AfuTmV-1 dsRNA3 shares 40% amino acid sequence similarity to a dsRNA-encoded protein found in another M. lini mycovirus dsRNA element, 1932 bp long, nominated B3 (Dickenson *et al*., 1993; accession number X64371; 54% similar; *E* value = $9e^{-120}$) and hypothetical protein from unclassified Cladosporium cladosporioides virus1 (Yu *et al*., unpublished; 28% similarity; *E* value = $9e^{-07}$), respectively. In addition, Part of ORF3, amino acid position 136-253, aligned with Pfam protein database identified domain PF12847.2 (Methyltransferase_18, *E* value = $4.5e^{-09}$). A search of Conserved Domain database also identified the domain in the same region of cd02440 (S-adenosylmethionine-dependent methyltransferases (SAM-MTases or AdoMet-MTases), class I, *E* value = $1.06e^{-05}$) at positions 131- 253 which is written in red (Fig. 25d). No signal peptide and trans-membrane helix were predicted.

a) Genome organisation of AfuTmV-1 dsRNA3

b) Complete sequence of AfuTmV-1 dsRNA3 (1970 nt)

CGAAUAAAAGGAAACUAGAACGCUUGCAUGCUCCAACGCCCCGCAACCACCAUGCUAGCGAGGAAAAGGGGAGUGCGGUCCGUCCCGU
CAACCGCAUCAAGGGGGCCUCCACCAUCGACGAGGCGACCAUCUGAGACACCGUCCCACCCGUCAACUUCGGGGGUGUCGAACCUGGC
GCUCUCGGUGCGGUCCCGCUCCACCACCACAGUCGCCGCCUCUGGACCAUCGGGCGAAAAGGUCGUUCUGCCCUUCUCGCUGUUCGAG
UUCGGCUUCCCACCCGACAUGCCUUCCUCUCCGCUACCGGAGGACUCUCAAUUCAAGGAUGAGGGUGGCCAGCGGUUCAUGGCGAGCG
AGGCUGGUACAAUCCUGCGCAGGGAUGAUCGGAGGUACGCUGCGCUACAGCGCCAAUUCAUCGUGCGCAGCGUCCGCCUUACCGGCUC
UUCCAUCCUGGUCCUCGGGUCCGGGUCGUCGAAAUCCAUAAUUCCCCUCCUCAGGAGGGGGGUCGCAACCGCCACCUUCGUCGACACC
UCCCAGGCUGCCCUCGACCGCAUGAGGCGCAACCUCACCGAGGCCGGGAUAACCGCCACCGUCGACGCCCAGUUCGUCUGCUGCGACG
CCUGGGAUUGGAUCUCCGGUGAUGACCAGCCCCUGUACGACGUCGUCAUCGCGACCAAGUGCCUCGGGCUGAUCUUCUCUAGCGACCC
GGCCCAGCGCGACGUGCAGUCGCUCCUCGACUACUGCUCUGCCAUCCUCCGGGAUGACGGAUCCGUGUUCGUCGACCACCACCUCGCU
UUCGCCUCCCCUUCCCCACGGUACGAGGGUCGCGUCCGCAGUAGAGCCCGAGUUGUUCGACCUGGCGACCAUUGCUGGCCGCUACGCUG
ACGACGUCGCGUACAAUGCUGAGGUUGACCACCAGGACUUCGAUCGGGUUGCCUCCUUUGUCUCUUCUGCGGCCGCGCACCUUGUCCA
GGUGUGGCAGUUCAUCCAUUACCGCCUCAAGAACGUCGGCCGAGCCAAGCCCGGGGCAACGUUCUCGCUCCAUAAAGCGCCAUGCCCC
ACGGAAUUCCCGACACCGCCCGCAUUGGAAUUCGACGCCCUCGCUGAUGCCAUGUACCCCGUCAACGGGCGUGGUGUCAAGCGCAUCC
CCACCGCCUCUGACAUCAAGGGGCAUCCUUACGCCACCGCGCUCGUCAAGUAUGAUGGUGAGCAUGGGGUCCUUGUGCUCGAUGGCGC
UAACGCCACCUUUAUAUCUGGCCGUUACCGCUUCGCACGCCAGCUGCACCUCUCAGUCCAGCCCGUUCUGGCCUGUACUGCCGAACUG
GUGUCAGUUUCCCCGCAGACCUCCGUGCUCAUCAUCACGGGGCUUAUCUCACUCGGCGAUGCCUUCGCCGACCCGCUCGAUUACGAGG
CCCUCCGUCCCCUCGUUCCCACCCUCGAGCGCCUUGCCCCAUCCGGGAUCGUUCCCACGAUCCCCGAGCACGUUCGGCUAGUCAAGGG
CUCCGCGGUCCACUUCUUUGGUCCCCACGGGUCAGUCCUCCGCGCCCCCGUUGACGGUGUCCAAGUCAACACCGGCGGCAAGGCUGGU
GUGUUCAUCAAGCCUGCCGCUGCCUGCACCGUCGACGCCACGACCAAUGAUGCUUGUGACGCCAUCGCUGAUGCAUACGCUGCCCUUG
GGCUUUCCGCGAUGCCCUACGUCCAACCAGGGGAGGGGGACAACAUAUACGAGUUCCACCGCGUUCCGGGAACGCACGUCUGGCGCCC
AGGACGCCCACGGCCUGACAAGAACCGCAGCGAUAAGCCUGGCACCGUGGUCCACACGGUCGCUGCCAGCAUCAUGGCGGAGCAACUA
GCCCUCACCAGUGACGUCGCCGCCCUCGUGUCCAAGAUCUUCCGGUGAGUGGAACGUUGGAGAUGCUCGCGUCCCGACGCGGCGGGGG
GGGCCCUGGCCCCCCCCCCGCGCCGGGGCGGGGG

c) Amino acid sequence encoded from ORF3 (66.66 kDa)

```
MLQRPATTMLARKRGVRSVPSTASRGPPPSTRRPSETPSHPSTSGVSNLALSVRSRSTTTVAASGPSGEKVVLPFSLFEFGFPPDMPS
SPLPEDSQFKDEGGQRFMASEAGTILRRDDRRYAALQRQFIVRSVRLTGSSILVLGSGSSKSIIPLLRRGVATATFVDTSQAALDRMR
RNLTEAGITATVDAQFVCCDAWDWISGDDQPLYDVVIATKCLGLIFSSDPAQRDVQSLLDYCSAILRDDGSVFVDHHLAFASLPHGTR
VASAVEPELFDLATIAGRYADDVAYNAEVDHQDFDRVASFVSSAAAHLVQVWQFIHYRLKNVGRAKPGATFSLHKAPCPTEFPTPPAL
EFDALADAMYPVNGRGVKRIPTASDIKGHPYATALVKYDGEHGVLVLDGANATFISGRYRFARQLHLSVQPVLACTAELVSVSPQTSV
LIITGLISLGDAFADPLDYEALRPLVPTLERLAPSGIVPTIPEHVRLVKGSAVHFFGPHGSVLRAPVDGVQVNTGGKAGVFIKPAAAC
TVDATTNDACDAIADAYAALGLSAMPYVQPGEGDNIYEFHRVPGTHVWRPGRPRPDKNRSDKPGTVVHTVAASIMAEQLALTSDVAAL
VSKIFR
```

d) Amino acid sequence similarity of AfuTmV-1 dsRNA3 and *M. lini* SP6 unnamed protein product

```
Query  191   SRSTTTVAASGPSGEKVVLPFSLFEFGFPPDMP-SSPLPEDSQFKDEGGQRFMASEAGTI  367
             S S T  +A      K+ +P+ +FEF FP  +  ++ LP DSQF D     FM SE G
Sbjct  44    SDSKTVRSAKDGQSLKMDVPYPIFEFDFPVKIAHNNKLPVDSQFSDPIVDAFMCSEPGQT  103

Query  368   LRRDDRRYAALQRQFIVRSVRltgssilvlgsgssksiiPLLRRGVATATFVDTSQAALD  547
             LRR++R Y A  RQFI R+ R+TG ++L +GSG+SK I+ +LRR   + TF+D S AAL+
Sbjct  104   LRRNERIYNATLRQFIYRNSRITGVNLLFVGSGASKEIMSILRRMPQSVTFLDISSAALN  163

Query  548   RMRRNLTEAGITATVDAQFVCCDAWDWISGDDQ-PLYDVVIATKCLGLIFSSDPAQRDVQ  724
             R+R  +  +G++A +D  +V  DA +W+    Q  ++DVVIATKC+G I          +
Sbjct  164   RLRVAIDASGLSAPIDISYVASDATEWLHTTGQEKVFDVVIATKCIGQIMRR--PGTSYE  221

Query  725   SLLDYCSAILRDDGSVFVDHHLAFASLPHGTRVASAVEPELFDLATIAGRYADDVAYNAE  904
              L+   S +L+ +G V+ DHH A  S P  T + S V P  +D+ATI GRYADD AY
Sbjct  222   YLMLGISNVLKPNGYVYADHHDALISAPENTPIGSCVSPSDYDVATICGRYADDRAYYCF  281

Query  905   VDHQDFDRVASFVSSAAAHLVQVWQFIHYRLKNVGRAKPGATFSLHKA-----PCPTEFP  1069
                 DF  VA+F +      VQ W     R     RA      SL+ A     P P  FP
Sbjct  282   TPVPDFTIVAAFTAKTTPTRVQQWTSFCIRYLPSVRADQ----SLYMANTAYNPKPRGFP  337

Query  1070  TPPALEFDALADAMYPVNGRGVKRIPTASDIKGHPYATALVKYDGEHGVLVLDGANATFI  1249
              PPAL +D +A+ + PVN RGVKRIPT  D   +       K+DG  G+L++ G  A F+
Sbjct  338   RPPALAYDEIAELLIPVNSRGVKRIPTPDDKLRYKLNAVFPKFDGTPGILLMHGNAAVFV  397

Query  1250  SGRYRFARQLHLSVQPVLACTAELVSVSPQTSVLIITGLISLGDAFADPLDYEALRPLVP  1429
             SG YRF   +  SV   L+   ELV        +++TG++ +GD   DPLD   LR ++P
Sbjct  398   SGTYRFEIAISYSVDTNLSMACELVESLEGLFYVVVTGIVEIGDNTTDPLDANTLRAVLP  457

Query  1430  TLERLAPSGIVPTIPEHVRLVKGSAVHFFGPHGSVLRAPVDGVQVNTGGKAGVFIKPAAA  1609
              L   A +GI+   P+ VR ++G  V      GS +R PVDG+ V   GK G FIK
Sbjct  458   LLSPYAENGILVNTPDLVRNIRGDQVVLTDYRGSTMRLPVDGLHVAIDGKNGTFIKSLPF  517

Query  1610  CTVDATTNdacdaiadayaaL-GLSAMPYVQ-PGEGDNIYEFHRVPGTHVWRPGRPRPDK  1783
             CTVDA  ++  + +A AY A  G+  +P V    EG++I+E+ R P T  W P R R DK
Sbjct  518   CTVDAAASEIKELLASAYKAADGVYNVPVVSGEVEGEDIWEYARAPSTDNWYPVRKRIDK  577

Query  1784  NRSDKPGTVVHTVAASIMAEQLALTSDVAAL  1876
               SD  G+V+HTVA+S  A +L   S V  L
Sbjct  578   RYSDPLGSVIHTVASSYAAGRLGAESTVEQL  608
```

Fig. 25 Diagrammatic representation of genome organization of AfuTmV-1 dsRNA3 (a); its complete sequence (b); amino acid encoded from the ORF3 (c); and its similarity to the *Melampsora lini* SP6 unnamed protein product (d). The region of S-adenosyl methionine-dependent methyltransferases (SAM or AdoMet-MTase) was written in red and protein region involving in 5'-capping of RNA is shaded in gray. Start and stop codons are shaded in pink and red, respectively. CAA repeats are underlined.

Since a methyltransferase domain is readily detectable on the dsRNA3, it is assumed that the molecule encodes a capping enzyme. Methyltransferase catalyses the transfer of methyl group to ensure nucleic acid and protein stability. The stability of mRNA, which usually occurs in nucleus in the eukaryotes, appears to be dependent on methyltransferase *via* capping of the strand to prevent 5'-3' exonucleolytic degradation and initiate efficient translation (Bokar and

Rottman, 1999). In eukaryotic cells, capping of mRNA is also crucial for splicing of mRNA, export of mRNA from the nucleus and non-self recognition of foreign RNAs for example viral transcripts. Similarly, some viruses that replicate in cytoplasm have evolved a capping mechanism to protect the 5'-RNA sequence. In addition, capping appears to be highly diverse activity, dependent on virus genetic components and protein organization.

During capping of all known eukaryotic and some viral mRNAs, the guanylyltransferase (GTP-mRNA guanylyltransferase; GTase), which is reversibly and covalently guanylylated, transfers a GMP of GTP to a 5'-diphosphate terminus of the transcribing mRNA (Ogino and Banerjee, 2007). Then, methyltransferase adds a methyl group to generate a 7-methylguanosine (m^7G) residue at the 5' end of RNAs to form a (m^7G[5']ppp[5']N-) cap structure. The RNA-capping methyltransferase domain is found in a broad range of viruses particularly in ssRNA viruses. The enzyme is conserved in the ssRNA alphavirus-like superfamily and present in some but not all picorna-like (+) ssRNA mycoviruses e.g. Botrytis virus F (Howitt *et al*., 2001) and some endornaviruses (Roossinck *et al*., 2011). In addition, capping enzymes have also been found in non-segmented (-) ssRNA viruses such as rabies, measles and Ebola. RNA capping has been reported in animal dsRNA viruses e.g. Reoviruses, however, mycoviruses with dsRNA genomes are not naturally capped but do furnish ssRNA transcripts with a cap structure by a novel cap-snatching mechanism from the host in at least one case in the *Totivirus*, Saccharomyces cerevisiae virus L-A (Fujimura and Estaban, 2011). In the GTase enzyme-GMP intermediate, GMP is linked to the epsilon-amino group of a lysine residue *via* a phosphoamide bond. This principal lysine residue has been identified in the AfuTmV-1 dsRNA3 ORF sequence as part of a characteristic catalytic motif KXDG(I/L) in the GTase for viral GTPs which is also present in *M. lini* mycovirus dsRNA B3, several DNA viruses including vaccinia virus, African swine fever virus (Fig. 26) and in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Decroly *et al*., 2012). With no clear relationship with the capping enzymes of any other RNA viruses the potential origin of AfuTmV-1 is unexpected.

In conclusion, it can be inferred that the protein encoded by AfuTmV-1 dsRNA3 might be associated with methyltransferase capping activity through either conventional or unconventional capping pathways. In addition, two interesting aspects were identified in this study that (I) AfuTmV-1 is found to be the first virus that combines a capping enzyme with a picorna-like RdRP, a striking case of chimerism; (II) no clear relationship with capping enzyme of any other RNA viruses was found and the sequence with strong similarity in the methytransferase domain in this study comes from bacteria and DNA viruses.

```
Vaccinia                        ENLYAVTKTDGIPITIRV
African swine fever virus       VGYYVTDKADGIRGIAVI
AfuTmV-1 dsRNA3                 PYATALVKYDGEHGVLVL
M. lini mycovirus dsRNA B       DLSTALPKFDGIAGTLII
                                    .  * **      :
```

Fig. 26 Amino acid sequence alignment of regions on proteins putatively encoded by AfuTmV-1 dsRNA3, *Melampsora lini* dsRNA B (accession number X64371), Vaccinia virus and African swine fever virus involved in 5'-capping of RNA. The principal lysine residue (K) is shown as part of a characteristic catalytic motif for viral GTPs. Asterisks signify identical amino acid residues, colons signify highly conserved residues and single dots signify less conserved but related residues.

2.9 Determination of the 5'- and 3'-termini of dsRNA4 using RLM-RACE

The complete sequence of AfuTmV-1 dsRNA4 (*ca.* 1.1 kbp in size) was obtained using RLM-RACE which was amplified using the combination of Lig-modified reverse and Lig forward primers as previously mentioned (Fig 19b).

To confirm the sequence of the 5'- and 3' termini of the fourth element, 5'- and 3' sequence specific primers were designed based on the known sequence information of both ends (5'AfuNK-dsRNA4-2; GTGTGGAGAAGGTCCAGGAG for 5'-end and 3'AfuNK-dsRNA4-2; GTATCGCTCGCGTCAAGG for 3'-end). The RLM-RACE amplicon generated from the 5'-end specific primers was *ca*. 400 bp in size whereas the 3'-terminal amplicon was *ca*. 500 bp in size (Fig. 27). The amplicons were cloned into pGEM-T Easy vector. Five cDNA clones of each amplicon were sequenced and analysed.

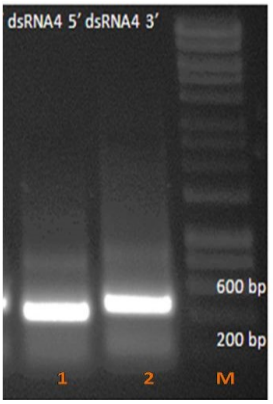


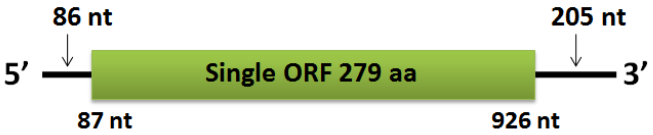
Fig. 27 Agarose gel electrophoresis showing the PCR products of AfuTmV-1 dsRNA4 generated by RLM-RACE. Lane 1, the amplicon produced from RLM-RACE amplification of the 5'-terminus using 5'AfuNK-dsRNA 4-2 as a primer; Lane 2, the amplicon produced from RLM-RACE amplification of the 3'-terminus using 3'AfuNK-dsRNA 4-2 as a primer; Hyperladder 1 (Lane M; 10 kbp; Bioline) was used as marker.

2.10 The genomic sequence of AfuTmV-1 dsRNA4

A schematic representation of the genetic organization is shown Fig. 28a and 28b. The sequence of AfuTmV-1 dsRNA4 comprises 1131 bp with 5'- and 3'- UTRs (of 86 and 205 bp in length, respectively). These UTRs flank a single ORF which encodes a putative protein of 279 amino acid residues of unknown function protein. The calculated molecular weight of the coding protein is 29.07 kDa. The GC content revealed that the genome contains 65% of GC residues. Interestingly, the sequence of 5'-terminus of dsRNA4 showed heterogeneity (U/C) at the very end of the sequence in three independent RT-PCR and cloning experiments. No ribosomal frameshifting signal was predicted from the sequence.

A BLAST search (BLASTX) of the GenBank database showed that the putative protein encoded by AfuTmV-1 dsRNA4 has significant homology to a hypothetical protein of C. cladosporioides virus 1 (Accession number: YP_009052473.1, 35% identity, *E* value = $1e^{-26}$). Interestingly, it can be observed that AfuTmV-1 dsRNA4 encodes a protein that itself is unusual being abundantly rich in proline (P), alanine (A) and serine (S) residues (termed PAS-rich protein (rp)) when compared to a non PAS-rich, structural coat protein found in Aspergillus fumigatus chrysovirus (AfuCV; Jamal *et al*., 2010). These PASrps have been documented as being potentially encoded by a number of dsRNA containing, unclassified insect, plant and fungal viruses (Fig. 29; Table 13). It was also possible to predict the presence of a hexosaminidase domain from amino acid 35-110 in the putative protein encoded by AfuTmV-1 dsRNA4. However BLAST searches failed to return subjects with significant *E* values to any known proteins in global databases and any similarities to hexosaminidase may be spurious.

a) Genome organisation of AfuTmV-1 dsRNA4



b) Complete sequence of AfuTmV-1 dsRNA4 (1131 nt)

(U/C)GAAUAAAAGAUUUGAUCUGAGCUUGCAUGCUUCUUCGUCGACCUACUACCGUAUAAGCGCCGAGUGAGUCCCCGAGAAGCCCA
CCAUGUCCCCUGCACCGAUCCUUUCCGCGGAACAGGCCGCGCACCUUGCGCAACUGACCGUCGAUGACGUCUCGCUCGUCAUCAAGCU
CGCCUCCCCUCGGCUUAAAAGCCAAGGAGAUUCAGCAGUACGCUGCUGAAAUCGCCGAGGGCGGCGACCCAGAGAUCGACCUCGCCCCG
GGGGCCCCCCGCCCUGUCGUCAUUCACGCGUGGUCCUUUGUCAAGGACCGCGGACAGUACCAUUCCACCUAUGGACUGUCGAGCGCCC
AGGCCGGUGAGCUCCUGGACCUUCUCCACACGGAUCCUAAGGACGCCGUCAAGCGGAUCCAGGCGAUCGUUUCCGAGCGCCUGCGAUU
GCGCGGCUCCUCCAGGCCCGUCGUGGUCGAUCUCGAGGGUGCGCCCUCAUCCACCUCGGCCAAAGGUGGUGCCCCCCCCUCCGGGGAC
GGCGGGCUUAGUCUCCUCCGCCAGGAGAUUAAGCGCAACUCUGGAGUCUACGGCUCCUACCAGUUCACCGCCUACCCCACAAGCCGCC
CCGGAGGGCAGGCAUUCGCCGUUGGCCUGGGUGGAGGCCUGUACGCCUUCGGCCAGUCGAAGGCAGUCGCGGUUGCCGCCGCCCGUAU
CGCUCGCGUCAAGGGGCGCAGCGAUGACCUCGUGCGGCCCUACAUCCAUUUCUGGGUGGAAGGGUCGUCACCACGCGCGGCAUCCGAC
AUCCCCGCGAAGGUUUGGGACGGCCGCCUAGGCCCCCACACUGAACCUCCUGCCGAGCCUGCUACCAAGGCCAAGACGGAGGCGCCCA
CCACAUCGGGUGGUAGGGCAUCCACGCCUAAGGGCGCGGAUGCCGCUUAGCGCUGCCCUUAGCACGCCCUUACCGUUCCCUGAACGGA
UGCGCCCGGGAGAUGCCGCAGUACUCUUUCUCCCGGGGAGGGACAUAGUGGGCUGCCUCAAGUGGUGAACUACACCUUGUCUGUGGGA
GUAGCGCCGAGGUUGCAUGUACUGUCUCACGUACCCCCCCCAAAAGGAUGGCUGACCCCUGUCCGGGGGGGGACGGGGG

c) Amino acid sequence encoded from ORF4 (29.07 kDa)

```
MSPAPILSAEQAAHLAQLTVDDVSLVIKLASLGLKAKEIQQYAAEIAEGGDPEIDLAPGAPRPVVIHAWSFVKDRGQYHSTYG
LSSAQAGELLDLLHTDPKDAVKRIQAIVSERLRLRGSSRPVVVDLEGAPSSTPPKGGAPPSGDGGLSLLRQEIKRNSGVYGSY
QFTAYPTSRPGGQAFAVGLGGGLYAFGQSKAVAVAAARIARVKGRSDDLVRPYIHFWVEGSSPRAASDIPAKVWDGRLGPHTE
PPAEPATKAKTEAPTTSGGRASTPKGADAA
```

d) Amino acid sequence similarity of AfuTmV-1 dsRNA4 and *C. cladosporioides* virus 1 hypothetical protein

```
Query  7    LSAEQAAHLAQLTVDDVSLVIKLASLGLKAKEIQQYAAEIAEGGDPEIDLAPGAPRPVVI  66
            L++ +A  L QL  + ++ ++KLAS GLK   I +YA+ +A+  +P    A   P+P+ I
Sbjct  6    LTSREAKLLMQLPDEAIAGILKLASHGLKPAGIIEYASAVAKDQEPAPFPALTGPKPLTI  65

Query  67   HAWSFVKDRGQYHSTYGLSSAQAGELLDLLHTDPKDAVKRIQAIVSERLRLRGSSRPVVV  126
               SF+ D  QY  TY L+      L  L  +D  +A K + A++S   R+  S +PV V
Sbjct  66   QLHSFLADPRQYSDTYSLAMDVVERLRALEMSDRDEARKELAAVISATPRIAASGKPVRV  125

Query  127  DLEGAPSSTPPKGGAPPSGD---GGLSL---LRQEIKRNSGVYGSYQFTAYPTSRPGGQA  180
              +GAP      GGAP S     G  S+   L  EI  NS +YGSY+F A+ T     +
Sbjct  126  VDDGAP------GGAPGSAQPLAGASSVTKSLAAEIASNSTLYGSYKFDAFETQLGVTER  179

Query  181  FAVGLGGGLYAFGQSKAVAVAAARIARVKGRSDDLVRPYIHFWVEGSSPRAASDIPAKVW  240
            FAV LG   Y    SK  A+A AR+  V GRS     PY+ +           DI ++
Sbjct  180  FAVDLGENFYVVTFSKKNAIAIARLVCVHGRSYASAAPYVVYVDSAGRMLRGGDIDSRRG  239

Query  241  DGRLGPHTEP  250
            D R GP   P
Sbjct  240  DCRRGPRDTP  249
```

Fig. 28 Diagrammatic representation of genome organisation of AfuTmV-1 dsRNA4 (a); and its complete sequences (b); and amino acid sequence encoded from the ORF4 (c) and its similarity to the hypothetical protein of C. cladosporioides virus 1 (d). Start and stop codons are shaded in pink and red, respectively. PAS residues are highlighted in grey.

Interestingly, proteins with a similar PA-rich profile are apparently encoded by most but not all of the infectious dsRNA agents described thus far as shown in Table 14 which are apparently unencapsidated in conventional virus particles. The presence of PASrps was first noted in the mycovirus Phlebiopsis gigantea large virus-1 (Kozlakidis *et al*., 2009) but their role in dsRNA replication in any of the examples described in mycoviruses is obscure. For instance it is known that P-rich domains can mediate protein-protein interactions by interacting with membrane components in the host cells and might thus serve as structural proteins or scaffold proteins anchoring replication complexes to the membrane surface (Spear *et al*., 2010). However these domain motifs are present in some but not all cases and as indicated by Cai *et al*. (2013) that they have a high possibility of random occurrence. Another proposed function of the proline-rich scaffold proteins was described in animal viruses that they involve in the formation of the vacuole-like membraneous structures in the virus-infected cells (Chen and Lamb, 2008). Interestingly, dsRNA viruses possessing PArp contain high GC content such as 59.0% GC in SpFV1, 57.2% GC in CiTV1 and 63.4% GC in AfuTmV-1.

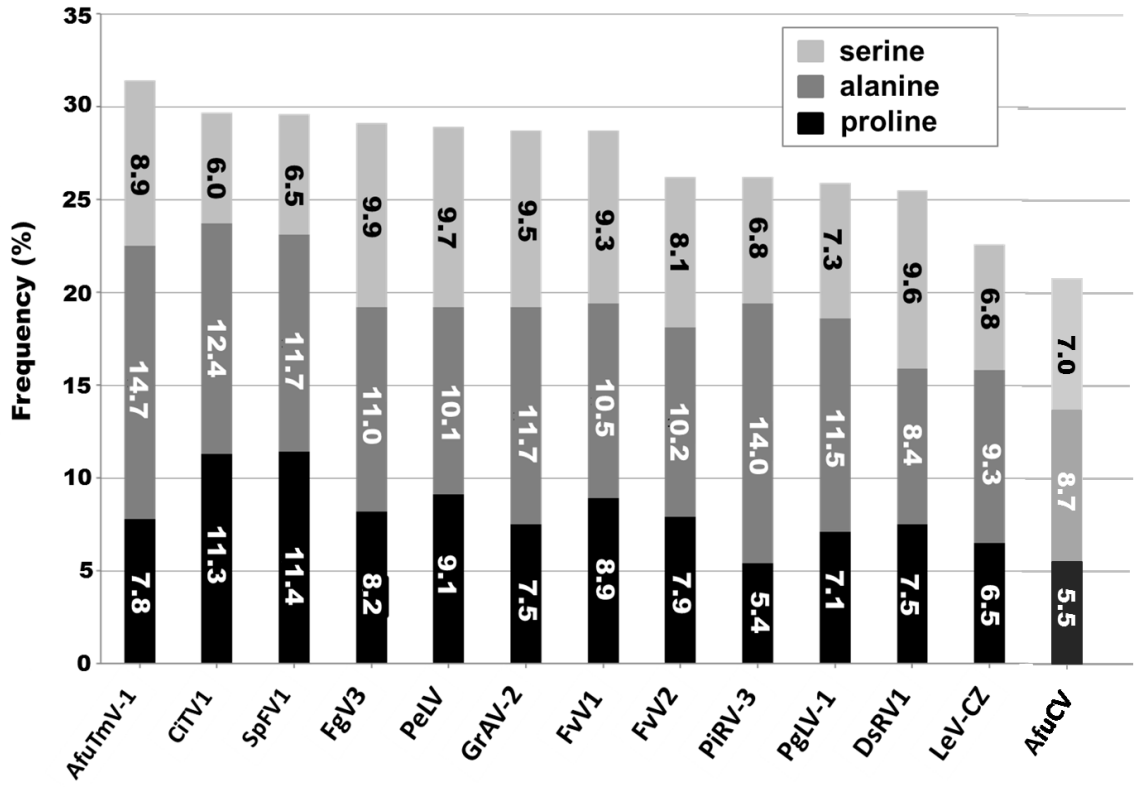


Fig. 29 Proline (P), alanine (A) and serine (S) ratios of PAS-rich proteins (PASrps) of AfuTmV-1 dsRNA4 and related unclassified viruses. Virus names, acronyms and GenBank accession numbers are listed in Table 14. Amino acid (aa) lengths of complete open reading frames and proline-alanine-serine (P-A-S) % ratios of selected, unclassified insect, plant and fungal viruses and a reference, non-PAS-rich coat protein (CP) from Aspergillus fumigatus chrysovirus are summarised in Table 13.

Since ribonucleoprotein (RNA/RdRP) complexes have been reported in a yeast narnavirus (Garcia-Cuella *et al*., 1997) and in a vesicle membrane associated unencapsidated hypovirus found in *C. parasitica* (Hansen *et al*, 1985; Fahima *et al*., 1993; Kazmierczak *et al*., 2012), it is plausible that AfuTmV-1 virus might also occur as an RNA/RdRP nucleoprotein complex in the fungal cytoplasm. Prediction of the subcellular localization of proteins encoded from the AfuTmV-1 dsRNA1-4 using Virus-mPLoc (Shen and Chou, 2010) revealed that all viral proteins were predicted to localize within the host cytoplasms. Furthermore, the virus does not seem to be associated with host organelles or membranous structures since transmembrane helices and signal peptides were not detected in any of the viral dsRNAs using computational analysis. To confirm the location of virus within the host cells, it would be very beneficial to experimentally elucidate more on protein function, protein targeting and localization of the protein encoded by AfuTmV-1 dsRNA4 to understand its replication mode and infectivity.

Table 13. Amino acid (aa) lengths of complete open reading frames and proline-alanine serine (P-A-S) % ratios of selected unclassified insect, plant and fungal viruses and a reference, non-PAS-rich coat protein (CP) from Aspergillus fumigatus chrysovirus.

Virus	Abbreviation	Protein-alanine rich protein (PArp)					
		Proline (P)	Alanine (A)	Serine (S)	P-A	P-A-S	Length (aa)
Aspergillus fumigatus tetramycovirus-1	AfuTmV-1	24 (8.6%)	40 (14.3%)	25 (8.9%)	64 (22.9%)	89 (31.9%)	279
Circulifer tenellus virus 1	CiTV1	128 (11.3%)	141 (12.4%)	68 (6.0%)	269 (23.7%)	337 (29.7%)	1133
Spissistilus festinus virus 1	SpFV1	125 (11.4%)	129 (11.7%)	72 (6.5%)	254 (23.1%)	326 (29.6%)	1101
Fusarium graminearum virus 3	FgV3	115 (8.2%)	151 (11.0%)	135 (9.9%)	266 (19.4%)	401 (29.3%)	1369
Persimmon latent virus	PeLV	98 (9.1%)	109 (10.1%)	104 (9.7%)	207(19.3%)	311 (28.9%)	1075
Grapevine associated totivirus-2	GrAV-2	99 (7.5%)	154 (11.7%)	125 (9.5%)	253 (19.3%)	378 (28.8%)	1313
Fusarium virguliforme dsRNA mycovirus 1	FvV1	117(8.9%)	137 (10.5%)	122 (9.3%)	254 (19.4%)	376 (28.7%)	1311
Fusarium virguliforme dsRNA mycovirus 2	FvV2	107 (7.9%)	138 (10.2%)	109 (8.1%)	245 (18.2%)	354 (26.3%)	1347
Phytophthora infestans RNA virus-3	PiRV-3	70 (5.4%)	182 (14%)	88 (6.8%)	252 (19.4%)	340 (26.2%)	1296
Phlebiopsis gigantea large virus-1	PgLV-1	142 (7.1%)	232 (11.5%)	147 (7.3%)	374(18.6%)	521 (25.9%)	2012
Diplodia scrobiculata RNA virus 1	DsRV1	31 (7.5%)	35 (8.4%)	40 (9.6%)	66 (15.9%)	106 (25.5%)	416
Lentinula edodes virus-CZ	LeV-CZ	129 (6.5%)	184 (9.3%)	135 (6.8%)	313 (15.8%)	448 (22.7%)	1975
Aspergillus fumigatus chrysovirus	AfuCV	52 (5.5%)	83 (8.7%)	67 (7.0%)	135 (14.2%)	202 (21.2%)	953

Table 14. Viruses selected for the amino acid analysis of PASrp.

Abbrev.	Virus	Family	Segment no.	Group	Accession no.	Features
AfuTmV-1	Aspergillus fumigatus tetramycovirus 1	unclassified virus	4	dsRNA virus	HG975302	Fungal virus/unencap.
PeLV	Persimmon latent virus	unclassified virus	-	dsRNA virus	BAN29037.1	Plant virus
CiTV1	Circulifer tenellus virus 1	unclassified virus	1 (8 kb)	dsRNA virus	YP_003800002.1	Insect virus/unencap.
SpFV1	Spissistilus festinus virus 1	unclassified virus	1 (8 Kb)	dsRNA virus	YP_003800000.1	Insect virus/unencap.
PgLV-1	Phlebiopsis gigantea large virus-1	Totiviridae	1 (>9 kb)	dsRNA virus	CAJ34331.3	Fungal virus/unencap.
FgV3	Fusarium graminearum virus 3	unclassified virus	1 (>9 kb)	dsRNA virus	YP_003288788	Fungal virus/unencap.
DsRV1	Diplodia scrobiculata RNA virus 1	unclassified virus	1 (>9 kb)	dsRNA virus	ACD91657	Fungal virus/unencap.
LeV-CZ	Lentinula edodes virus-CZ	unclassified virus	1 (>11 kb)	dsRNA virus	AEB96149	Fungal virus/unencap.
FvV1	Fusarium virguliforme dsRNA mycovirus 1	unclassified virus	-	dsRNA virus	AEZ54147.1	Fungal virus
FvV2	Fusarium virguliforme dsRNA mycovirus 2	unclassified virus	-	dsRNA virus	AEZ54148	Fungal virus
GrAV-2	Grapevine associated totivirus-2	Totiviridae	1 (>9 kb)	dsRNA virus	ADO60932.1	Fungal virus/unencap.
PiRV-3	Phytophthora infestans RNA virus-3	unclassified virus	-	dsRNA virus	AEX87901.1	Oomycete fungal virus
AfuCV	Aspergillus fumigatus chrysovirus	Chrysoviridae	4	dsRNA virus	CAX48751	Fungal virus

2.11 Multiple nucleotide sequence alignment of AfuTmV-1 dsRNAs

The sequences of 5'- and 3'-UTRs of all four AfuTmV-1 dsRNAs were aligned using the MAFFT multiple alignment programme. The alignment showed that the sequences shared extensive similarities at both 5'- and 3'- UTRs (Fig. 30). The 5' UTRs of AfuTmV-1 dsRNA1-4 were similar in length and sequence and ranged in size from 35-86 nucleotides (nt), whereas the extreme 5' terminal nucleotides showed sequence heterogeneity (U/C) while the next nine 5' nt of all four dsRNAs were strictly conserved (5'-YGAAUAAAAG) (Fig. 30a). The presence of sequence heterogeneity and variability at the very 5'-terminus of mycovirus dsRNAs has also been reported in *Penicillium chrysogenum* mycovirus (PcV) (Edmondson *et al*., 1984), Rosellinia necatrix megabirnavirus *1* (RnMBV1) (Chiba *et al*., 2009) and Rosellinia necatrix quadrivirus (RnQV1) (Lin *et al*., 2012).

A) 5'-terminal sequences (dsRNA1-4)

```
dsRNA1   5'-CGAAUAAAAGGAAACUUCUCCGCUUGUGUGCUCCA-----------------------------------
dsRNA3   5'-CGAAUAAAAGGAAACUAGAACGCUUGCAUGCUCCAACGCCCCGCAACCACC-------------------
dsRNA2   5'-UGAAUAAAAGAUUUGUUCUGGGCAAGCACGCUGCAACGUCGAGCAACUCCCAUAUAGGGGACGA-UGAGAC-
dsRNA4   5'-YGAAUAAAAGAUUUGAUCUGAGCUUGCAUGCUUCUUCGUCGACCUACUACCGUAUAAGCGCCGAGUGAGUCC
             *********.          **  * . *** *

dsRNA1      -------------- (35 nt)
dsRNA3      -------------- (51 nt)
dsRNA2      -------------- (70 nt)
dsRNA4      CCGAGAAGCCCACC (86 nt)
```

B) 3'-terminal sequences (dsRNA1-4)

```
dsRNA4     GCGUGCCCUUAGCACGCCCUUACCGUUCCCUGAACGGAUGCGCCCGGGAGAUGCCGCAGUACUCUUUCUCCCGG
dsRNA3     ------------------------------------------------------------------------------
dsRNA1     ------------------------------------------------------------------------------
dsRNA2     ------------------------------------------------------------------------------

dsRNA4     GGAGGGACAUAGUGGGCUGCCUCAAGUGGUGAACUACACCUU-GUCUGUGGGAGUAGCGCCGAGGUUGCAUGUA
dsRNA3     ----------------GUGGAAC--GU--UGGAGAUGCUCGCGUCCCGACG--------C-GGCGGGGGGGGCC
dsRNA1     ------------------UCACC--GCAC---ACAAGC---CGGCCCGGCG--------G-GAGGGGGCGCGCG
dsRNA2     ----------ACUGCGCUGCAUC--GUGCUUGACCAGC-----GUCCGACG--------G-GGAGGGG-----C
                             *  *      *           * *  *          *. *  *

dsRNA4     CUGUCU---CACGUACCCCCCCCAAAAGGAUGGCUGACCCCUGUCCGGGGGGGGGACGGGGG- -3'(205 nt)
dsRNA3     CUG-------GCCCCCCCCCCGCG--------------------CCGGG-----GCGGGGG- -3' (74 nt)
dsRNA1     UUGUCGAUCCGCGUGCCCCCCCCG--------------------CCGGG-----GCGGGGGG -3' (76 nt)
dsRNA2     UU--------AAUUGCCCCCCCCG--------------------CCGGG-----CCUGGGGG -3' (72 nt)
            *        .    ****** *.                    *****      * ****
```

C) 3'-terminal sequences (dsRNA1-3)

```
dsRNA3     ------GUGGAACGU--UGGAGAUGCUCGCGUCCGACGCGGCGGGGGGGGGCCCUG-------GCCCCCCCCC
dsRNA1     --------UCACCGCAC---ACAAGC---CGGCCCGGCGGGAGGGGGCGCGCGUUGUCGAUCCGCGUGCCCCC
dsRNA2     ACUGCGCUGCAUCGUGCUUGACCAGC-----GUCCGACGGGGAGGGG-----CUU--------AAUUGCCCCC
                     * **      *   **       ***.** *. ****       *        .    *****

dsRNA3     CGCGCCGGGGCGGGGG- -3' (74 nt)
dsRNA1     CCCGCCGGGGCGGGGGG -3' (76 nt)
dsRNA2     CCCGCCGGGCCUGGGGG -3' (72 nt)
           * ******* * ****
```

Fig. 30 Comparative nucleotide sequence alignment of the 5-UTRs (a) and 3-UTRs (b and c) of AfuTmV-1

dsRNAs using MAFFT program. CAA repeats on the 5'-terminal region are underlined. Asterisks signify identical nucleotides at the indicated positions and colons specify that three out of four nucleotides are identical at the indicated positions. Y represents C or U.

Also "CAA" repeats were found within the 5'-UTR of dsRNA2 and 3 but at a low frequency (as underlined in Fig. 30a). It has been suggested that CAA repeat sequences upstream of the initiation codon possibly serve as a translational enhancer as similar to observations in some mycoviruses in the families *Partitiviridae* and *Chrysoviridae* (Jiang and Ghabrial, 2004; Tavantzis, 2008). Conversely the absence of CAA repeat sequences in dsRNA1 and 4 could possibly contribute to a low level of translational activity of the components.

The 3' terminal sequences of AfuTmV-1 dsRNA1-3 were also highly conserved and comprised 18/22 identical nucleotides including 4 G residues at the extreme 3' terminus of each dsRNA (-CCCCCCSCGCCGGGSCKGGGG-3') (Fig. 30c). Unlike other dsRNA viruses the coding strands of AfuTmV-1 dsRNAs lack poly(A) tails at the 3'-termini similar to the dsRNA genomes of mycoviruses in the families *Totiviridae* or *Chrysoviridae* and in the yeast narnaviruses which are also not poly-adenylated (Solorzano *et al*., 2000). The lengths of 3' UTRs of the AfuTmV-1 dsRNA1-3 were similar and ranged in size from 74-76 nucleotides. However, the 205 nt at the 3' UTR of the AfuTmV-1 dsRNA4 was considerably longer with a poorly conserved sequence as compared to the other three dsRNAs apart from the very 3' terminal 4 G residues (Fig. 30b).

Conserved terminal sequences of viral RNA has been reported to be a feature of viruses with multipartite genomes and it is crucial for viral RNA replication, transcription and packaging in both ssRNA and dsRNA viruses (Wei *et al*., 2004). In addition, the conserved feature of the UTRs also supports the notion that the four AfuTmV-1 dsRNAs represent interdependent segments of a single viral genome. Beside the strictly conserved sequences of the 3'-termini, the sequences also showed stretches with high GC levels (Fig. 30b and 30c).

The 5'- and 3'-UTRs of AfuTmV-1 dsRNAs contain nucleotide inverted sequences and were predicted to be highly structured (Fig. 31). Multiple stem-loop structures were predicted at the 5'-termini of dsRNA2-4 and 3'-terminus of dsRNA4 whereas potential long stem-loops resembling panhandle structures were predicted at the 5'-terminus of dsRNA1 and 3'- termini of dsRNA1-3 analysed as evidenced using the Mfold programme. Hillman and Cai (2013) suggested that the potential of RNA to form a secondary panhandle structure at the 5' and 3' ends may be a diagnostic feature for ssRNA mitoviruses (mitochondrial RNA viruses), ssRNA yeast narnaviruses and many RNA viruses and may protect RNA from exonuclease digestion. The presence of the stem-loop at 5'-UTR of dsRNA1 could serve as an RdRP recognition

structure whereas the stem-loop structure at the 3'-terminus could act as an inhibitor or terminator in the translation or replication process driven by the RNA polymerase. As the dsRNAs of the AfuTmV-1 virus lack the poly(A) stretches the stem loop structure at the 3'-terminus could also serve as a self-priming element, allowing viral RNA replication (Fujimura and Esteban, 2004; Yu and Markoff, 2005). Thus, it is possible that they may also function in a similar fashion for AfuTmV-1.

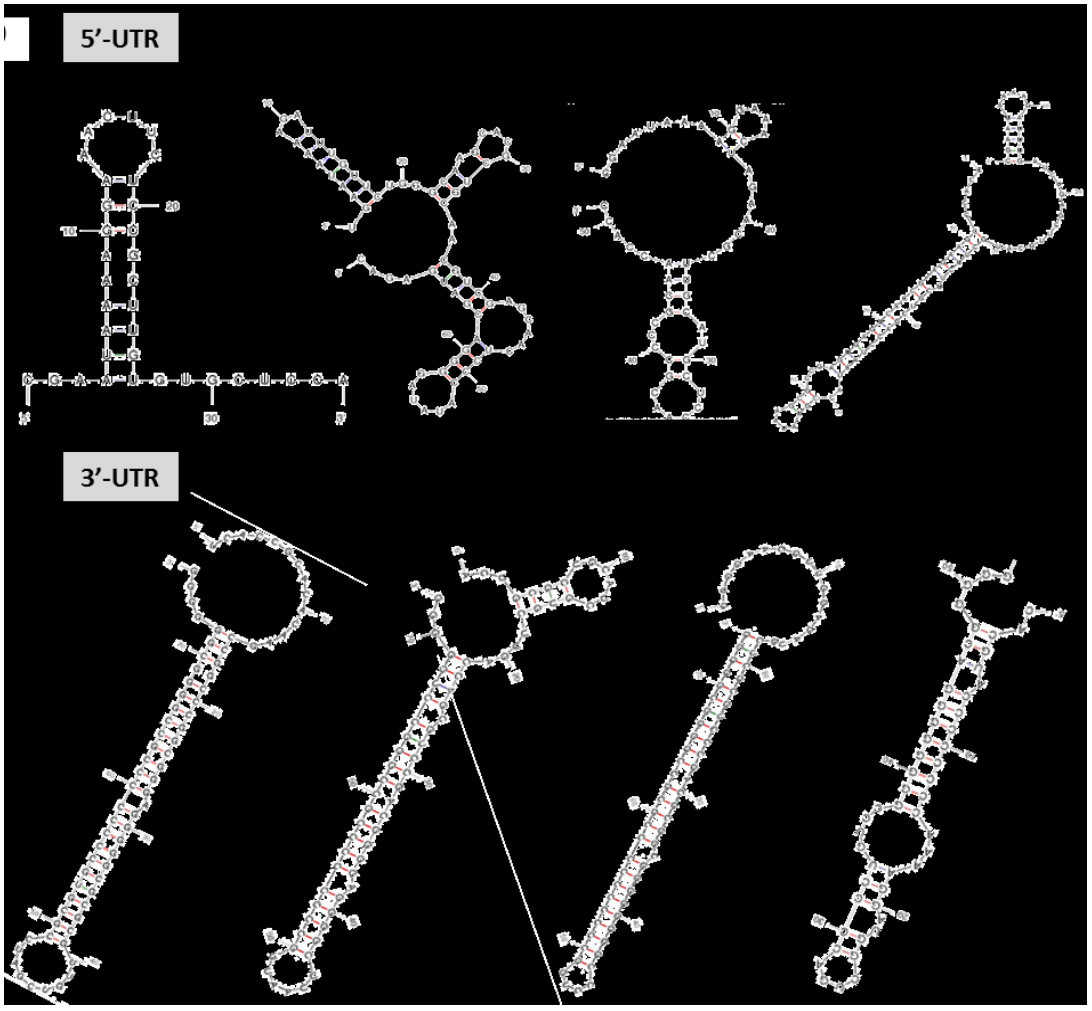


Fig. 31 Predicted secondary structure of the 5'- UTRs and 3'- UTRs of AfuTmV-1 dsRNAs using Mfold programme (http://mfold.rna.albany.edu/?q=mfold) (Zuker, 2003).

2.12 A DNA copy of AfuTmV-1 does not reside in the Af293 genome

An NCBI BLASTN search against the *A. fumigatus* Af293 reference genome sequences retrieved from the NCBI and other sources available online such as The Central *Aspergillus* Resource (CADRE; www.cadre-genomes.org.uk, Mabey *et al*., 2004) and the *Aspergillus* Genome Database (*Asp*GD; http://www.aspgd.org, Arnaud *et al*., 2010) with all four dsRNA sequences revealed no significant similarity. Southern blot analysis was also conducted to confirm the

homology between the DIG-lebelling PCR probes of four AfuTmV-1 sequences and *Hind*III- or *Eco*RI-digested genomic DNA of the virus-free *A. fumigatus* NK125. The result showed undetectable signal from the digested DNA samples hybridized with individual virus probes (Fig. 32), indicating no similarity between the virus and fungus genome. The *FKS* gene which is present as a single copy in the fungal genome was used as a control. Thus, it can be concluded that *A. fumigatus* genome does not have a DNA copy of any of the AfuTmV-1 mycovirus dsRNAs described.

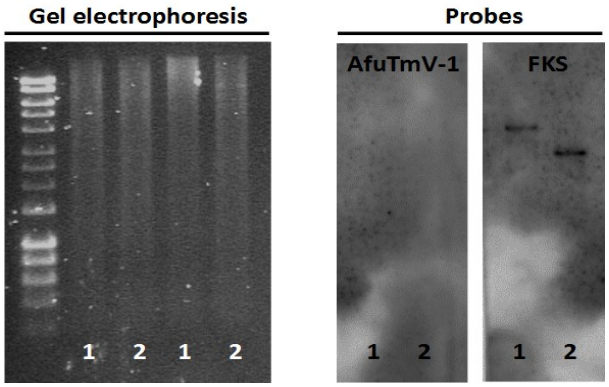


Fig. 32 Analysis of DNA copy of AfuTmV-1 in the NK125 genome by southern blot hybridization. The genomic DNA was digested with either *Hind*III (lane#1) or *Eco*RI (lane#2).

2.13 Phylogenetic analysis of the putative RdRP of AfuTmV-1

To investigate phylogenetic relationships between AfuTmV-1 and other RNA mycoviruses an alignment of 40 seed members of the RdRP domain in Pfam database 22.0 and NCBI was performed. These 40 seed members belong to the members of the families such as *Reoviridae, Totiviridae*, *Chrysoviridae, Alternaviridae, Quadriviridae* and *Partitiviridae*, as well as ssRNA viruses in the family *Hypoviridae, Picornaviridae* and *Narnaviridae*. Previously reported *Aspergillus* mycoviruses and unencapsidated mycoviruses were also included in the alignment which some of them yet to be characterized. The predicted RdRP region of AfuTmV-1 was aligned to the seed alignment using CLUSTAL Omega whereas phylogenetic analyses were carried out using Fast Fourier Transform MAFFT programme L9INS-1 and CLUSTAL Omega.

Examination of the deduced amino acid (aa) sequence of the AfuTmV-1 dsRNA1 ORF revealed the presence of three conserved motifs characteristic of RdRPs of dsRNA viruses of simple eukaryotes. The resulting multiple alignment of the putative RdRP of AfuTmV-1 and representative RNA viruses indicated that AfuTmV-1 RdRP shares homology with the viral RdRP families of the three main conserved motifs, A-C (Fig. 33), which corresponded to motifs IV-VI,

respectively. Theoretically, the RdRP motifs were shown to comprise eight motifs (Koonin and Dolja, 1993) but only motifs IV-VI were distinctly conserved throughout the entire RdRP class.

Interestingly, the highly conserved GDD or ADD motif, which is normally invariant for positive-stranded RNA and dsRNA viruses, was replaced by the GDN triplet in the AfuTmv-1 virus. The GDN triplet in RdRP genes followed immediately by a Q residue is a characteristic sequence motif normally found in the L genes of non-segmented negative-stranded (-) ssRNA viruses such as rabies virus (rhabdoviruses) and measles virus (paramyxoviruses) within the order of *Mononegavirales* both of which are animal viruses (Delarue *et al*., 1990; Schnell and Conzelmann, 1995). Recently, a (-) ssRNA virus in fungi has first been found in the ascomycete fungus *Erysiphe pisi* (pea powdery mildew fungus). This newly identified (-) ssRNA mycovirus was named Erysiphe pisi mononegavirus which possesses L protein-like sequences 1-4 and a GDNQ motif (Kondo *et al*., 2013). This includes a recently described negative-stranded mycovirus of *Sclerotinia sclerotiorum* (Lui *et al*., 2014). The presence of a GDN motif was also found in a positive-stranded RNA virus such as CHV1 isolated from *Cryphonectria parasitica* and DaRV isolated from the fruit tree fungal pathogen, *Diaporthe ambigua* (Preisig *et al*., 2000) including the unclassified Cladosporium cladosporioides virus 1 (Accession number: YP_009052470.1).

Experimental evidence revealed that the GDN motif of negative-strand, non-segmented RNA viruses is a variant of the GDD motif of positive-strand RNA viruses, compatible with the notion that the second D is not essential, but its discovery in the sequence of a dsRNA virus is unique. It has been described that conservation of the two aspartic acids (DD) in positive-strand RNA viruses are crucial for the RdRP function since the residues are involved in catalytic activity and metal coordination (Mg^{2+}) of the enzyme at the catalytic site (Delarue *et al*., 1990). However, Jablonski and Morrow (1995) reported that change of the second aspartic acid (GDD) to asparagine (GDN) by site-direct mutagenesis in Poliovirus RdRP resulted in a different metal utilization (preferable to Mn^{2+} or Fe^{2+} instead of Mg^{2+}) for enzyme activity. Thus, differencies in the conserved motifs may alter enzyme activity. It would be interesting to investigate the role and activity of the intrinsic GDN catalytic site in AfuTmV-1 whether it behaves in the same manner as observed in the mutated poliovirus RdRP.

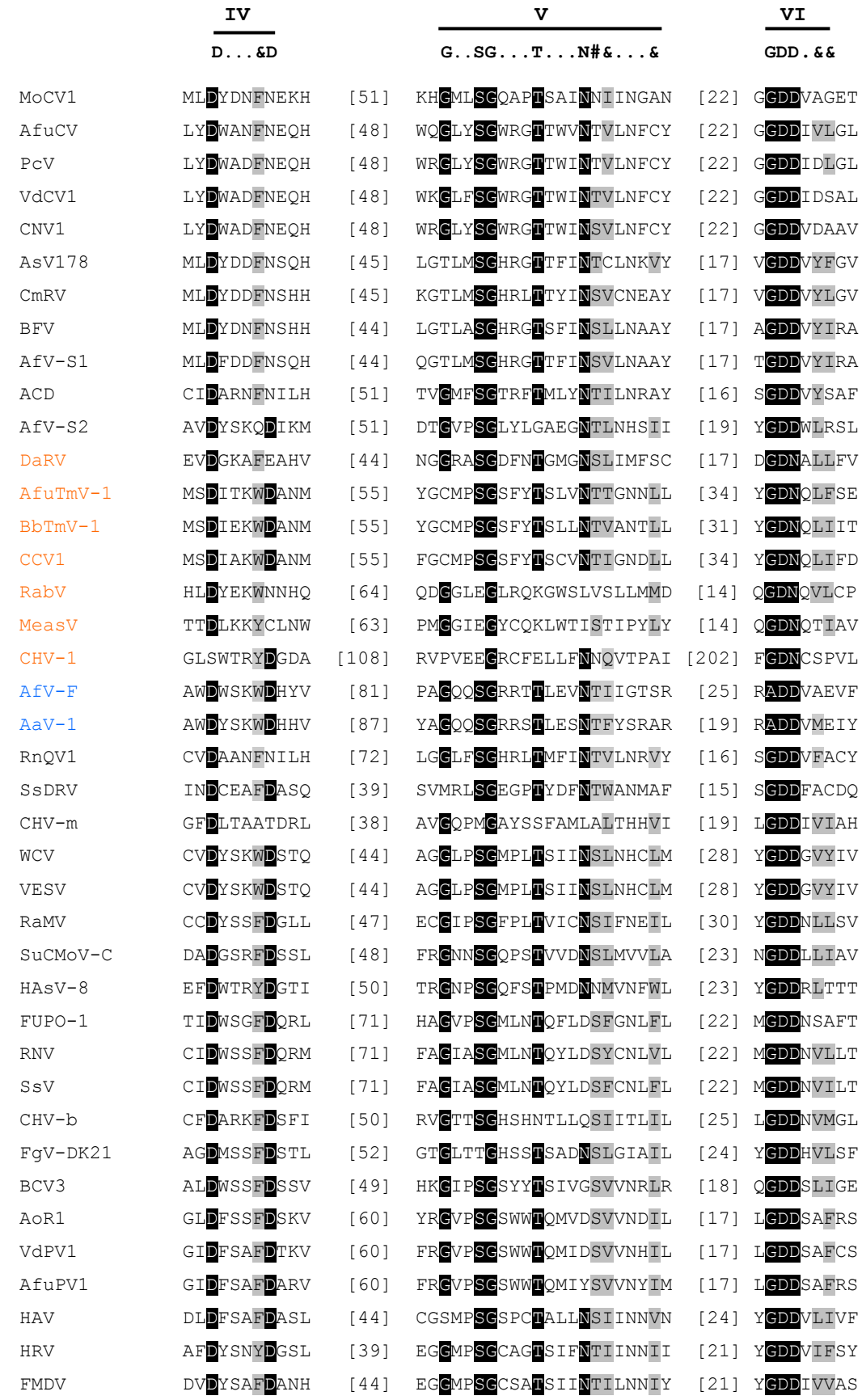


Fig. 33 Amino acid sequence alignment of RdRP regions of AfuTmV-1 dsRNA1 and selected RNA viruses. The alignment was generated using CLUSTAL Omega. Three conserved RdRP motif corresponding to respectively motifs IV, V and VI according to Bruenn (1993) are shown, and consensus sequences are highlighted either with black or grey shading. Black shading indicates identical amino acid residues and

grey shading indicates those which belong to the same group of consensus sequences. The "#" symbol refers to polar uncharged residues S or T and "&" denotes any bulky hydrophobic residue (L, I, V, M, F, Y, W). Numbers within the brackets indicate the number of amino acids that are not shown. Virus names, acronyms and GenBank accession numbers are listed in Table 15.

Although the PSI-BLAST search of amino acid sequence encoded from dsRNA1 revealed the best match to, respectively a putative RdRP of Cladosporium cladosporioides virus1 and a putative RdRP of walrus calicivirus, phylogenetic analysis based on core RdRP motifs of the AfuTmV-1 against selected mycoviruses yielded dissimilar results. A bootstrap test was performed with 1000 replicates for the neighbor-joining (NJ) tree. The analysis using Clustal Omega (Fig. 34) showed that the AfuTmV-1 clustered in the same group to two recently discovered Cladosporium cladosporioides virus1 (CCV1) and Beauveria bassiana tetramycovirus-1 (BbTmV-1) which are similar in profile and size but not sequence to their AfuTmV-1 counterpart. It indicated that these dsRNA viruses might have evolved from the same origin and shared a common ancestor, however, they are found to be distantly related to any classified dsRNA families. In addition, they showed a close phylogenetic relationship to the unclassified Diaporthe ambigua RNA virus 1 (DaRV) which is a (+) ssRNA virus. It can be noted that mycoviruses clustered in this clade contain the GDN motif in their RdRP molecules and are yet characterized to a known virus family. Interestingly, a new clade consisting of AfuTmV-1, CCV1, BbTmV-1 and DaRV appeared to be a sister clade of (-) ssRNA viruses, measles virus and rabies virus, which also possess the GDN motif but it is notably distant from viruses in the families *Caliciviridae* and *Picornaviridae*. However, the resemblance of the motif in *Mononegavirales* as such is spurious. Thus, AfuTmV-1 appears to be evolutionarily related to but not a member of (-) ssRNA virus families when the analysis was carried out using Clustal Omega. However, the presence of the Aspergillus foetidus slow virus 2 (AfV-S2), which is a dsRNA virus, in this cluster was unexpected.

Construction of the phylogenetic tree using the Fast Fourier Transform MAFFT programme L9INS-1 revealed similar results to that of Clustal Omega programme which showed that AfuTmV-1 clustered in the same group (monophyletic group) to CCV1 and BbTmV-1 and was phylogenetically divergent from other viruses selected in the analysis (Fig. 35 and 36). However, the resulting NJ tree placed AfuTmV-1 and CCV1 as sister taxa, suggesting a close relationship between these two newly discovered viruses. Interestingly, AfuTmV-1 RdRP appeared to be a derivative of the picorna-like superfamily with technically, the highest similarity to human hepatitis A virus including viruses in the family *Picornaviridae* (human rhinovirus A and foot and mouth disease virus) and less similarity to the walrus calicivirus in the family *Caliciviridae*. Excluding its counterparts, the closest grouping (although still remote)

of the AfuTmV-1 were the picornaviruses but it was clearly different both in genome organization and particle morphology to the picornaviruses. Picornaviruses have monopartite (+) ssRNA genomes with an icosahedral capsid whereas the AfuTmV-1 is an unencapsidated virus with a tetrapartite dsRNA genome. Only one gene, the RdRP, is shared between the picornaviruses and AfuTmV-1. It has been proposed that the RdRP is shared by all (+) ssRNA and dsRNA viruses and also found to be the only gene that is present in all capsid-less RNA viruses (Koonin and Dolja, 1993). Another feature which discriminates AfuTmV-1 from other (+) ssRNA viruses is that the coding strand of the four AfuTmV-1 dsRNAs does not have the poly (A) tail which does not resemble to the replicative form of ssRNA viruses. The small genomic size (*ca*. 1.2-2.4 kbp) of AfuTmV-1 excludes the virus from other unencapsidated ssRNA viruses in the families *Hypoviridae* and *Endonarviridae* which contain *ca*. 9-13 kbp and *ca*. 14-17 kbp genomes, respectively. It is thus possible that the AfuTmV-1 is a dsRNA virus and phylogenetically a relative of the (+) ssRNA viruses in the family *Picornaviridae* and independently evolved from the (+) ssRNA virus ancestors. The virus subsequently lost its capsid and genes involved in viral transmissibility and infectivity *via* long reductive evolution of the virus.

From the alignment, although the viral RdRP genes are highly conserved among RNA viruses, phylogenetic analysis of the sequence indicates a polyphyletic origin of dsRNA viruses which contributes to grouping of the dsRNA viruses in different supergroups of the positive(+) ssRNA viruses. Mycovirus RdRPs are known to be related to the RdRPs of viruses in the picorna-like superfamily of eukaryotic, positive-strand RNA viruses (Koonin *et al*., 2008) and numerous examples can be found in the literature including Ustilaginoidea virens non segmented virus 1 (Zhang *et al*., 2014), a number of viruses in the *Partitiviridae* family (Nibert, 2014) and Aspergillus foetidus virus-S2, which is closely related to several members of the *Caliciviridae* family and *Comoviridae* subfamily (Kozlakidis *et al*., 2013). Thus, the phylogenetic analysis results described here support the contention that AfuTmV-1 should also be classified as a picorna-like virus. The evidence to support this conclusion has been documented which stated that the origins of dsRNA viruses are contributed from diverge lineages of (+) ssRNA viruses (Ahlquist, 2006).

Fig. 34 Phylogenetic analysis of AfuTmV-1 and other mycoviruses on the basis of core RdRP sequences. A phylogenetic tree was constructed using Clustal Omega. A bootstrap test was conducted with 1000 re-samplings for the neighbor-joining trees. Virus genera and families are indicated by normal and embolded type respectively. Numbers on the nodes indicate percentage of bootstrap support from 1000 replicates with branch length values indicated. GenBank accession numbers and abbreviations for each sequence are shown in Table 15.

Fig. 35 Phylogenetic analysis of AfuTmV-1 and other mycoviruses (excluding viruses from the family *Alternaviridae*) on the basis of core RdRP sequences. A phylogenetic tree was constructed using the Fast Fourier Transform MAFFT program L9INS-1 (Katoh *et al*., 2005). A bootstrap test was conducted with 1000 re-samplings for the neighbor-joining trees. Virus genera and families are indicated by normal and embolded type respectively. Numbers on the nodes indicate percentage of bootstrap support from 1000 replicates with branch length values indicated. GenBank accession numbers and abbreviations for each sequence are shown in Table 15.

Fig. 36 Phylogenetic analysis of AfuTmV-1 and other mycoviruses (including viruses from the family *Alternaviridae*) on the basis of core RdRP sequences. A phylogenetic tree was constructed using the Fast Fourier Transform MAFFT program L9INS-1 (Katoh *et al*., 2005). A bootstrap test was conducted with 1000 re-samplings for the neighbor-joining trees. Virus genera and families are indicated by normal and embolded type respectively. Numbers on the nodes indicate percentage of bootstrap support from 1000 replicates with branch length values indicated. GenBank accession numbers and abbreviations for each sequence are shown in Table 15.

Fig. 34 Phylogenetic analysis of RdRP using Clustal Omega

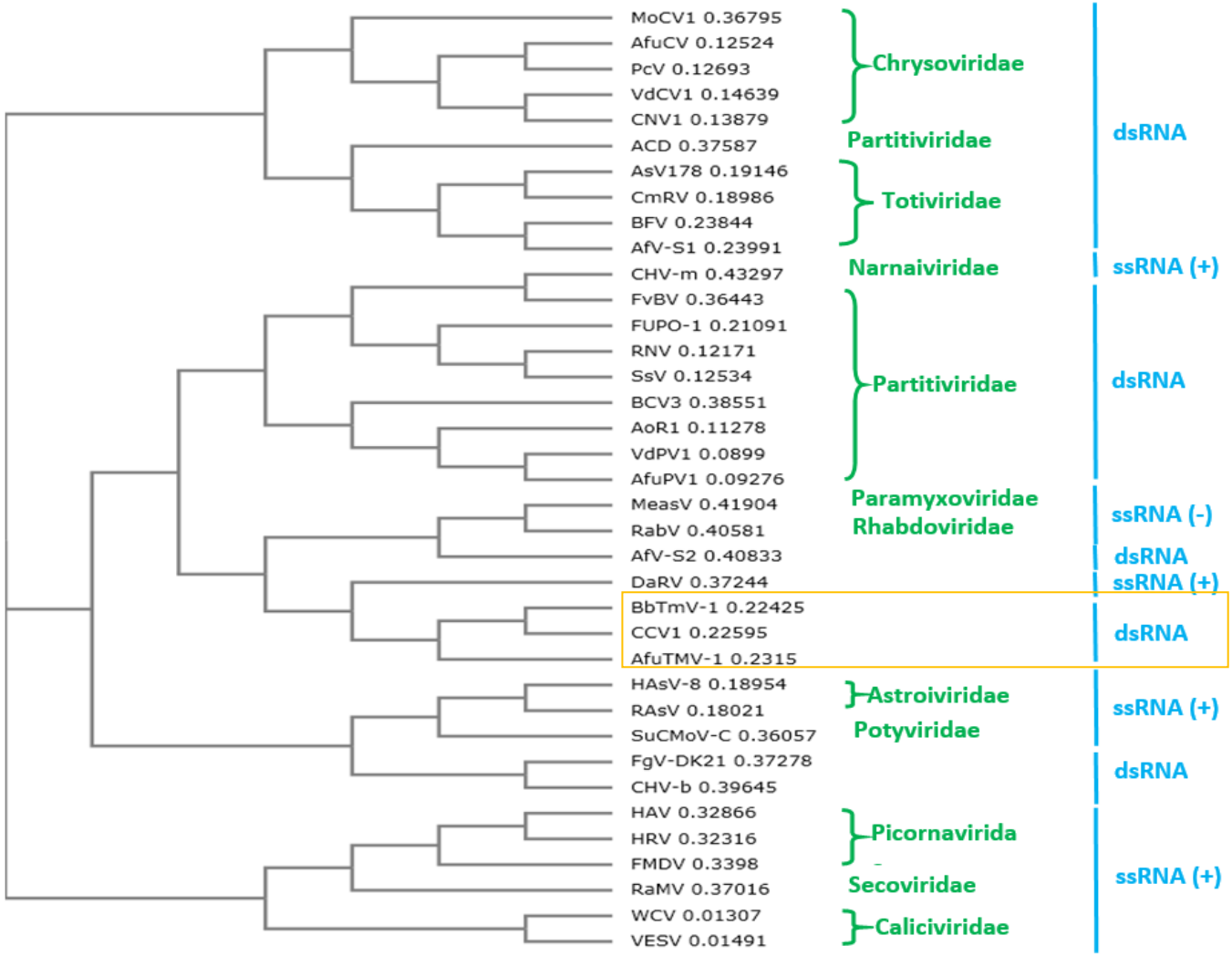


Fig. 35 Phylogenetic analysis of RdRP using MAFFT program

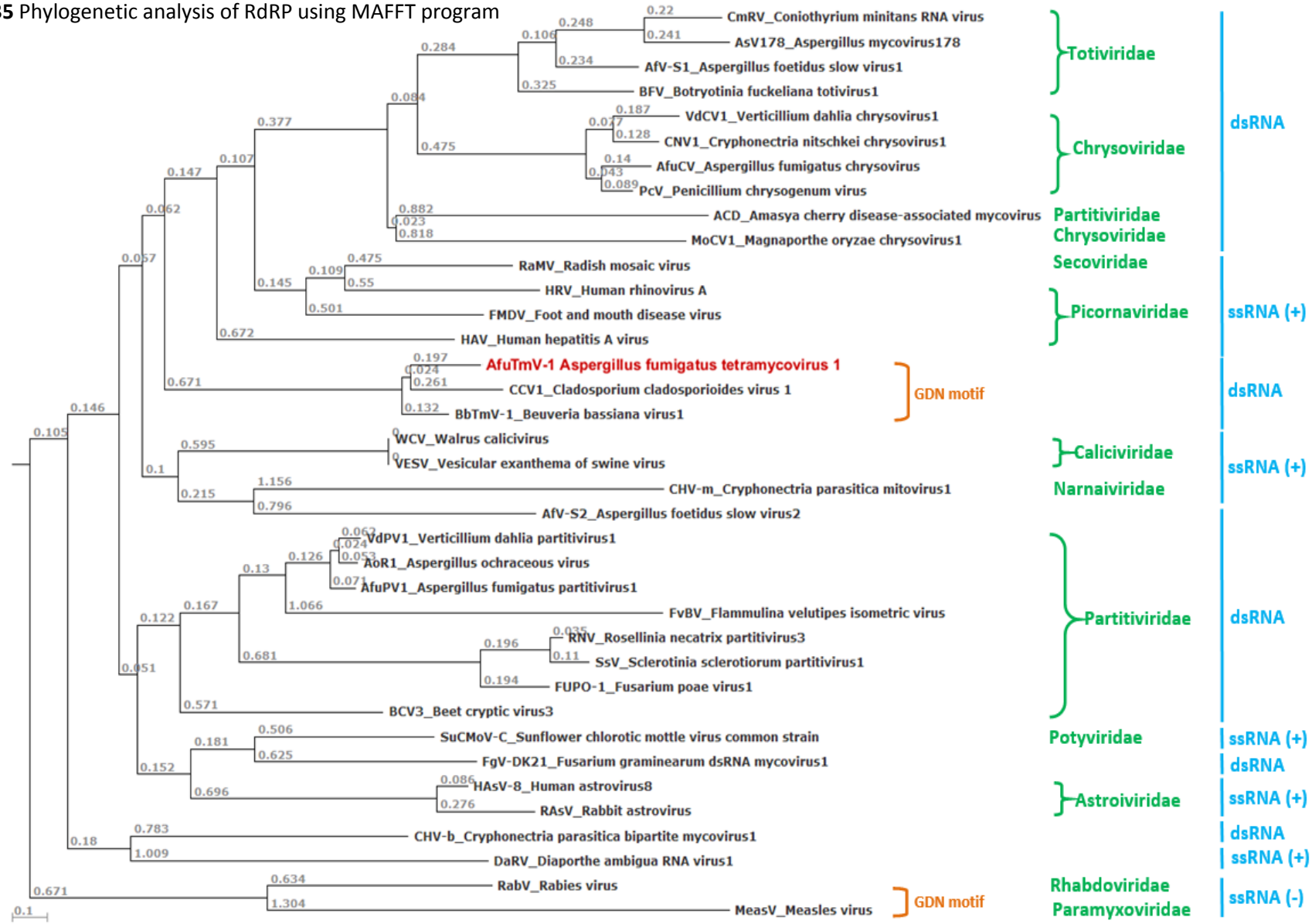


Fig. 36 Phylogenetic analysis of RdRP using MAFFT program

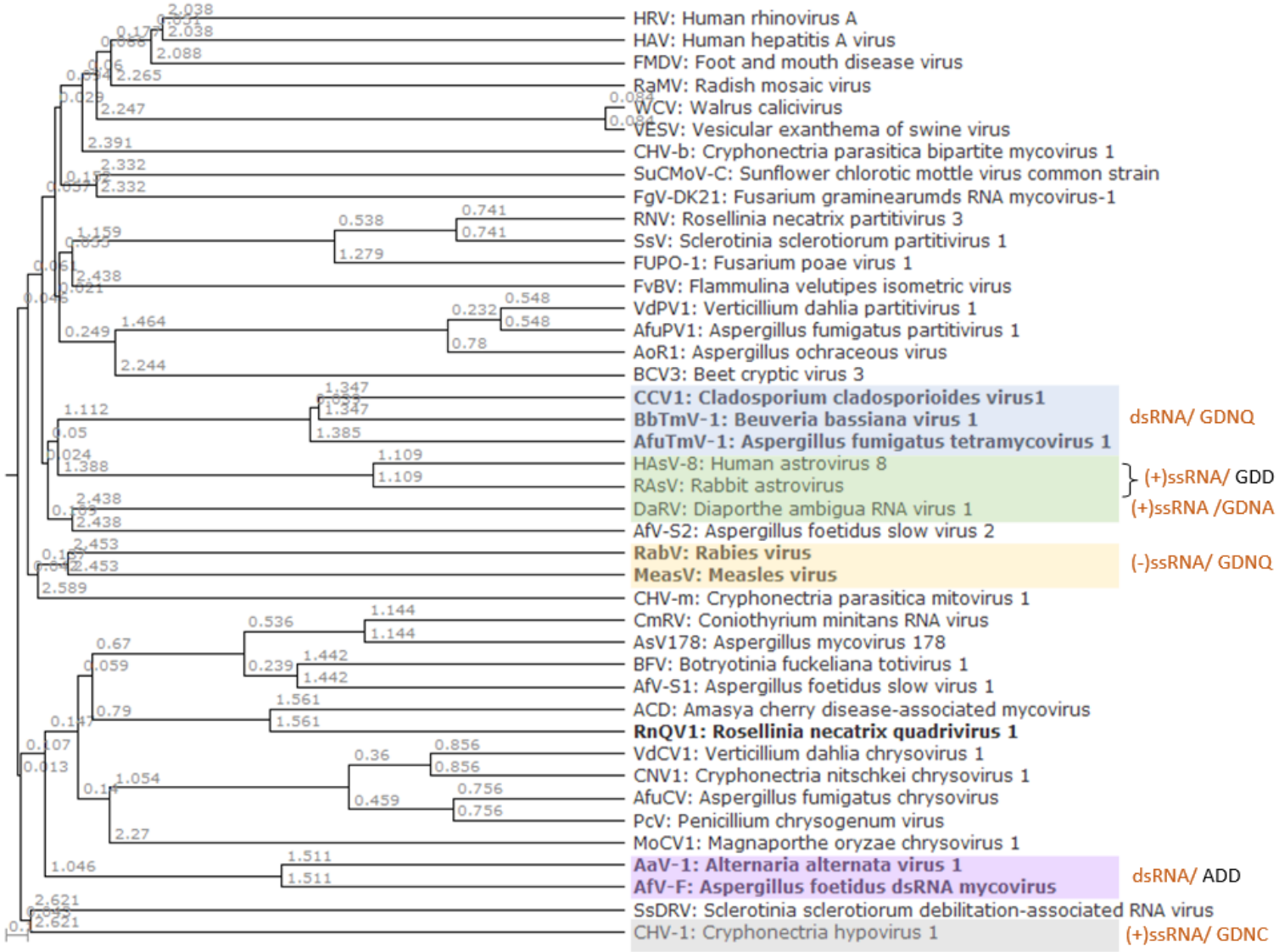


Table 15. Viruses selected for phylogenetic analysis of RdRP.

Abbrev.	Virus	Family	Genus	Virus Group	Accession no.	Features
AfuTmV-1	Aspergillus fumigatus tetramycovirus 1	unclassified	-	dsRNA	HG975302	4 segments (1.2-2.4 kbp) Unencapsidated
CCV1	Cladosporium cladosporioides virus 1	unclassified	-	-	YP_009052470	4 segments (1.3-2.4 kbp) + 1 defective RNA (0.9 kb)
BCV3	Beet cryptic virus 3	Partitiviridae	Alphacryptovirus	dsRNA	S63913.1	
SuCMoV-C	Sunflower chlorotic mottle virus common strain	Potyviridae	Potyvirus	ssRNA (+)strand	JN863233	
RaMV	Radish mosaic virus	Secoviridae	Comovirus	ssRNA (+)strand	NC_010709	
FgV-DK21	Fusarium graminearumds RNA mycovirus-1	unclassified	hypovirus	dsRNA	YP_223920.2	Unencapsidated
WCV	Walrus calicivirus	Caliciviridae	-	ssRNA (+)strand	NP_786919.1	
VESV	Vesicular exanthema of swine virus	Caliciviridae	Vesivirus	ssRNA (+)strand	NP_786896.1	
AsV178	Aspergillus mycovirus 178	Totiviridae	Totivirus	dsRNA	ABX79995.1	
CmRV	Coniothyrium minitans RNA virus	Totiviridae	Victorivirus	dsRNA	YP_392467.1	
BFV	Botryotinia fuckeliana totivirus 1	Totiviridae	Totivirus	dsRNA	CAM33265.1	
AfV-S1	Aspergillus foetidus slow virus 1	Totiviridae	Victorivirus	dsRNA	CCD33024.1	
AfV-S2	Aspergillus foetidus slow virus 2	unclassified	-	dsRNA	HE588148	
BbTmV-1	Beuveria bassiana virus 1	unclassified	-	dsRNA	-	
ACD	Amasya cherry disease-associated mycovirus	Partitiviridae	Partitivirus	dsRNA	CAJ29958.1	

RNV	Rosellinia necatrix partitivirus 3	Partitiviridae	Partitivirus	dsRNA	BAM36401.1	
FvBV	Flammulina velutipes isometric virus	Partitiviridae	unclassified	dsRNA	BAH08700.1	
VdPV1	Verticillium dahlia partitivirus 1	Partitiviridae	Partitivirus	dsRNA	AGI52210.1	
AfuPV1	Aspergillus fumigatus partitivirus 1	Partitiviridae	Partitivirus	dsRNA	CAY25801.2	
SsV	Sclerotinia sclerotiorum partitivirus 1	Partitiviridae	Partitivirus	dsRNA	AFR78160.1	
FUPO-1	Fusarium poae virus 1	Partitiviridae	Partitivirus	dsRNA	NP_624349.1	
AoR1	Aspergillus ochraceous virus	Partitiviridae	Partitivirus	dsRNA	ABV30675.1	
MoCV1	Magnaporthe oryzae chrysovirus 1	Chrysoviridae	Chrysovirus	dsRNA	YP_003858286.1	
VdCV1	Verticillium dahlia chrysovirus 1	Chrysoviridae	Chrysovirus	dsRNA	ADG21213.1	
AfuCV	Aspergillus fumigatus chrysovirus	Chrysoviridae	Chrysovirus	dsRNA	CAX48749.1	
CNV1	Cryphonectria nitschkei chrysovirus 1	Chrysoviridae	Chrysovirus	dsRNA	ACT79257.1	
PcV	Penicillium chrysogenum virus	Chrysoviridae	Chrysovirus	dsRNA	AAM95601.1	4 segments (2.9-3.6 kbp). Isometric (35-40 nm)
CHV-b	Cryphonectria parasitica bipartite mycovirus 1	-	-	dsRNA	YP_007985675.1	
CHV-m	Cryphonectria parasitica mitovirus 1-NB631	Narnaviridae	Mitovirus	ssRNA (+)strand	NP_660174.1	Unencapsidated
HAsV-8	Human astrovirus 8	Astroviridae	Mamastrovirus	ssRNA (+)strand	AAF85963.1	
RAsV	Rabbit astrovirus TN/2208/2010	Astroviridae	Mamastrovirus	ssRNA (+)strand	AFD34222.1	
DaRV	Diaporthe ambigua RNA virus 1	unclassified	-	ssRNA (+)strand	AAF22958	Unencapsidated

HRV	Human rhinovirus A	Picornaviridae	Enterovirus	ssRNA (+)strand	NP_740401.1	No 5' cap, poly(A)
FMDV	Foot and mouth disease virus	Picornaviridae	Aphthovirus	ssRNA (+)strand	AHW45732	No 5' cap, poly(A)
HAV	Human hepatitis A virus	Picornaviridae	Hepatovirus	ssRNA (+)strand	GNNYHB	No 5' cap, poly(A)
MeasV	Measles virus	Paramyxoviridae	Morbillivirus	nonsegmented (-)stranded RNA	NC_001498	5' cap, poly(A)
RabV	Rabies virus	Rhabdoviridae	Lyssavirus	nonsegmented (-)stranded RNA	AFP36387	5' cap, poly(A)
AfV-F	Aspergillus foetidus virus-F	Alternaviridae	Alternavirus	dsRNA	CCD33020	4 segments, encapsidated, poly (A)
AaV-1	Alternaria alternata virus 1	Alternaviridae	Alternavirus	dsRNA	YP_001976142	4 segments (1.5-3.6 kbp), Isometric (33 nm), poly (A) 33-50 nt, hypovirulent
RnQV1	Rosellinia necatrix quadrivirus 1	Quadriviridae	Quadrivirus	dsRNA	YP_005097975	4 segments (3.7-4.9 kbp), Isometric (45 nm)
SsDRV	Sclerotinia sclerotiorum debilitation-associated RNA virus	Alphaflexiviridae	Sclerodarnavirus	ssRNA (+) strand	YP_325662	5.47 kb
CHV1	Cryphonectria parasitica hypovirus 1	-	Hypovirus	ssRNA (+) strand	NC_001492	Capsid-less (12.73 kb)

2.14 Northern blot analysis of AfuTmV-1 dsRNAs from Af293 isolate

Northern hybridisation analysis was performed to verify the authenticity of the cDNA clones generated with the individually purified dsRNAs. The RNA probes were produced from representative cDNA clones (Frous 1kb[1], BamF, Race3a and Bamr11F) of dsRNAs1-4 using the transcription-labelling of RNA with digoxigenin (DIG)-11-dUTP (Roche Applied Science, Germany), facilitating the detection of single dsRNA bands. Hybridisation showed strong signals for the four dsRNAs with individual probes specific for each component. The result confirmed that sequences obtained corresponded to all of the four dsRNAs (Fig. 37).

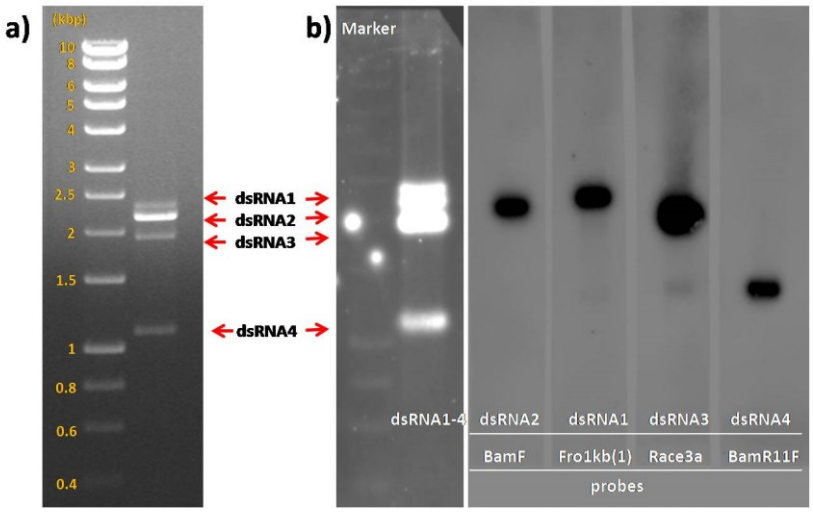


Fig. 37 Northern hybridisation of AfuTmV-1 genomic dsRNAs. Purified viral dsRNAs were separated in 1% agarose gel in 1xTAE, denatured and blotted onto nylon membrane. Transferred RNA were hybridised to each probe is shown on the right (b). Agarose gel electrophoresis of dsRNAs is shown on the left (a). Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

2.15 Purification of the AfuTmV-1 VLP and SDS-PAGE analysis

AfuTmV-1 VLP from both wild type Af293 (virus-infected) and NK125 (virus-free) strains were subjected to virus purification protocol with PEG ultracentrifugation, following CsCl density equilibrium centrifugation. The VLPs was seen as a clear colloidal band in the CsCl gradient (Fig. 38a). Prior to investigation of the VLPs by transmission electron microscopy (TEM), purified VLP fractions of wild type Af293 (virus-infected) and NK125 (virus-free) strains were dialysed in TE buffer.

Negative staining and TEM of VLPs with 1% (w/v) uranyl acetate of the fraction purified from the Af293 isolate revealed some amorphous electron-dense material (small molecules with a particular round shape) (LEO 906E, Zeiss, Germany) but nothing in the NK125 isolate (Fig. 38b), indicating the absence of virus particles in the samples. This observation suggests that

AfuTmV-1 is not conventionally encapsidated in virions and is consistent with the fact that none of the AfuTmV-1 dsRNAs encode a protein identified in BLASTX searches as a capsid protein. It can also be assumed that AfuTmV-1 dsRNAs are associated with or enveloped in colloidal proteinaceous components.

Subsequently, atomic force microscopy (AFM) imaging was performed by A. Patrick Gunning at the Institute of Food Research (IFR). A total 10 µl of PEG-purified dsRNA sample was dropped onto a muscovite mica, dried and observed under a silicon cantilever OMCL-AC240TS (Olympus). RNase III -treated sample was also performed to confirm that the AFM imaging was dsRNA. Following AFM of purified AfuTmV- 1, chain-like molecules were visualised in virtually every scanned area in the virus sample but not in the RNase III-treated control sample (Fig. 38c). Numerous linear strands of different lengths were observed with structures very similar to AFM images of dsRNA observed previously by Magae (2012) of an unusual, apparently unencapsidated mycovirus of *Lentinula edodes* which appears to be associated with a protein which has some similarities to the AfuTmV-1 PAS-rich protein. The lengths of a number of RNA molecules were measured and processed using the Image J programme and allowing for the occurrence of damaged RNA strands, representative dsRNAs of all 4 AfuTmV-1 dsRNAs based on a value of 1 µm = 3 kb dsRNA were observed.

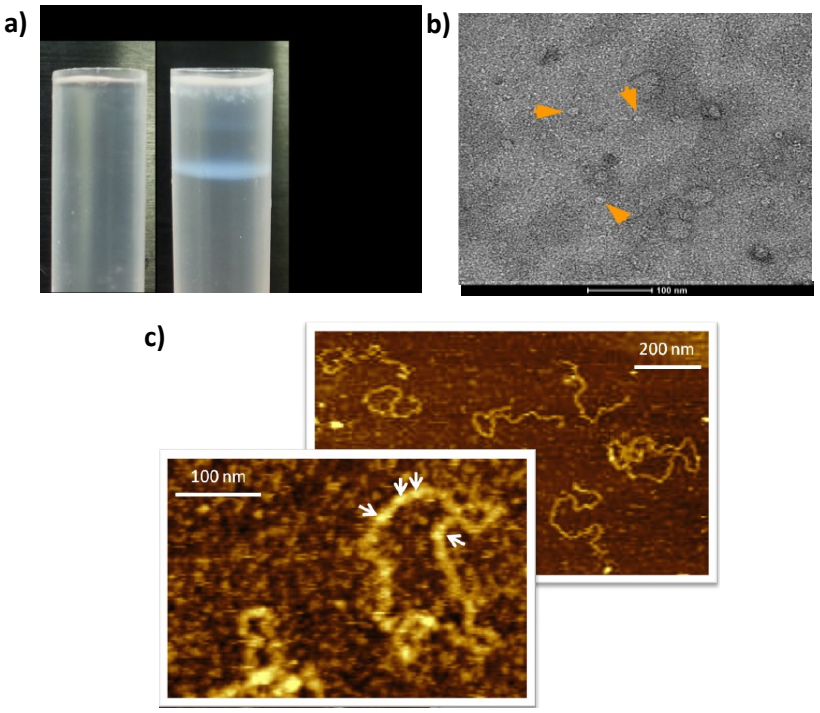


Fig. 38 Virus-like particles of AfuTmV-1. CsCl density (1.45 g/cm^3) centrifugation of AfuTmV-1 fraction (a). TEM of AfuTmV-1 contaning dsRNAs from *A. fumigatus* Af293 (b). Bar = 100 nm. AFM images of purified AfuTmV-1 (c). AfuTmV-1 is visualised as numerous chain-like, linear nucleic acids of different lengths corresponding to those predicted from the genomic size of the four genomic dsRNAs (right

panel). In the left panel, arrows indicate the AfuTmV-1 PAS-rich ORF 4-encoded protein putatively associated with dsRNA 1.

SDS-PAGE analysis revealed that the purified virus preparations contained several major proteins >50 kDa which appeared to be fungal host proteins co-fractionated with the virus samples. One distinct band (Fig. 39 as arrowed) is only present in the virus sample purified from the Af293 virus-infected isolate but not from the NK125 virus-free isolate in denaturing polyacrylamide gels, suggesting that the protein is associated with the virus. The estimated size of the corresponding band in SDS-PAGE analysis was consistent to the deduced amino acid sequence of AfuTmV-1 dsRNA4 with a calculated molecular weight of 29.07 kDa. To identify the protein coding segment, viral protein of *ca*. 30 kDa was subsequently characterized using peptide mass fingerprinting (PMF) by mass spectrometry (MS/MS) analysis. The result yielded 9 peptide fragments derived from tryptic digestion which matched the deduced polypeptide sequence predicted from the sequence of AfuTmV-1 dsRNA4 with ion scores higher than 46 (95% confidence index), accounting for 53% of the entire coverage (279 aa) (Fig. 40).

The investigation on protein encoded by dsRNA4 demonstrated by PMF for the first time that the unusually small AfuTmV-1 PASrp, as compared to the other examples (Table 14), is associated with and possibly coats the dsRNA genome of the virus in an unconventional manner. It is also possible that the AfuTmV-1 dsRNAs are enveloped in colloidal proteinaceous components in the fungal host. However, it is not insignificant that all dsRNA viruses that encode PASrps appear to be unencapsidated, including a *M. lini* mycovirus (Dickinson and Pryor, 1989; Dickinson *et al*., 1993), which has a similar dsRNA profile to AfuTmV-1 and potentially encodes some similar proteins.

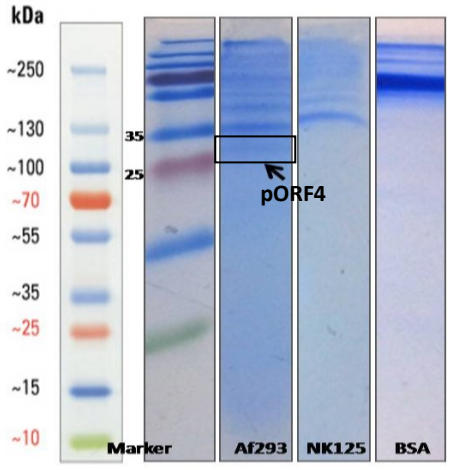


Fig. 39 SDS/PAGE analysis of AfuTmV-1 proteins. Wild type Af293 virus-infected and NK125 virus-free samples were collected from PEG ultracentrifugation. BSA (conc. 1 µg/µl, Promega) was used as a control. Size of the Coomassie blue-stained protein was estimated using protein marker (Page Ruler^{TM} plus prestained protein ladder #26619, Thermo Scientific, USA).

MSPAPILSAEQAAHLAQLTVDDVSLVIKLASLGLKAKEIQQYAAEIAEGGDPEIDLAPGAPRPVVIHAWSFVKDRGQY
HSTYGLSNAQAGELLDLLHTDPKDAVKRIQAIVSERLRLRGSSRPVVVDLEGAPSSTSAKGGAPPSGDGGLSLLRQEIKR
NSGVYGSYQFTAYPTSRPGGQAFAVGLGGGLYAFGQSKAVAVAAARIARVKGRSDDLVRPYIHFWMEGSSPRAASD
IPAKVWDGRLGPHTEPPAEPATKAKTEAPTTSGGRASTPKGADAA

Fig. 40 Coverage of the partial digested protein sequences on the translated protein sequence of AfuTmV-1 dsRNA4. Coverage probability ID; green = very good, blue = moderate, red = weak, and grey = not seen.

Unencapsidated mycoviruses are currently classified in the family *Hypoviridae* or in the genus *Endornavirus* (Nuss *et al*., 2005a and 2005b; Gibbs *et al*., 2005). Unencapsidated dsRNA viruses generally contain multiple genetic elements and are associated with host cell membranes. In addition to AfuTmV-1, a number of unencapsidated dsRNA viruses have been reported such as dsRNA virus associated with spherical membrane vesicles in *Alternaria solani* (Zabalgogeazcoa *et al*., 1997); a well-known vesicle membrane associated dsRNA hypovirus found in *Cryphonectria parasitica* (Hansen *et al*., 1985; Fahima *et al*., 1993); a cytoplasmic unencapsidated dsRNA mycovirus in the rice blast fungus *Magnaporthe grisae* strain MG01 (Chun and Lee, 1997); mitochondrial unencapsidated dsRNA viruses of the plant pathogenic fungus, *Ophiostoma novo-ulmi* (Cole *et al*., 1998); Phlebiopsis gigantea virus 2 (PgV2) (Kozlakidis *et al*., 2009); Fusarium graminearum virus 3 (FgV3) (Yu *et al*., 2009); and Lentinula edodes mycovirus HKB (LeV) found in shiitake mushroom (Magae, 2012). Two dsRNA insect viruses have been reported to be unconventionally encapsidated nominated Spissistilus festinus virus 1 (SpFV1) and Circulifer tenellus virus 1 (CiTV1) (Spear *et al*., 2010). DsRNA viruses from the green alga *Bryopsis* are also capsidless with predicted RdRP genes that are closely related to partitiviruses (Zhang *et al*., 2014). Unencapsidated viruses containing (+) ssRNA and (-) ssRNA have also been described such as in fruit tree fungal pathogen, *Diaporthe ambigua* which is a vesicle membrane associated positive-stranded RNA virus (DaRV; Preisig *et al*., 2000), mycoviruses (e.g. CHV1) infecting *C. parasitica* (Choi and Nuss, 1992; Kazmierczak *et al*., 2012) and yeast narnaviruses (Garcia-Cuellar *et al*., 1997).

2.16 Quantitation of AfuTmV-1 dsRNAs and ssRNAs in infected mycelia

Northern blot hybridisation was performed to examine whether the AfuTmV-1 genome is predominantly present in dsRNA or ssRNA form and to determine the relative abundance of both forms of viral RNAs. Total nucleic acids were prepared from virus-infected Af293 and isogenic virus-free isolate (NK125) and fractionated using LiCl.

The ssRNA (2M LiCl fraction) and dsRNA (4M LiCl fraction) were pre-treated with DNase 1 alone or in combination with S1 nuclease. The viral RNAs are present almost exclusively in the dsRNA form and were insensitive to DNase 1 and S1 nuclease (Fig. 41; top panels). AfuTmV-1 ssRNA could not be visualized by ethidium bromide staining of nucleic acids separated by agarose gel electrophoresis. Northern blot analyses using riboprobes specific for plus-stranded and minus-stranded AfuTmV-1 dsRNA1, were used to individually hybridise with the nuclease-treated ssRNA and dsRNA of two isogenic *A. fumigates* lines (Af293 and NK125) on duplicate blots. The hybridisation results for virus-infected Af293 showed that ssRNA1 and dsRNA1 migrated with the same mobilities under the electrophoresis condition employed in this study. It can be seen that signal intensities of the pre-treated dsRNA1 detected with plus(+)-strand specific riboprobe were distinctly stronger than those detected by minus(-)-strand specific riboprobe, suggesting that the positive stranded RNA1 was present in a greater quantity than the negative stranded RNA1 (Fig. 41; lane 1, 2, 5 and 6; bottom panels). However, differential signal intensity between plus(+)- and minus(-)-stranded RNAs is unlikely which could be a result from the riboprobe length and specificity used for hybridization. Weak signal was detected from the hybridisation of DNase-treated ssRNA to plus(+)-strand specific riboprobe (Fig. 41; Lane 3) and none from the other pre-treated ssRNA preparations (Lanes 4, 7 and 8), confirming that only small amounts of the (+)ssRNA form of dsRNA1 are present in infected mycelia. Since ssRNA and dsRNA were both extracted from the same amounts of total nucleic acid, it could be inferred that most of the AfuTmV-1 dsRNA1 is predominantly present in the form of genomic dsRNA rather than the replicative form of (+) ssRNA. This observation is similar to a finding previously reported for the AaV-1 dsRNA mycovirus (Aoki *et al*., 2009). No signals were detected in fractionated nucleic acids isolated from the virus-free NK125 isolate. However, these findings differ from those for the dsRNA mitochondrial viruses (mitoviruses) in the Dutch elm disease fungus O*phiostoma novo-ulmi* (Hong e*t al*., 1999). Here the (+) ssRNA form of the mycovirus RNA predominates over dsRNA. However this result is anticipated since the genome of mitoviruses is comprised of ssRNA rather than dsRNA where the reduced amounts of the latter would represent the replicative form of the virion RNA.

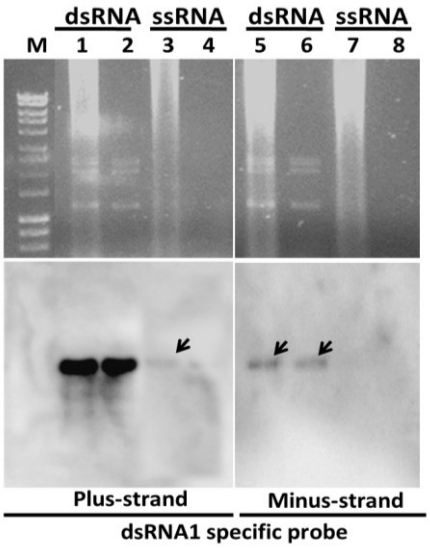


Fig. 41 Comparison between the amount of viral ssRNA and dsRNA of AfuTmV-1 dsRNA1. Amounts of viral ssRNA and dsRNA of the virus-infected isogenic line (Af293), equivalent to 10 mg of total nucleic acids before LiCl fractionation, were electrophoresed in 1.0% non-denaturing agarose gel. Samples were pre-treated with 1 unit of DNase alone (lane 1, 3, 5 and 7; top panels) or DNase+S1 nuclease (lane 2, 4, 6, and 8; top panels); Northern blot hybridisation of viral ssRNA and dsRNA of the virus-infected isogenic line (Af293) using plus(+)-strand specific riboprobe and the minus(-)-strand specific riboprobe (bottom panels). Hyperladder 1 (M; 10 kbp; Bioline) was used as a marker.

Northern blot hybridization of four dsRNAs was also performed with the probes specific for positive-stranded and negative-stranded RNAs. The result showed that dsRNAs of AfuTmV-1 are present in much larger amounts than positive-stranded RNAs for all four elements while no signal for negative-stranded RNA was detected (Fig. 42). This confirmed that dsRNA is the genome of AfuTmV-1, not the replicative form of an RNA virus genome.

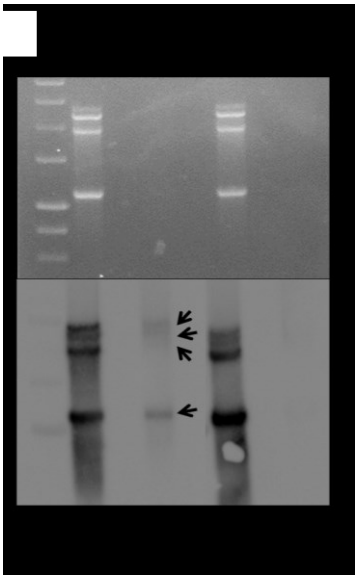


Fig. 42 Comparison between the amount of viral ssRNA and dsRNA of AfuTmV-1 dsRNAs. Amounts of viral ssRNA and dsRNA of the virus-infected isogenic line (Af293), equivalent to 20 mg of total nucleic acids before LiCl fractionation, were electrophoresed in 1.0% non-denaturing agarose gel. Samples were pre-treated with 1 unit of DNase+S1 nuclease for dsRNA samples and Dnase+RNaseIII for ssRNA samples

(top panel); Northern blot hybridisation of viral ssRNA and dsRNA of the virus-infected isogenic line (Af293) using plus(+)-strand specific riboprobe and the minus(-)-strand specific riboprobe (bottom panels). Hyperladder 1 (M; 10 kbp; Bioline) was used as a marker.

2.17 *In vitro* replication of the AfuTmV-1 dsRNAs

In order to explain the mode of replication and transcription of the AfuTmV-1 in fungal cells, RdRP activity was examined using the AfuTmV-1 VLPs to investigate whether RNA replication could occur in a cell-free replication system. VLPs were incubated in the presence of nucleotide triphosphates (ATP, CTP, GTP and UTP) and Actinomycin D serving as host RdRP and DdRP inhibitor, then incubated at 37°C for 16 h. DIG-labelled ssRNAs synthesized during *in vitro* replication were subsequently used as hybridization probes in dot blot hybridization against each ssRNA strand (positive and negative strands) obtained from four plasmids corresponding to dsRNAs1-dsRNA4.

As a result, no viral RNA replication product was synthesized in this system because no hybridization signals were detected. The reason for the failure to synthesize RNA may result from incapability of the viral RNA to generate RdRP or other viral proteins required for the replication process. Since attempted replication was performed in a cell-free system, lack of host factors such as host replication proteins or host endomembranes essential for membrane-bound replication complex formation could also affect the replication of the virus *in vitro,* particularly for an unencapsidated virus. To circumvent the problem, *in vitro* translation of the AfuTmV-1 replication proteins required for RNA replication in either fungal cell extract or non-host extract such as commercial rabbit reticulocyte lysate could be done prior to using the translation reaction to *in vitro* replication system. Another strategy which might be used to study *in vitro* replication of the virus RNA is using purified RdRP isolated from the virus infected host to catalyse complete replication of template RNA.

2.18 *In vitro* translation of the AfuTmV-1 dsRNAs

The coding potential of the AfuTmV-1 viral genome was investigated by *in vitro* translation in the TNT® coupled rabbit reticulocyte lysate system using heat denatured total dsRNAs (1 µg). The reactions were incubated over a time courses (30, 60 and 90 min) and the products analyzed by PAGE. The results showed that translation products of the four dsRNAs were

actively synthesized after a 60-90 min incubation period. No AfuTmV-1 related products were synthesized when denatured dsRNA was omitted from the reactions (Fig. 43a).

In all cases translation products with similar sizes to those predicted from amino acid sequences of the four genomic dsRNAs were observed. A prominent translation product at 76 kDa corresponding to dsRNA2 ORF2 was synthesized in the lysate containing denatured dsRNAs. This result suggests that the translation capability of dsRNA2 was greater than those of other dsRNA segments since the incorporation of [^{35}S] methionine was greater, especially as compared to dsRNA1. Two smaller translation products were detected with estimated sizes of 67 and 29 kDa corresponding to the predicted amino acid sequences of dsRNA3 ORF3 and dsRNA4 ORF4, respectively (Fig 43a). The size of the translation product of dsRNA4 ORF4 obtained following *in vitro* translation corresponded to that of SDS-PAGE analysis of the purified viral proteins. Smaller translation products (e.g. *ca*. 45 kDa), which were not virus encoded, were also found. These polypeptides are presumably incomplete translation products resulting from premature termination of translation and some of them were found to be translation backgrounds as found in the control assay.

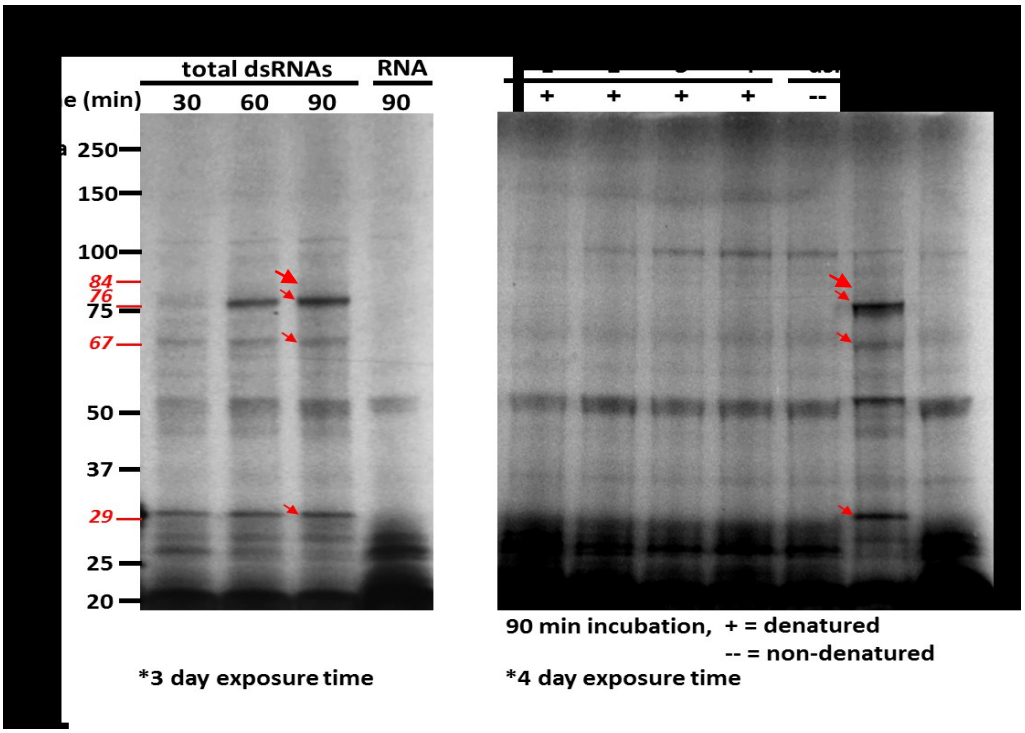


Fig. 43 Autoradiograms of *in vitro* translation products of the AfuTmV-1 dsRNAs. Total *in vitro* translation products from unfractionated (a) and fractionated dsRNAs (b) were analyzed on NuPAGE® Novex® 4-12% Bis-Tris protein gel (Invitrogen). Translation mixes in panel a were incubated in varied time course (30, 60 and 90 min) whereas the mixtures in panel b were incubated for 90 min. Negative control sample (no RNA) was included. Precision plus protein™ all blue standard (BIO-RAD) was used as marker proteins. Arrows indicate major translation products from each sample.

Translation of non-denatured dsRNAs and heat denatured dsRNAs was performed to determine the difference in translation capacity between them. The results revealed that translation products were synthesized only when denatured dsRNAs were used as template while no products were detected in non-denaturated sample (Fig. 43b; lane 5 and 6). Simultaneously fractionated individual denaturated dsRNAs (1 µg) were *in vitro* translated to determine translation products of the four individual dsRNA ORFs. These results indicate that dsRNA formation might inhibit translation and *in vivo* transcription of AfuTmV-1 dsRNAs might require specific helical host proteins to unwind the dsRNA strands to initiate translation. However, it would be more beneficial to directly perform the same experiment in the fungal host extract to investigate whether host cellular proteins are required for virus transcription and translation processes.

In conclusion the four AfuTmV-1 dsRNAs appear to be monocistronic since each dsRNA contains a single ORF and translates into a single major product of the expected size. *In vitro* translation of denatured dsRNA has been demonstrated for several dsRNA viruses including *Ustilago maydis* virus strain P1, beet crytic virus 1 (BCV 1) and a segmented dsRNA virus from *Rhizoctonia solani* (Dalton *et al*., 1985; Accotto *et al*., 1987; Finkler *et al*., 1988; Tavantzis and Bandy, 1988). To my knowledge, *in vitro* translation of non-denatured dsRNA has not been reported thus far.

---

3. Biological comparison of virus-cured and virus-infected fungal strains

3.1 Curing *A. fumigatus* Af293 wild-tpye isolate with cycloheximide

The experiment aimed to eliminate AfuTmV-1 dsRNA elements from the *Aspergillus fumigatus* clinical isolate Af293 by cycloheximide treatment. Sequential hyphal tip isolation was performed to reduce virus titre by inoculation on *Aspergillus* complete medium (ACM) agar supplemented with different concentrations of cycloheximide (0, 75, 100, 125 and 150 mM) and passaged for 5 generations.

Fungal cultures showed a correlation between the fungal growth and concentrations of cycloheximide used (Fig. 44). The highest concentration of cycloheximide (150 mM) inhibited fungal growth therefore this concentration was excluded for further passage. After culture passage for 5 generations, cycloheximide-treated inocula of each concentration were

recovered on normal ACM plates. A total of 3 isolates (nominated NK105, NK106 and NK125) were selected only from the plates containing 100mM and 125mM cycloheximide based on a change in pigmentation of the fungus.

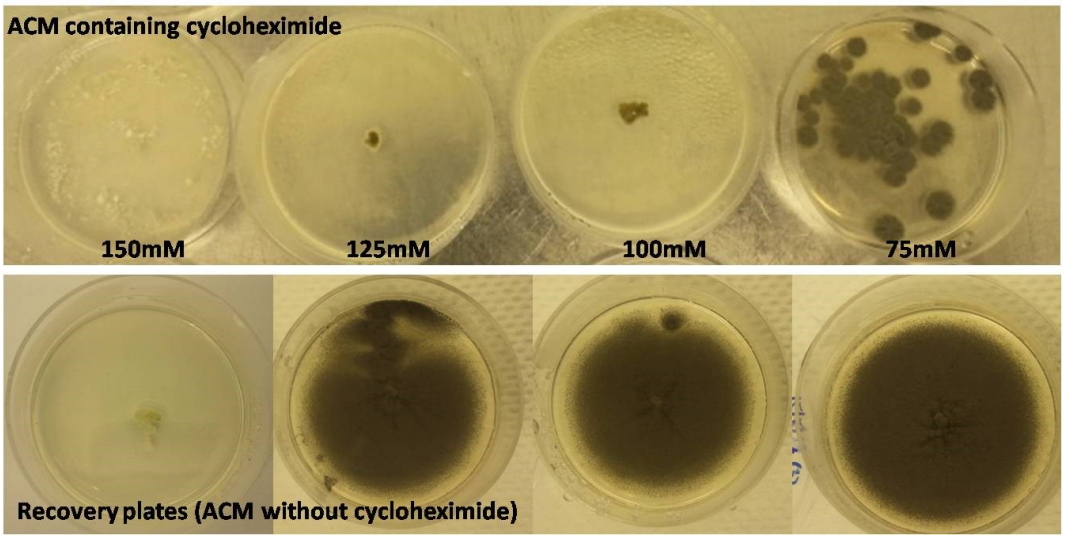


Fig. 44 Comparison of fungal growth on the different concentrations of cycloheximide (75-150 mM), top panel; the recovery of fungal culture on normal ACM plates after treatment with cycloheximide, bottom panel.

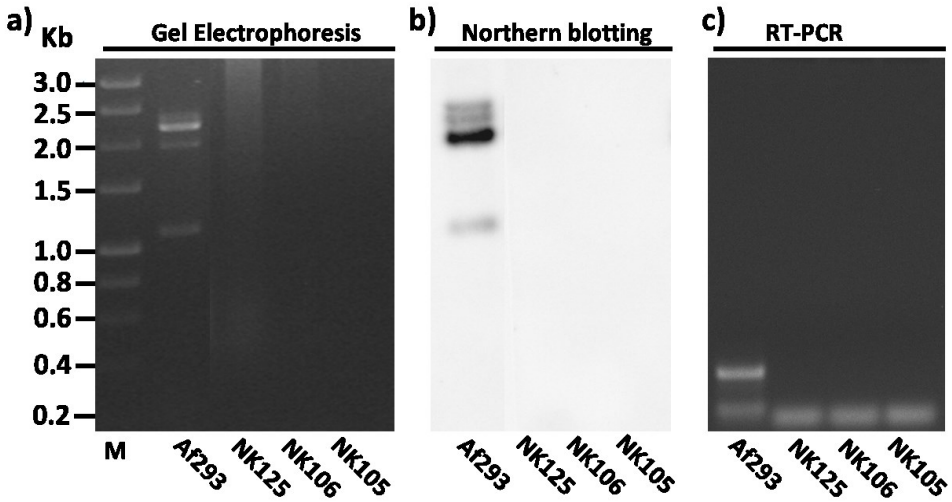


Fig. 45 Agarose gel electrophoresis profile of AfuTmV-1 dsRNAs isolated from *A. fumigatus* Af293 wild-type and cycloheximide-cured isolates (a); Northern blot hybridization (b); RT-PCR amplification of a fragment of AfuTmV-1 dsRNA2 (c). Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

Following cycloheximide treatment of *A. fumigatus* Af293 isogenic line, virus-cured isolates were confirmed as being virus-free by electrophoretic analysis of dsRNAs, northern blotting and RT-PCR amplification assay of a fragment of AfuTmV-1 dsRNA2. No dsRNA bands or RT-PCR amplicons of the expected size (340 bp) was visible as compared to the dsRNA control

samples extracted from the Af293 wild-type, suggesting that AfuTmV-1 was eradicated completely from all three cured isolates. Northern blot hybridisation was also performed to confirm the loss of all four genomic AfuTmV-1 dsRNAs (Fig. 45).

Cycloheximide has been used to eliminate dsRNA viruses from many fungi such as Alternaria alternata virus-1 (AaV-1) and mycoviruses infecting *Rhizoctonia solani* and *Curvularia protuberate* (Strauss *et al*., 2000; Aoki *et al*., 2009). However, those viruses were not completely eliminated from the fungal host in contrast to the AfuTmV-1 which required relative lower concentration of 100mM cycloheximide to completely eliminate the virus from the Af293 isolate after five passages. Additionally cycloheximide has been used to cure *Aspergillus* spp. of mycovirus infection including *A. fumigatus* chrysovirus (AfuCV; Jamal *et al*., 2010) where high concentrations of cycloheximide >150 mM and six passages were required to completely eliminate the virus. Since cycloheximide inhibits initiation and elongation steps of protein synthesis, it is feasible that viral RNAs lacking a poly (A) tail e. g. AfuTmV-1 are more vulnerable to cycloheximide than RNAs with a poly (A) tail. The assumption correlates with the observation that AaV-1 and partitiviruses found in *R. solani* which are polyadenylated at the 3'-termini were resistant to cycloheximide treatment plus the fact that individual dsRNA elements with short tails rather than long tails are more susceptible (Aoki *et al*., 2009).

3.2 Effects of AfuTmV-1 on growth and phenotype of *A. fumigatus* Af293

The phenotypes of the virus-infected Af293 and virus-free *A. fumigatus* isolates, together with their growth rate and biomass production were respectively observed and measured over a 5-7 day incubation period in both solid and liquid AMM and ACM media.

Morphological comparison revealed that the three isogenic virus-free lines showed sectored growth and slightly darker spore pigmentation (Fig. 46). However, it is unclear whether these phenotypic alterations were due to an artefact related to a side-effect stimulated by cycloheximide treatment resulting in genetic instability and morphological variation or if they were due to an interaction with AfuTmV-1 infection. In addition to sectoring and slight changes in pigmentation, there were no significant microscopic phenotypic changes found in virus-infected as compared to virus-free fungal isolates. For instance in the virus-infected Af293 isolate there were no alterations to conidial shape or mycelial integrity as compared to the virus-free NK125 isolate. However a minor change of abnormally enlarged vesicles appeared in the mycelia of the virus-infected Af293 isolate whereas only a few small vesicles were

observed in the mycelia of the virus-free NK125 isolate. This is similar to a finding reported for another unencapsidated dsRNA mycovirus of *Alternaria solani* (Chun and Lee, 1997; Zabalgogeazcoa *et al*., 1997). Since the virus-infected and virus-free *A. fumigatus* Af293 isolates do not show clear differences in colony morphology, ultrathin section of both isolates could be investigated to determine if the differences of organelles and ultrastructures reside within the hyphae to confirm an effect of the virus on the fungus. Such effects have been reported in several virus-infected fungi such as in *F. graminearum* isolate 98-8-60 (Chu *et al*., 2004) where virus-infected fungal mycelia possessed smaller nuclei and thicker cell walls as compared to the virus-free isolate despite no external morphological differences between the two. At first glance it would appear that AfuTmV-1 is latent/cryptic in its host but pathogenicity assays need to be performed to assess this.

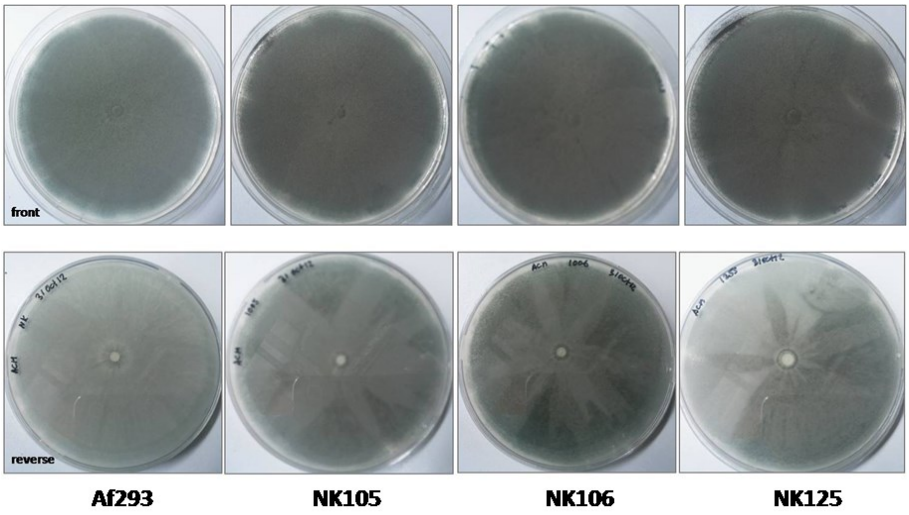


Fig. 46 Colony comparison between the virus-infected Af293 and three isogenic virus-free isolates (NK105, NK106 and NK125) grown on ACM at 37 ˚C for 5 days.

From the radial growth assay, three virus-free isolates appeared to grow slightly faster than the virus-infected control but statistical analysis using Student's t test revealed no significant differences in radial growth rates between *A. fumigatus* Af293 and virus-free isogenic lines in either medium. Similar results were obtained when biomass assays were performed which showed that three virus-free isogenic lines apparently generated higher biomass than the virus-infected line. However Student's t test revealed no significant differences between the two. Only the virus-free NK125 isolate grown in ACM medium showed significantly greater biomass production than the virus-infected Af293 when statistical analysis was performed (Fig. 47). Altogether reviewing the observations of colony morphology, mycelial growth and biomass production, it would appear that AfuTmV-1 has a minor effect on the virulence level and growth of the host fungus.

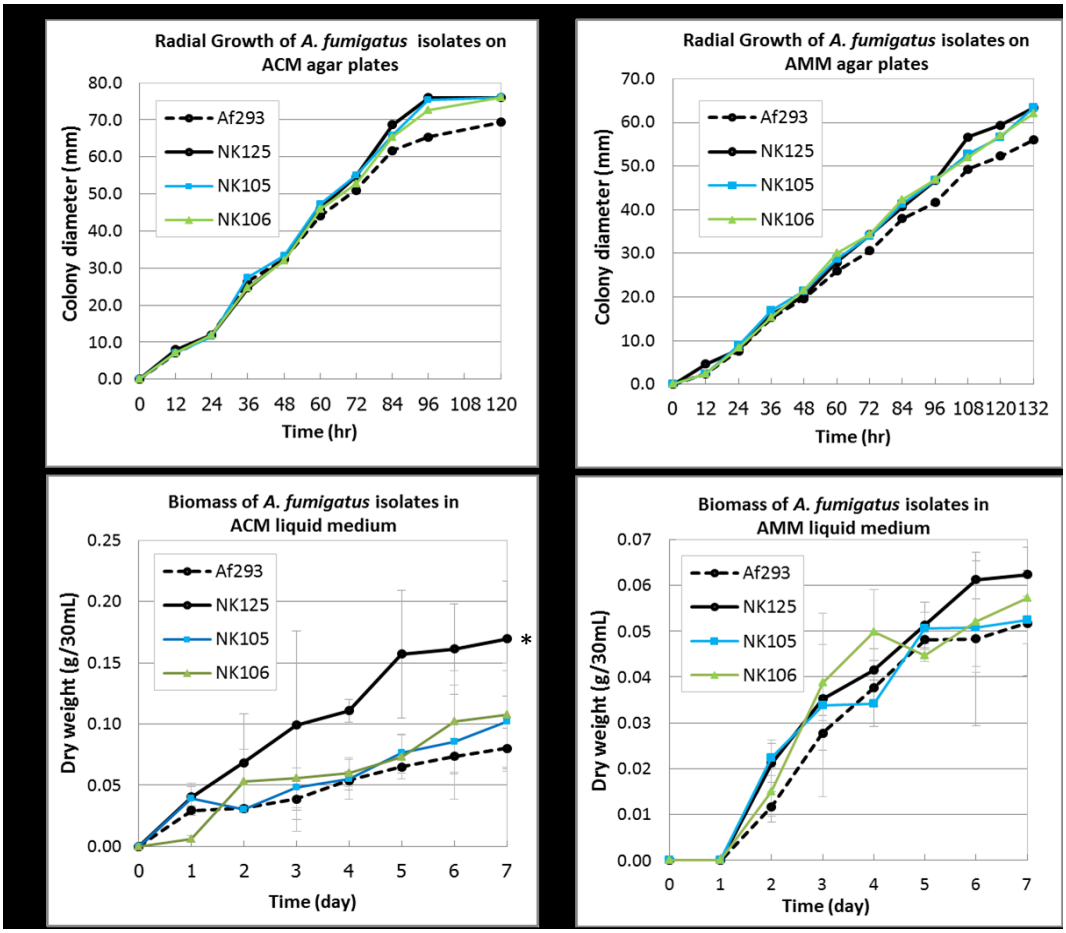


Fig. 47 Comparison of the fungal growth under the different cultivation conditions. Radial growth of *A. fumigatus* on the ACM plates (a); Radial growth of *A. fumigatus* on AMM plates (b); Biomass production of *A. fumigatus* in ACM broth (c); and Biomass production of *A. fumigatus* in AMM broth (d).

4. Infectivity of purified AfuTmV-1 and its dsRNAs

In nature RNA mycoviruses are not infectious for fungi. However, extracellular transmission of mycoviruses can be achieved by introducing the transfecting purified virus into fungal protoplasts followed by replication of the virus in the cytoplasm and eventual spread through the hyphal network. Protoplasts from the virus-free *A. fumigatus* isolate NK125 were produced using a modified protocol to that originally described by Szewczyk *et al*. (2006) which yielded of *ca.* 10^7 protoplasts/50ml (Fig. 48). The protoplasts were transfected with in the presence of PEG and used several inocula including purified AfuTmV-1 VLPs and AfuTmV-1 dsRNAs. Transfectants were plated onto an SMM recovery medium and finally ACM agar plates following initial incubation at 30 °C overnight followed by incubation at 37 °C for 1-2 days. No colonies were found when only virus and dsRNA preparations were plated on SMM medium, indicating that these preparations were free from fungal contaminants.

In initial experiments the transfection efficiency of protoplasts was very low and 0/48 subcultures derived from 4 regeneration plates (12 regenerants from each plate) appeared to be virus-free when extracts were examined using small scale isolation procedure and gel eletrophoresis of the AfuTmV-1 dsRNAs. Since the number of the regenerants recovered from each plate was high (>50 colonies at 10^{-3} serial dilution) with no obvious morphological or growth rate differences between virus-infected and virus-free regenerants it is likely that some transfected colonies were missed. To circumvent this eventuality in further experiments, regenerants from each regeneration plate were pooled and examined for the presence of virus. Mycelial and conidial mass were taken from regenerated *A. fumigatus* transfected protoplasts and subsequently transferred into ACM broth and incubated at 37 °C for 7 days. Pooled regenerated colonies, transfected with purified AfuTmV-1 and AfuTmV-1 dsRNAs were nominated ATV-1 and ATR-1 respectively.

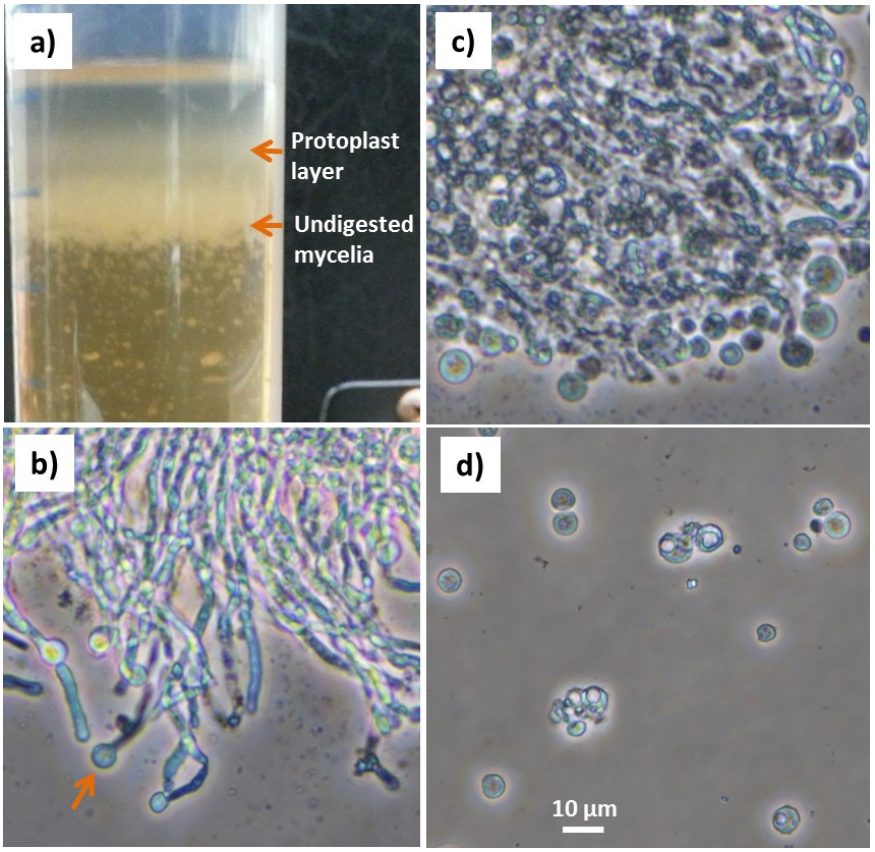


Fig. 48 Protoplast production from the virus-free *A. fumigatus* NK125 isogenic line. Separation of the protoplasts from digested and undigested mycelia in OM protoplasting solution (a). Protoplast formation at the hyphal tips seen as bulges and swelling forms (b). Digested hyphal mat showing several small round-shaped protoplasts released from hyphal mat (c). Protoplasts after purification on the sorbitol cushion, showing swelling of the protoplasts and vacuole formation (d). Scale bar = 10 µm.

The colony morphology and hyphal growth pattern of all recovered transfectants were identical to those of isolate Af293 and only differed slightly from the NK125 isolate in terms of reduced pigmentation under the same culture conditions, suggesting that AfuTmV-1 has no overt effect on colony morphology and fungal pathogenicity (Fig. 49d). Since morphology of the transfected lines and the virus-free strain are indistinguishable, the presence and integrity of viral dsRNA all transfectants was examined and analysed by gel electrophoresis. Total nucleic acids were extracted and purified from 25 g of transfectants using LiCl. As a result, four dsRNA bands were clearly visualized from the ATV-1 and ATR-1 samples with the same mobility as AfuTmV-1 dsRNAs purified from the naturally infected mycelia and were resistant to DNase and S1 nuclease (Fig. 49a). It was observed that yields of dsRNA from transfected lines ATV-1 and ATR-1 were approximately 3-5 fold less than those obtained from naturally infected mycelia from strain Af293 when equal amounts of 25 g wet weight of tissue were processed. RT-PCR amplification was also performed to confirm the integrity of AfuTmV-1 using specific primers to amplify a fragment of dsRNA2, yielding a 360 bp PCR product (Fig. 49b). This confirmed that *A. fumigatus* Af293 protoplasts could be successfully infected with both AfuTmV-1 VLPs and dsRNAs.

Northern blot analysis of RNA enriched in dsRNA or ssRNA of ATR-1 was performed with plus- or minus-strand specific DIG-labelled riboprobes derived from all four AfuTmV-1 dsRNAs. Total nucleic acid was purified using the RNeasy Plant Mini Kit (QIAGEN) and treated with either S1 nuclease or RNAse III to recover dsRNA and ssRNA, respectively. This analysis revealed that under the conditions of the electrophoresis employed here the ss and ds forms of AfuTmV-1 dsRNAs had similar electrophoretic mobilities (Fig. 49c). The ssRNAs and dsRNAs were distinguished by their capability to hybridize with the plus-strand specific riboprobes or minus-strand specific riboprobes. The results indicated that nucleic acid extracts from ATR-1 transfectants contain, for all four RNAs, both dsRNA and positive-stranded ssRNA. Because the amounts of ssRNA and dsRNA analyzed were derived from the same amounts of total nucleic acids, the relatively strong signals of dsRNA and weak signals of ssRNA from blotting indicated that each of the RNAs exists predominantly in the dsRNA form. No significant amounts of negative-stranded ssRNAs were found for all four RNAs when hybridised with the minus-strand specific riboprobes, suggesting no expression of the short ORFs on the negative strand. These experiments were performed on three separate occasions with identical results and confirm that the genome of AfuTmV-1 is indeed dsRNA. The structural characteristics of the AfuTmV-1 sequence is further evidence in support of the conclusion that AfuTmV-1 dsRNA is not a replication intermediate of an ssRNA virus. This is dissimilar to the situation with *C. parasitica*

hypoviruses where dsRNAs isolated from infected mycelia are indeed the replicative form of an ssRNA virus (Nuss, 2005b). In conclusion the transfection experiments revealed that both purified AfuTmV-1 and AfuTmV-1 dsRNA are infectious for *A. fumigatus*.

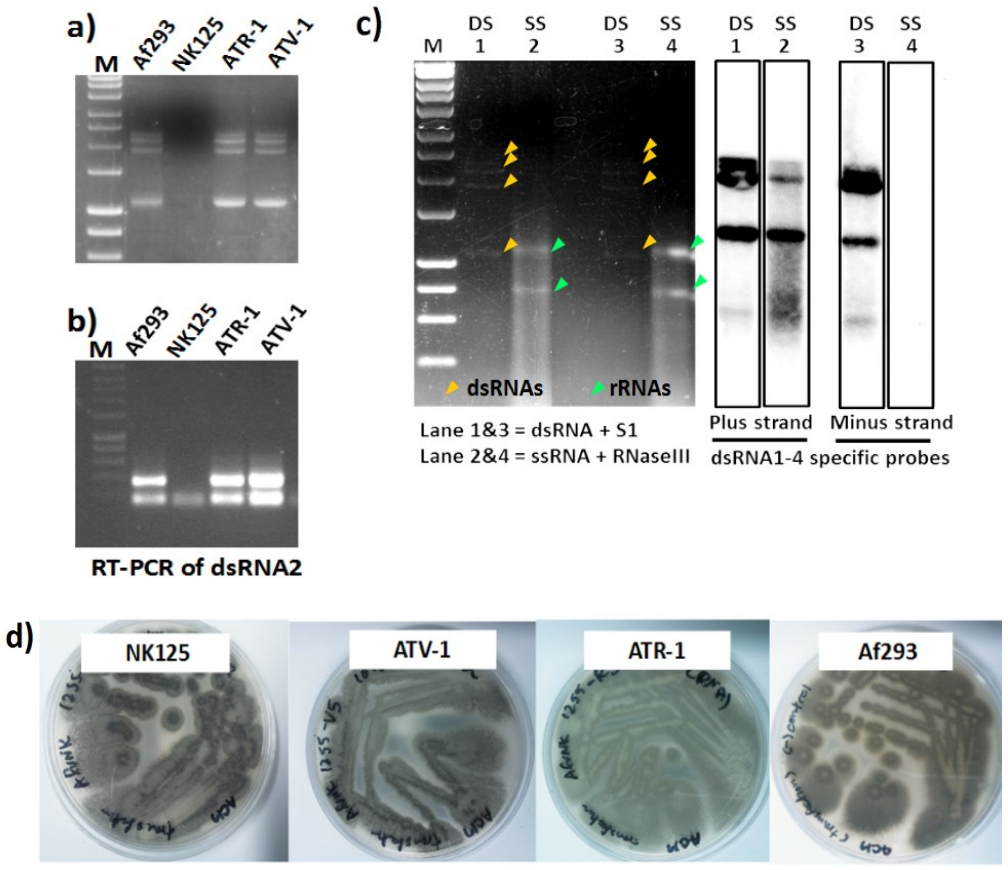


Fig. 49 Transfection of PEG-mediated protoplast produced from virus-free *A. fumigatus* isolate NK125 with purified AfuTmV-1 VLPs and dsRNAs. Gel electrophoresis of dsRNA extracted by LiCl extraction from virus-infected wild type Af293, virus-free isolate NK125, ATV-1 and ATR-1 transfectants (a). Total dsRNAs of ATV-1 and ATR-1, transfected lines were loaded onto the gel, representing whole quantity of dsRNA extracted from 25 g mycelia. While dsRNA bands from the wild-type and virus-free isolate showed amount of dsRNA extracted from *ca*. 5 g mycelia. RT-PCR amplification of dsRNA using sequence specific primers designed to amplify dsRNA2, yielding an amplicon of *ca*. 360 bp (b). Northern blot hybridization of the ATR-1 dsRNAs and ssRNAs to all four plus-strand or minus-strand specific riboprobes (c). Colony morphology of the NK125 virus-free strain, Af293 wild type strain and transfected isolates (d). Hyperladder 1 (M; 10 kbp; Bioline) was used as a marker.

Successful transfection of *A. fumigatus* protoplasts with AfuTmV-1 suggests that the virus replicates in the cytoplasm of the fungal cells and migrates to new hyphal compartments and eventually conidia. Currently there is little information available on the replication of unencapsidated dsRNA mycoviruses. From northern blot analysis, asymmetrical accumulation of positive-stranded (+)ssRNA and dsRNA within host cells could be explained by the lack of the 3'-poly (A) tail required for the initiation of RNA replication in the genomic sequences of

AfuTmV-1, resulting in the existence of low amounts of (+) ssRNA (Spagnolo and Hogue, 2001). However initiation of virus replication might be substituted by stem loop structures at the 3' termini serving as *cis*-acting elements or a self-priming elements to enhance replication and protein expression (Fujimura and Esteban, 2004; Yu and Markoff, 2005). In addition, the infectivity assay demonstated that AfuTmV-1 mycovirus is associated with cryptic/latent infection in *A. fumigatus* Af293 since no morphological alterations of the regenerated colonies were observed.

Proposed replication process of AfuTmV-1

As AfuTmV-1 is unencapsidated and contains 4 genomic segments, it is not known how VLPs are formed and maintained in fungal cells. As described earlier the protein encoded by dsRNA4 might function as a structural protein for the virus, however it remains to be fully elucidated how the protein associates with all four dsRNAs. Low transfection rates of protoplasts with AfuTmV-1 might be the result of poor hit kinetics where all four genomic dsRNAs are necessary for efficient replication of the virus. The absence of any of the four genomic dsRNAs in transfected cells would lead to no infection (Fig. 50).

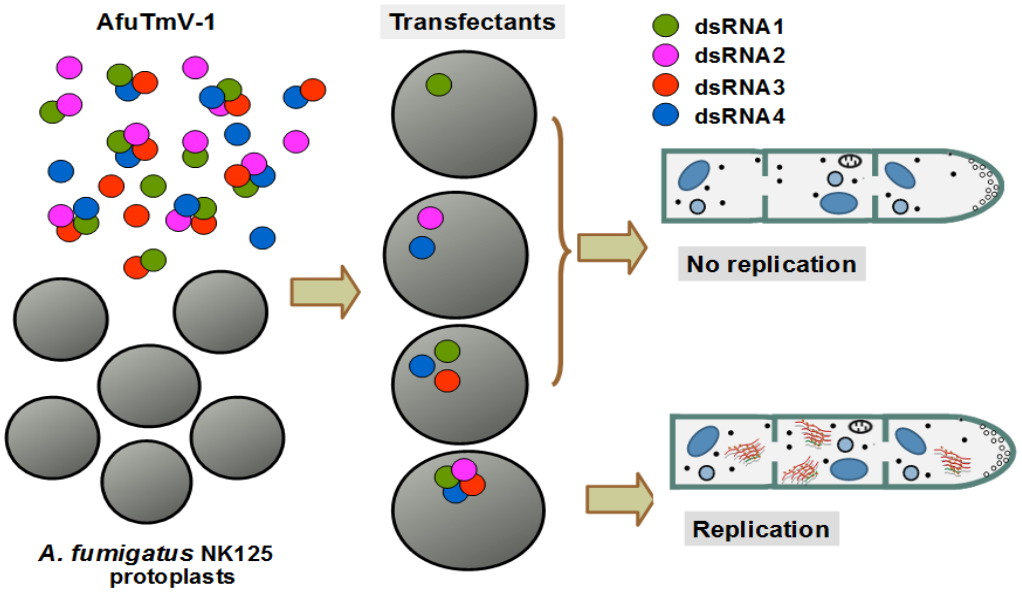


Fig. 50 A putative schematic for the transfection of *A. fumigatus* NK125 protoplasts with AfuTmV-1.

Replication of AfuTmV-1 dsRNAs is initiated by transcription of (+) ssRNA using viral (-) RNA template by viral RdRP and subsequently the translation of viral RNA transcripts using the host translation machinery to synthesize viral RdRP and proteins required for replication initiation. The principal regarding how the (-) RNA template is selected from dsRNAs in the transcription

step is not known. However, it is evident that 5'-capping on only the (+) RNA strand of dsRNAs drives 5'-cap binding site on the polymerase to move close to the 3' terminus of the (-) RNA strand, the process of which subsequently permits initiation of transcription (McDonald *et al*., 2009).

By analogy with other investigations on plant and animal viruses it is conceivable that AfuTmV-1 dsRNA1, which encodes an RdRP, might be membrane associated in a bound complex. To complete the process of RNA replication, helicase is required to separate template and product strands. However, none of the AfuTmV-1 dsRNAs encode a protein associated with conventional helicase activity. Thus partial unwinding of the dsRNA strands may be reliant on some host cellular proteins such as a specific-RNA helicase to allow some transcription. The replication of the viruses with small genomes usually relies on host components, and it can thus be inferred that AfuTmV-1 dsRNA replication likely requires host polypeptides to enhance RdRP and replication activity. The replication of AfuTmV-1 dsRNAs2-4 must depend on the RdRP encoded by dsRNA1. Proteins encoded by dsRNA2 and dsRNA3 might be involved in dsRNA or protein binding activity during virus replication and 5'-RNA-capping methyltransferase activity, respectively, whereas the PAS-rp encoded by dsRNA4, apart from coating VLPs might also be associated with interacting with host membrane components or serve as a structural or scaffold protein anchoring replication complexes or probably assist a host RNA helicase to partially unwind dsRNA. Since the dsRNA3 molecule encodes the putative Mtase capping enzyme, it could be possible that some (+) ssRNA strands are capped to allow initiation of translation whereas some (+) ssRNA strands are not capped and used as substrates for dsRNA replication.

Maintenance of AfuTmV-1 in the fungal cells

Since AfuTmV-1 lacks a CP to protect its genome and does not form conventional virus particles, the mechanism concerning its replication, evasion from host immunological surveillance, survival and maintenance in the host cells are unclear. AfuTmV-1 might form ribonucleoprotein (RdRP/RNA) complexes together with virus encoded proteins (PASrp) or host proteins to form replication complexes and accumulate in the cytoplasm. Alternatively the virus might form membranous-associated structures and reside within the host organelles e.g. endoplasmic reticulum, nucleus and plasma membranes. In this context, if AfuTmV-1 is maintained in the host cells as a membrane-associated virus, its replication process possibly occurs within the endomembrane vesicles where the viral RdRP replication complex is

embedded as has been reported for some ssRNA capsid-less mycoviruses (Solorzano *et al*., 2000; Kazmierczak *et al*., 2012).

Endomembrane vesicles are a favourable site for the replication of some ssRNA mycoviruses where they also protect the virus from fungal antiviral responses (Fahima *et al*., 1993). However, the localization of the AfuTmV-1 RdRP is obscure as the signal peptide and transmembrane helix could not be predicted from the protein sequence. In contrast, it has also been described that the majority of unencapsidated dsRNA viruses are found to locate to the cytoplasm (Nuss and Koltin, 1990; Chun and Lee, 1997). As a consequence viral dsRNA is likely to be degraded by silencing host nucleases following recognition as an invasive RNA. It has been suggested that the formation of ribonucleoprotein complexes could protect the capsid-less narnavirus RNA from cytoplasmic exoribonucleases and their RdRPs might act as a substitute for the viral CP (Solorzano *et al*., 2000). Apart from RNA replication formation of a ribonucleoprotein complex could also be essential for the virus to escape the host immune response and contribute to its stability in host cells and AfuTmV-1 may act similarly. In further work it would be of interest to fractionate virus-infected mycelia to obtain nuclear, cytoplasmic and cell debris extracts to ascertain the location of AfuTmV-1 dsRNAs in the cells. Also ultrathin sections and electron microscopy of virus-free and virus-infected fungal mycelia could also be performed to observe the location of the virus inside the fungal cells.

Application of AfuTmV-1 and Its Future Prospects

The development of transfection system in this study demonstrated that PEG was suitable to use for transfection of AfuTmV-1 and that the procedure could be applicable for infecting with other viruses or introducing genetically-modified viral dsRNA sequences directly to *A. fumigatus* protoplasts. This would promote studies on fungus-virus interaction to investigate effects of viruses on fungal virulence, as well as exploiting viral dsRNA sequences to study fungal pathogenicity *via* RNA silencing mechanisms. Since the colony morphologies of the AfuTmV-1-free and AfuTmV-1-infected *A. fumigatus* Af293 strains are indistinguishable, an effective selection method should be developed to simplify screening of a large numbers of rescued colonies on regeneration plates such as efficient selective marker genes, immunological assays or *in situ* nucleic acid hybridization. Additionally limitations on host range extension caused by the *vic* system of AfuTmV-1 to closely related fungal species could also be overcome using transfection. An important finding on the association between fungal cell surface protein (CSP) type in 86 clinical isolates of *A. fumigatus* and dsRNA mycovirus infection revealed that virus infection of *A. fumigatus* is not restricted to a certain CSP type, signifying that mycoviruses may be capable of infecting all *A. fumigatus* isolates (Refos *et al*., 2013). As a consequence, dsRNA mycoviruses are currently regarded as a therapeutic tool for *A. fumigatus* infections because the virus is prone to infect a wide variety of fungal strains. The approach could be achieved by introducing a strongly deleterious/hypovirulent mycovirus into any *A. fumigates* strain, allowing the replication of the virus within an infectious fungal population. In the future prospect, it would be worthwhile to conduct an experiment on virus transmission between closely related fungal strains in order to extend the virus host range.

Although an ability to cause hypovirulence of the fungal host is one of the important requirements for mycovirus as a possible therapeutic tool and biological control agent, in the case of symptomless AfuTmV-1 it would be possible to genetically modify this virus into a killer-phenotype carrying killer toxins as described for yeast mycoviruses (Schmitt and Breinig, 2002, Magliani *et al*., 1997). The implication of exploiting mycoviruses as therapeutic tools and biological control agents particularly in medicine to control human fungal infections has become one of the striking fields in medical mycology research. Curing human microbial infections with viruses is known for instance the treatment of bacterial infection with lytic bacteriophage (Biswas *et al*., 2002; Bradbury, 2004). Likewise, the same strategy can be

applied to hypovirulent-associated *Aspergillus* mycoviruses in order to treat fungal infections. Transposable elements, whose genetic structures resemble viruses has been used for vaccine development against aspergillosis. Also the *S. cerevisiae Ty1* element was genetically modified to express *Aspergillus* peptides which act as stimulators of memorised immune responses (Marchenko *et al*., 2003). Again, a similar strategy can be applied to develop mycovirus recombinants to express foreign antigenic epitopes of fungal pathogens.

Vector construct derived from mycovirus has become another interesting and promising scenario which can be used as an alternative therapeutic agent against fungal infection and for functional genomic studies. In this regard another approach which is focused in this study is developing mycovirus-based silencing vectors using the AfuTmV-1 sequences as gene carriers. This could assist investigations on fungal pathogenicity and modulation of gene expression in fungal pathogens, including exploitation of the potential virus vector as a future biological control agent. The development of vector construction is further described in chapter 3, part IV. Simultaneously a silencing vector based on a short interspersed nuclear elements (*SINE*) was also developed to silence a gene in an *A. fumigatus* clinical isolate. Characterisation and identification of *A. fumigatus SINEs* is further described in chapter 3, part III.

Conclusion (Part II)

Recently, increasing number of novel mycoviruses have been described. The discoveries of new virus structure and genome organisation have enhanced knowledge on virus diversity and virus-host evolution. This study investigated molecular properties of a novel dsRNA mycovirus with a tetrapartite genome designated as Aspergillus fumigatus tetramycovirus 1 (AfuTmV-1) isolated from the human pathogenic fungus *A. fumigatus* Af293 clinical isolate. Complete nucleotide sequences of the four AfuTmV-1 dsRNAs were determined by assembling and combining a series of cDNA clones, including those generated by RNA linker-mediated (RLM)-RACE, which spanned the entire length of each dsRNA. The full length genomic cDNA sequences of AfuTmV-1 dsRNA1-4 were deposited in the GenBank database with accession numbers HG975302-HG975305.

Inspection of the coding potential of the AfuTmV-1 dsRNAs revealed that each dsRNA element is monocistronic with a single open reading frame (ORF) on the positive-strands of all four dsRNAs. The genome of AfuTmV-1 is enriched in GC content (average value 63.40%) which is considerably similar to the yeast narnaviruses (average value 60%), however higher than those of most characterised mycoviruses including dsRNA partitiviruses and totiviruses, and also most ssRNA plant viruses. The presence of high GC content of this virus is not well-understood but it might result from adaptability of the virus against high mutation rates in order to maintain its genome.

The 5’ termini of untranslated regions (UTRs) of AfuTmV-1 dsRNA1-4 were highly conserved where the extreme 5’ terminal nt showed sequence heterogeneity whereas the 3’ termini were also highly conserved which comprised 4 G residues at the extreme 3’ terminus of each dsRNA. In addition, 5’ and 3’ UTRs of all four AfuTmV-1 dsRNAs were predicted to be highly structured following analysis with the mfold programme which consisted of long stem-loops resembling panhandle structures, especially at the 3’-terminus. As only dsRNA1 encodes an RdRP and terminal regions of the four dsRNAs are conserved, it can be inferred that the four dsRNAs do not originate from four independent mycoviruses.

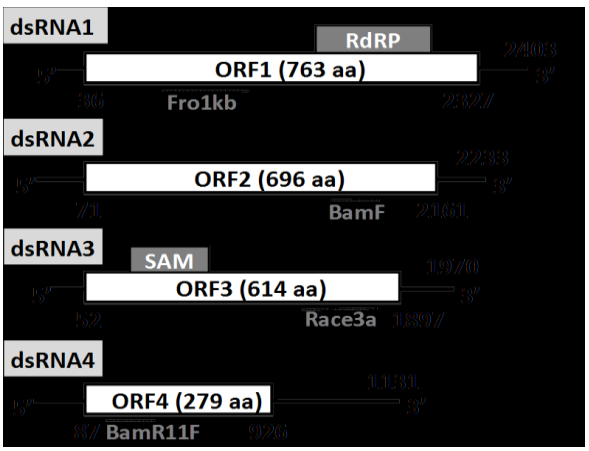
SDS-PAGE analysis and peptide mass fingerprinting (PMF) have revealed a major protein corresponding to protein translated from dsRNA4 which is proline-alanine rich (PArp). The protein encoded by dsRNA4 appeared to be associated with interaction to host membrane

components or serve as structural proteins or scaffold proteins anchoring replication complexes. The protein encoded by dsRNA1 appeared to associate with the synthesis of a membrane-bound RdRP whereas proteins encoded by dsRNA2 and dsRNA3 might be involved in dsRNA or protein binding activity during virus replication and 5'-RNA-capping methyltransferase activity, respectively. In addition, AfuTmV-1 was found to be unencapsidated as evidenced by BLASTX searches of the coat protein, transmission electron microscopy and SDS-PAGE analysis.

The phylogenetic analysis using RdRP sequences placed AfuTmV-1 in the same monophyletic group with newly discovered Cladosporium cladosporioides virus 1 (CCV1) and Beauveria bassiana tetramycovirus-1 (BbTmV-1) and they are phylogenetically distant from other dsRNA viruses reported thus far. The close relationship is supported by their dsRNA profile and conserved GDNQ motif on their RdRP sequences which also distinguishes these mycoviruses from other dsRNA mycoviruses. In addition, AfuTmV-1 RdRP appeared to be a derivative of the picorna-like superfamily and AfuTmV-1 dsRNA virus appeared to be phylogenetically a relative of the (+) ssRNA viruses in the family *Picornaviridae.* Characteristics of AfuTmV-1 dsRNAs and some properties of the representative dsRNA mycoviruses are summarized and compared in Table 16 and 17.

Table 16. Properties and genome organization of AfuTmV-1 dsRNAs.

Segment	Accession no.	ORF size (aa,kDa)	Putative function	UTR length (nt)	
				5'-UTR	3'-UTR
dsRNA1 (2403 bp)	HG975302	2292 (763, 84.2)	RdRP	35	76
dsRNA2 (2233 bp)	HG975303	2091 (696, 75.8)	Unknown	70	72
dsRNA3 (1970 bp)	HG975304	1845 (614, 66.7)	Methyltransferase	51	74
dsRNA4 (1131 bp)	HG975305	840 (279, 29.1)	PAS-rich protein (rp)	86	205



AfuTmV-1 is a member of a new lineage of quadripartite dsRNA mycoviruses

From The International Committee on Taxonomy of Viruses (ICTV) database, a recently established family *Quadriviridae* (e.g. Rosellinia necatrix quadrivirus1; RnQV1) (Lin *et al*., 2012) exhibits some similar features to AfuTmV-1. Viruses in the *Quadriviridae* family have four dsRNAs with sequence heterogeneity in the extreme terminal ends and ability to cause latent infections which are similar to the AfuTmV-1 (Table 17). However, some properties observed of this new virus family distinguish the quadriviruses from AfuTmV-1 such as its encapsidation in spherical particles, larger genome sizes, 72-82% sequence identity between the corresponding proteins and the GDD motif RdRP. Similar to the *Quadriviridae*, Alternaria alternata dsRNA virus 1 (AaV-1; Aoki *et al*., 2009) also contains a 4 dsRNA segmented genome which are encapsidated in isometric particles. Additionally, each dsRNA segment of this virus is polyadenylated and its dsRNA4 sequence has a similar size to AfuTmV-1 dsRNA4 (*ca*. 1.5 kb) but shorter than those of viruses from the family *Quadriviridae* and *Chrysoviridae*. Besides, the AaV-1 RdRP conserved motif is ADD instead of GDD. Based on these properties, AaV-1 is classified into another novel lineage of quadripartite viruses, the family *Alternaviridae*. Taken together the information on the AfuTmV-1 profile and other quadripartite viruses, the newly characterized AfuTmV-1 mycovirus could be classified as a member of a novel lineage of quadripartite dsRNA viruses. However the unencapsidated nature of AfuTmV-1, non-poly adenylated termini and the presence of GDN motif distinguish the virus from other quadripartite dsRNA viruses. Summarizing the information presented, AfuTmV-1 represents a new lineage of quadripartite dsRNA viruses and can be practically grouped in a new proposed virus family of tetramycovirus.

In addition to phylogenetic analyses, the finding of a homolog for AfuTmV-1 dsRNA2 in *Alternaria* sp. FA0703 (AltR1; accession number FJ595830) and our recent discovery of four dsRNAs in several isolates of *Beauveria bassiana* and in *Cladosporium cladosporioides* (CCV1, accession numbers YP_009052470-009052474) which their infecting viruses are similar in profile and size but not sequence to their AfuTmV-1 counterparts, suggest that tetraviruses having 4 dsRNA genome segments are prevalent in *Ascomycota* and are also present in the *Basidiomycota* such as *Melampsora lini* (Table 17).

In conclusion and in emphasis of the increasing diversity of mycovirus-related dsRNA viruses, the discovery in this study of a segmented dsRNA virus with a genome derived in part from unsegmented viral ancestors with reference to a capping enzyme and RdRP activity is unique as is the fact that the dsRNA is infectious. How AfuTmV-1 replicates is unknown but it is

conceivable that the AfuTmV-1 dsRNA4 gene product (PArp) may be intimately involved in perhaps assisting a host RNA helicase to partially unwind the dsRNA and this allows some translation. The replication cycle of AfuTmV-1 appears to be intermediate between dsRNA and positive-strand RNA viruses, as well as between *bona fide* viruses and capsid-less RNA elements. The AfuTmV-1 appears to exhibit a minor effect on the virulence level and phenotype of the fungal host. Thus, the virus exhibits latent/cryptic infection. From the novelty of the AfuTmV-1, the virus is accordingly proposed the establishment of a new genus (Tetramycovirus) and family (*Tetraviridae*) and to nominate AfuTmV-1 as the prototype species.

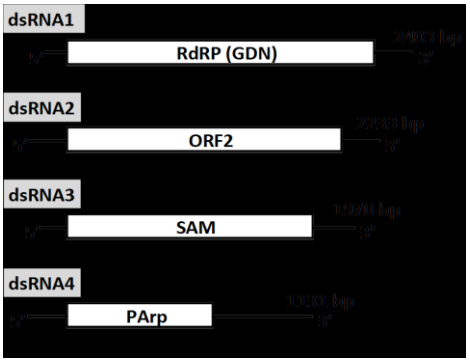
Table 17. Comparison of representative viruses in different ssRNA and dsRNA viral lineages.

Family	Segment No.	Genome Size (kb)	Poly(A)	Particles	Virus Sample	Abbrev.	RdRP motif	Host Taxa
*Totiviridae*****	1, dsRNA	4.6-7.0	N/A	Isometric (30-40 nm)	Aspergillus mycovirus 178 Botryotinia fuckeliana totivirus 1 Coniothyrium minitans RNA virus	AsV178 BFV CmRV	GDD	Ascomycota
Partitiviridae	2, dsRNA	1.4-3.0	Yes	Isometric (30-40 nm)	Aspergillus fumigatus partitivirus 1 Fusarium poae virus 1 Sclerotinia sclerotiorum partitivirus 1	AfuPV1 FUPO-1 SsV	GDD	Ascomycota
Chrysoviridae	4, dsRNA	2.4-3.6	No	Isometric (35-40 nm)	Aspergillus fumigatus chrysovirus Magnaporthe oryzae chrysovirus 1 Penicillium chrysogenum virus	AfuCV MoCV1 PcV	GDD	Ascomycota
Alternaviridae	4, dsRNA	1.5-3.6	Yes (33-50 nt)	Isometric (33 nm)	Alternaria alternata virus 1** Aspergillus foetidus virus-F	AaV-1 AfV-F	ADD	Ascomycota
Quadriviridae	4, dsRNA	3.7-4.9	No	Isometric (45 nm)	Rosellinia necatrix quadrivirus 1	RnQV1	GDD	Ascomycota
Tetramycovirus	**4, dsRNA**	**1.2-2.4**	**No (G stretch)**	**No**	**Aspergillus fumigatus tetramycovirus 1**	**AfuTmV-1**	**GDNQ**	**Ascomycota**
Unclassified virus	4, dsRNA (+1 DR)*	1.3-2.4 (0.9)	No (G stretch)	N/A	Cladosporium cladosporioides virus 1	CCV-1	GDNQ	Ascomycota

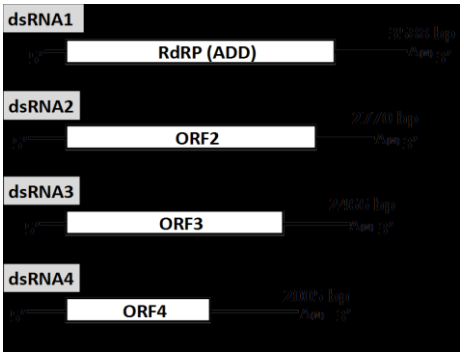
*Rhabdoviridae****	1, (-)ssRNA	11-15	Yes	Bacilliform	Rabies virus	RabV	GDN	human
Picornaviridae	1, (+)ssRNA	7.2-9.0	Yes	Icosahedral	Human hepatitis A virus	HAV	GDD	human
Narnaviridae	1, (+)ssRNA	2.73	N/A	No	Cryphonectria parasitica mitovirus 1-NB631	CHV-m	GDD	Ascomycota
Unclassifed hypovirus	1, (+)ssRNA	12.73	Yes (22 nt)	No	Cryphonectria parasitica hypovirus 1	CHV-1	GDN	Ascomycota

* Defective RNA ** Hypovirulent-associated virus ***5' capped RNA **** Not 5' capped RNA

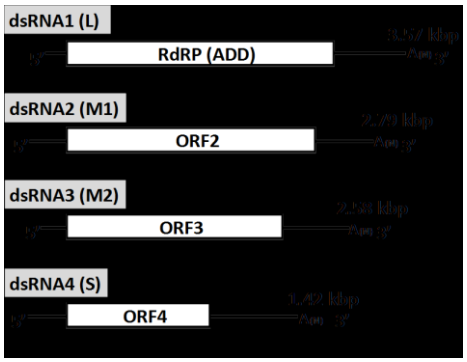
Comparison of genome organization of AfuTmV-1 and quadripatite dsRNA viruses.



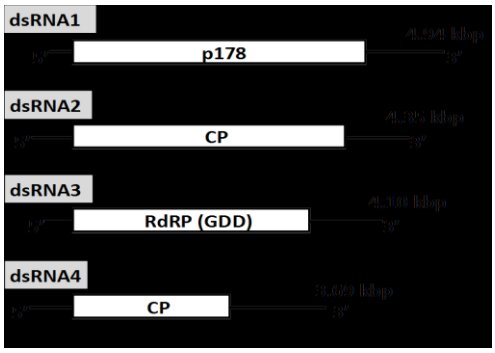
AfuTmV-1
(Aspergillus fumigatus tetramycovirus 1)



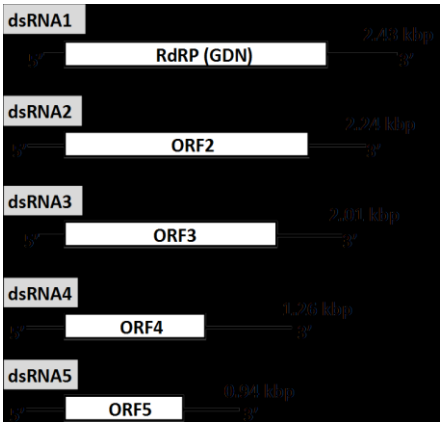
AfV-F
(Aspergillus foetidus Fast virus)



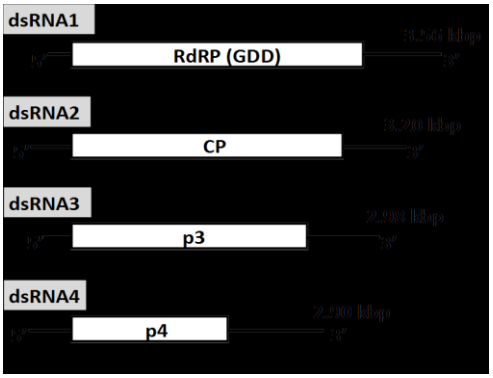
AaV-1
(Alternaria alternata virus 1)



RnQV1
(Rosellinia necatrix quadrivirus 1)



CCV-1
(Cladosporium cladosporioides virus 1)



PcV
(Penicillium chrysogenum virus)

Chapter 3

Results and Discussion (Part III)

(*SINE* Identification)

Short interspersed nuclear elements (*SINEs*) in the genome of the human pathogenic fungus, *Aspergillus fumigatus* Af293

Novel families of short interspersed nuclear elements (*SINEs*) in the human pathogenic fungus *Aspergillus fumigatus* clinical isolate Af293 were identified which were categorized into tRNA-related and 5S rRNA-related *SINE* families. Eight predicted tRNA-related *SINEs* nominated as *AfuSINE2* (for the *A. fumigatus SINE*) show target site duplication of short direct repeat sequences (4-14 bp) flanking the elements, extended tRNA-unrelated region, typical feature of RNA polymerase III promoter sequences and are found to originate from different tRNA origins. Each element ranges in size from 140-493 bp and represents low copy numbers in the genome based on southern blot analysis. Four out of 8 were identified as being actively transcribed by RT-PCR, indicating that the elements are functionally active in the genome. It can be found that a putative $tRNA^{Arg}$-derived *AfuSINE2-1a* exhibits a unique feature of terminal tetranucleotide repeats, ACT, at the 3' end and shows sequence similarity to the *I-4_AO* identified in *A. oryzae* and *I-1_AF LINE*-like element identified in *A. fumigatus* Af293. A family of 5S rRNA-related *SINEs* nominated as *AfuSINE3* are recognized and their 5'-5S rRNA-related regions show 50-65% and 60-75% similarity to the *A. fumigatus* 5S rRNAs and *SINE3-1_AO* present in *A. oryzae*, respectively. The genome contains approximately 5-6 copies of *AfuSINE3* ranging in size from 259-343 bp. The RT-PCR analysis revealed that only two out of five *AfuSINE3s* are actively transcribed. Investigation on *AfuSINE* distribution revealed that the elements are enriched in the pericentromeric and subtelomeric regions and inserted within gene-rich regions.

Short interspersed nuclear elements (*SINEs*) | *Aspergillus* | *SINE* distribution.

Chapter 3: Results and Discussion (Part III)
(*SINE* Identification)

5. Identification of *SINEs* in the *A. fumigatus* Af293 genome

5.1 Search strategy and identification of *SINEs* by computational analyses

The availability of the sequence of the complete genome of *A. fumigatus* Af293 provides an opportunity to investigate short interspersed nuclear elements (*SINEs*). In this study a computational search strategy was developed to predict the occurrence of *SINEs* in the *A. fumigatus* Af293 genomic sequences (designated as *AfuSINE*) using various bioinformatics software packages. The process framework used for computational analysis of *AfuSINEs* is shown in Fig. 51.

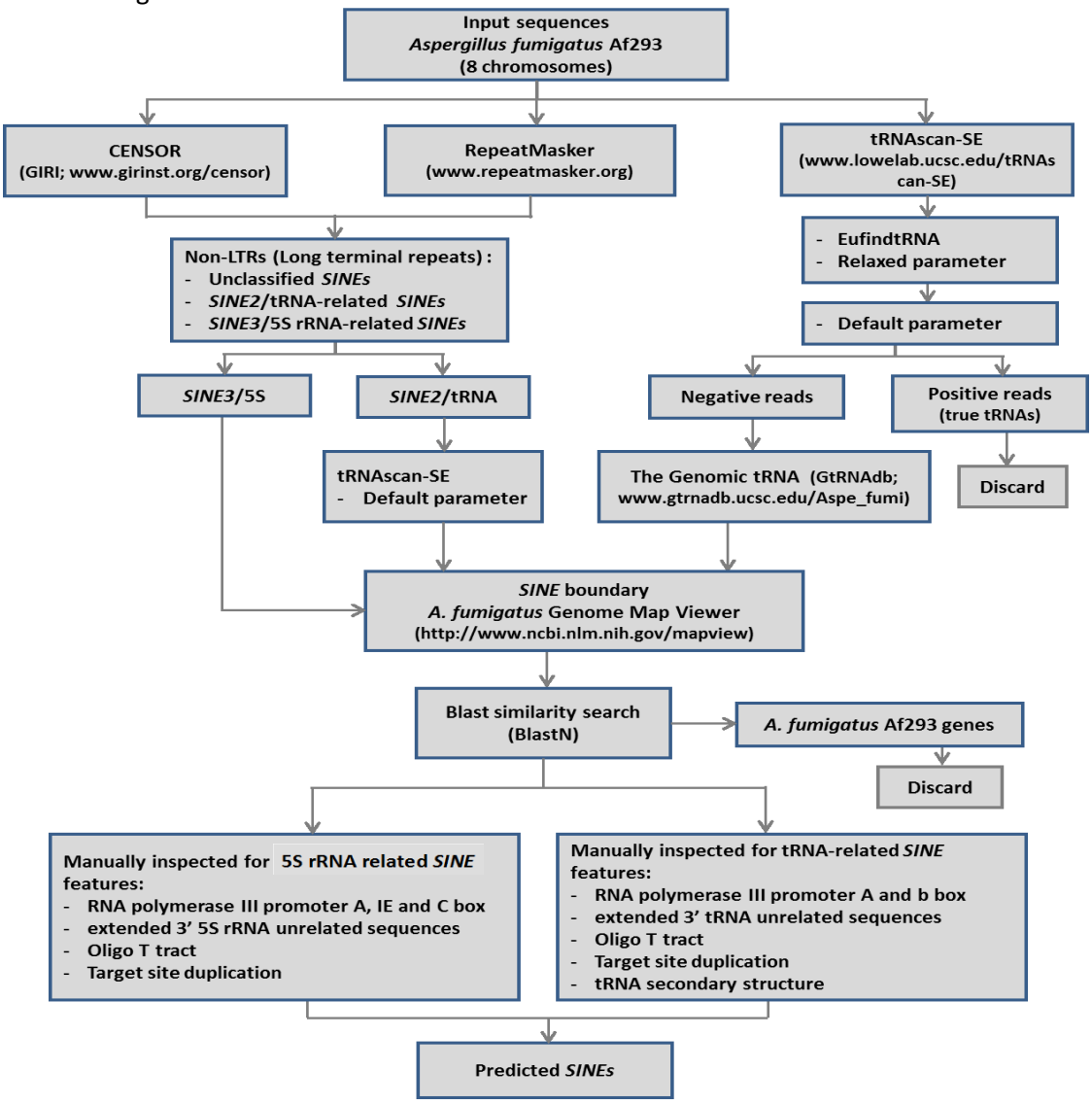


Fig. 51 Processing framework used for computational analysis of *AfuSINEs*.

The Af293 genomic sequences were initially screened for the interspersed repeat consensus sequences of *SINEs* using Censor and RepeatMasker against the available sequences on the Repbase Update including the sequences of other *LINEs*, *SINEs*, tRNAs, 5S rRNAs and 7SL RNA. The sequences classified as non-long terminal repeat (non-LTR); *SINE*, *SINE2*/tRNA-derived and *SINE3*/5S rRNA-derived *SINE*, were selected as candidate *SINEs*. In total, returned hits homologous to a reference collection of repeats in the Repbase database was retrieved, yielding *ca.* 17269 candidates using CENSOR and RepeatMasker. The sequences categorized as *SINE2* (tRNA-derived *SINE*) and *SINE3* (5S rRNA-derived *SINE*) were selected, yielding 145 hits for further investigation. It was found that all predicted *SINEs* showed similarity to the masked sequences of various tRNA-derived and 5S rRNA-derived *SINEs* from a broad spectrum of reference organisms including *A. oryzae*, *A. nidulans*, *Blumeria graminis*, *Cryptococcus neoformans*, arabidopsis, hedgehogs, mosquitos and humans.

The boundaries of predicted *SINEs* were retrieved by scanning the masked sequences with the NCBI BLAST and the *A. fumigatus* Genome Map Viewer servers to obtain 1000 bp upstream and 2000 bp downstream sequence. The expanded sequences classified as *SINE2*/tRNA from Censor and RepeatMasker were subsequently subjected to a scan with the tRNAScan-SE programme and the Genomic tRNA Database (GtRNAdb) using default settings to distinguish tRNA-related *SINEs* from tRNAs. Since the programme allows identification of the basic features of tRNAs such as the presence of a RNA pol III promoter, base separations for A and B box sequences (25-50 bp) and the presence of tRNA cloverleaf secondary structures, the predicted *SINE* showing positive reads were discarded as they were considered as true tRNAs. As a result 2 tRNA-related *SINEs* were predicted which showed similarity to *LFSINE-Vert* identified in *Latimeria menadoensis* (Bejerano *et al*., 2006) and *MIRc* identified in mammals (Smit and Riggs, 1995).

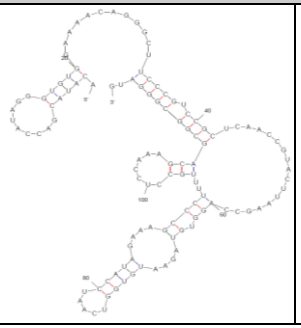
Sequences classified as *SINE3*/5S which were retrieved from the Censor and RepeatMasker programmes were aligned against the sequence of *SINE3-1_AO* from *A. oryzae*, the only fully characterised *SINE* element in Aspergilli, since most of the masked sequences corresponded to this *SINE* family. Using these programmes, 5S rRNAs of *A. fumigatus* Af293 were also characterized. It was found that the 5S rRNAs of *A. fumigatus* are closely related to the 5S rRNA from *A. nidulans*. Subsequently, the sequences of 5S rRNA-related *SINEs* were aligned against the *A. fumigatus* 5S rRNA. Of the 8 sequences, only 5 were found to contain an RNA polymerase III promoter (A, IE and C boxes) and possess 5S rRNA-related *SINE* features.

To intensively identify tRNA-related *SINEs*, the *A. fumigatus* Af293 genomic sequence was subjected to the tRNAScan-SE program to search for the presence of basic tRNA features using relaxed parameters set as follows; Search Mode= EufindtRNA only, Source= Eukaryotic, EufindtRNA search parameters= Relaxed. These settings were employed to search only for canonical promoter regions in the absence of structural tRNA features. Subsequently, the sequences predicted from the programme were subjected to the tRNAScan-SE using default settings. As a consequence, sequences yielding positive reads were discarded as they are considered as true tRNAs while the remaining sequences yielding negative reads, which showed no support for 5'tRNA-related region, were considered for further analyses. The tRNA cloverleaf secondary structures and tRNA origin of the remaining sequences were also identified from the programme. In addition, Genomic tRNA Database against the *A. fumigatus* tRNAs (GtRNAdb; http://gtrnadb.ucsc.edu/Aspe_fumi) was used to separate tRNAs from candidate *SINEs*. The boundaries of the predicted *SINEs* was also scanned as described above.

The remaining 25 sequences were manually inspected for the presence of degenerate RNA polymerase III promoter A and B boxes, nucleotide distance between identified A and B boxes, extended 3' tRNA-unrelated sequence before an oligothymidine tract and also target site duplication flanking the *SINEs*. However, it has been suggested that the RNA polymerase III promoter sequences can only be used as a rough indicator of *SINEs* but cannot be used as proof that the sequences belong to *SINEs*. Of the 25 sequences, only 6 were found to contain a full or partially intact RNA polymerase III promoter and possess tRNA-related *SINE* features. The candidate *SINEs* were also checked for the tRNA cloverleaf secondary structures and their origin by manually upload the sequence to tRNAScan-SE and mfold. Finally, BLASTN searches against the NCBI Genbank was performed to confirm that the predicted *SINEs* do not belong to other types of RNA Polymerase III transcripts, such as snRNAs, or other retrotransposons possessing a tRNA-like sequence as part of their organisation.

As a result, 13 candidate *SINEs* were recognized in the *A. fumigatus* clinical isolate Af293. The organization and structure of the representative *AfuSINEs* are shown in Table 18 and 19.

Table 18. List of 5 putative *A. fumigatus* Af293 5S rRNA-related *SINE* (*AfuSINE3)* sequences. Sequence alignment was performed using (CLUSTAL Omega).

Chr. #	*AfuSINEs*	Origin/ Similarity	2nd structure (mfold)	Length (bp)	Copy no.	Primers
5s-rRNA-related *SINEs* (*SINE3*)						
1	**>*AfuSINE3-1a#1_AO*** [3628551 to 3628830, plus strand] *TAGGCTC*CAACCAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAACCGTACTTAAGCCAGGTGTAGAATGTGGTCAATCCATAGAAAGCCCTTTTGCCTCCAAAGCAGCGGCGGGATGCTTTTCGTATGAGTGCTGGTGATGTGGGTGAGAACTCCAGGGGGCCCCATCACCCCTGGTGATGATGGCTCTAGGGGTTTGATGAAGGCTAACTTAACAGCTGACTGCAGAAGAAAACAGTGATTCAGCTGCTTTTAA*TAGGCTC*	5S rRNA/ *SINE3-1_AO*		**280** (Active)	**5**	**SINE3F** GACCATAGGGTGTGGAAAAC Tm=52 C GC=50 % **SINE3R1a** CCAGCACTCATACGAAAAGCA Tm=52 C GC=48 % Product size = 132 bp

Sequence alignment of *AfuSINE3-1a#1_AO* with *SINE3-1_AO* from *A. oryzae* = 75 % similarity

```
AfuSINE3-1a#1_AO      TAGGCTCCAACCAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAACCGTACTTAAGCCAGGTGTAGAATGTGGTCAATCCATAG
SINE3-1_AO            -------------AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGCGGCCAATTATCAA
                                   ** ********************** ********************** *************** * *  ** * ** ****    *

AfuSINE3-1a#1_AO      AAAGCCCTTTTGCCTCCAAAGCAGCGGCGGGATGCTTTTCGTATGAGTGCTGGTGATGTGGGTGAGAACTCCAGGGGGCCCCATCACCCCTGGTGATGATG
SINE3-1_AO            TGTGCCCTTAGCTCTNAAAGA-GCGRTGGCTTCGRTNTTAGGGTGAGTGCGGGTGATTTGGGTGAGAABTCTGGAGACCGGGGTGAGAGAGGTGAGAACTG
                         ******    **  **         *    * * ** *  ******* ****** ********** **  * *  **   *       *       **

AfuSINE3-1a#1_AO      GCTCTAGGGGTTTGATGAAGGCTAACTTAACAGCTGACTGCAGAAGAAAACAGTGATTCAGCTGCTTTTAATAGGCTCTGTGCCCTGAAGGTTAAAAA
SINE3-1_AO            GGTGAGAGTGGGTGAGAA-----------------------------------------------------------------------------------
                      * *    * *  ***  *
```

Sequence alignment of *AfuSINE3-1a#1_AO* with Afu5SrRNA1_1 = 64 % similarity

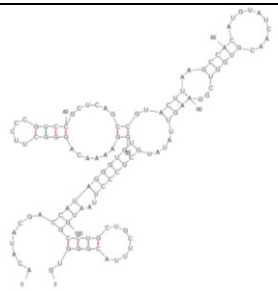
```
AfuSINE3-1a#1_AO      TAGGCTCCAACCAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAACCGTACTTAAGCCAGGTGTAGAATGTGGTCAATCCATAG
Afu5SrRNA1_1          ----------------ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGCCGGCCGGTTAGTAGTTGGGT
                                       ****************************************** **************   *  *   *      *

AfuSINE3-1a#1_AO      AAAGCCCTTTTGCCTCCAAAGCAGCGGCGGGATGCTTTTCGTATGAGTGCTGGTGATGTGGGTGAGAACTCCAGGGGGCCCCATCACCCCTGGTGATGATG
Afu5SrRNA1_1          G------GGTGACCACCAGCGAATCC--CGGCTGTTGTATGT-------------------------------------------------------------
                             *  ** ***  * * *    ** ** * *  **
```

3

>*AfuSINE3-3a#2_AO* [639030 to 639304, plus strand]

*AGCTT*G*AA*AACAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGT
CCGCTCAGCCGTACTTAAGCCACATGTATCAACGTGGTCGGAAGTTATATGCG
CCCTTAATGGCCTGCTGCTTTACGGGTGAGAGACGCTAAACCATAGGGTGAGT
GTGGGTGAGTGAAGGTGAGCTGAGTGATGCTATCCAACGCCATGGGTGAGCAA
GGGTGATCAGTGGGTGACAGTGTGATTCTATAATGACTAGACTCCGTCTATCG
CT*AGCTT*A*AA*

5S rRNA/ *SINE3-1_AO*



275

NA

SINE3F
GACCATAGGGTGTGGAAAAC
Tm=52 C GC=50 %
SINE3R3a
CCCTATGGTTTAGCGTCTCTCA
Tm=55 C GC=50 %

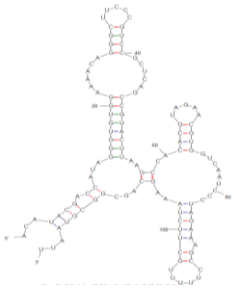
Product size = 133 bp

Sequence alignment of *AfuSINE3-3a#2_AO* with *SINE3-1_AO* from *A. oryzae* =73 % similarity

```
AfuSINE3-3a#2_AO      AGCTTGAAAACAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACATGTATCAACGTGGTCGGAAGTTATATGCGCC
SINE3-1_AO            ------------AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGCGGCCAATTATCAATGTGCCCT
                                  ** ********************** ************************************  * *   *    *      *   ** *** *

AfuSINE3-3a#2_AO      CTTAATGGCCTGCTGCTTTACGGGTGAGAGACGCTAAACCATAGGGTGAGTGTGGGTGAGTGAAGGTGAGCTGAGTGATGCTATCCAACGCCATGGGTGAGCAAGGGT
SINE3-1_AO            TAGCTCTNAA------------AGAGCGRTGGCTTCGRTNTTAGGGTGAGTGCGGGTGATT----------TGGGTGAGAABTCTGGAGACCGGG-G-TGAGAGAGGT
                                           * *      *      *********** ****** *          ** ****         *  **  * *     *  ***

AfuSINE3-3a#2_AO      GATCAGTGGGTGACAGTGTGATTCTATAATGACTAGACTCCGTCTATCGCTAGCTTAAA
SINE3-1_AO            GAGAACTGGGTGAGAGTGGGTGAGAA---------------------------------
                      **  * ******* **** *     *
```

Sequence alignment of *AfuSINE3-3a#2_AO* with Afu5SrRNA1_1 = 65 % similarity

```
AfuSINE3-3a#2_AO      AGCTTGAAAACAAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACATGTATCAACGTGGTCGGAAGTTATATGCGCC
Afu5SrRNA1_1          ---------------ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCACCGTACTTAAGCCACACGCCGGCCGGTTAGTAGTTGGGTGGGTGGACC
                                    ************************************************************ *       **     *  *         **

AfuSINE3-3a#2_AO      CTTAATGG---CCTGCTGCTTTACGGGTGAGAGACGCTAAACCATAGGGTGAGTGTGGGTGAGTGAAGGTGAGCTGAGTGATGCTATCCAACGCCATGGGTGAGCAAG
Afu5SrRNA1_1          ACCAGCGAATCCCGGCTGTTGTATGT-------------------------------------------------------------------------------------
                        *  *     ** **** * ** *

AfuSINE3-3a#2_AO      GGTGATCAGTGGGTGACAGTGTGATTCTATAATGACTAGACTCCGTCTATCGCTAGCTTAAA
Afu5SrRNA1_1          ---------------------------------------------------------
```

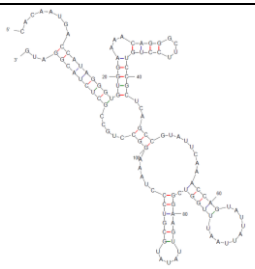
3	**>*AfuSINE3-3c#1_AO*** [1089477 to 1089819, minus strand] *ATTCAGT*TATCTTTCTCCCAGCACATACGACCATAGGGTGTGGAAAACAGGGC TTCCCGTCCGCTCAGCCGTACTTAAGCCACACGTAGAACGTGGTCAATCCCTA GAAAGCCTTTTTGCTTCTAAAGCAGCGGCGTAATTCTTTTAGTATGAGTGCTG GTGATACGGGTGAGGACCCCGTCACCCCTGGTGATGATGGCTTACTGGTGGTG ACGAGGGCTCCAGGGGGGTGACGAAGGCGTACTCAGAAGTCAACCACAGAAGA TACAGTGATTGAGCTACTTTTAATGGGCTCTGCGCCCCGCAGAAGTAATGCAC TTGATTATTCCTGAGTCC*ATA*C*AGT*	5S rRNA/ *SINE3-1_AO*		**343** (Active)	**5**	**SINE3F** GACCATAGGGTGTGGAAAAC Tm=52 C GC=50 % **SINE3R3c** CCCGTATCACCAGCACTCAT Tm=54 C GC=55 % Product size = 141 bp

Sequence alignment of *AfuSINE3-3c#1_AO* with *SINE3-1_AO* from *A. oryzae* = 74 % similarity

```
AfuSINE3-3c#1_AO      ATTCAGTTATCTTTCTCCCAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGTAGAACGTGGTCAATCC
SINE3-1_AO            -------------------AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGCGGCCAATTA
                                        ** ********************** ************************************    *  **** ** ****

AfuSINE3-3c#1_AO      CTAGAAAGCCTTTTTGCTTCTAAAGCAGCGGCGTAATTCTTTTAGTATGAGTGCTGGTGATACGGGTGAGGACCCCGTCACCCCTGGTGATGATGGCTTACTG
SINE3-1_AO            TCAATGTGCCCTTAGCTCTNAAAG-AGCGRTGGCTTCGRTNTTAGGGTGAGTGCGGGTGATTTGGGTGAGAABTCTGGAGACCGGGGTGAGAGAGGTGAGAAC
                        *    *** **     *  **         *      * ****  ******* ******  ******* *  * *    **  *****    **

AfuSINE3-3c#1_AO      GTGGTGACGAGGGCTCCAGGGGGGTGACGAAGGCGTACTCAGAAGTCAACCACAGAAGATACAGTGATTGAGCTACTTTTAATGGGCTCTGCGCCCCGCAGAA
SINE3-1_AO            TGGGTGAGAGTGGGTG----------AGAA----------------------------------------------------------------------------
                          *****    ** *           *  *

AfuSINE3-3c#1_AO      GTAATGCACTTGATTATTCCTGAGTCCATACAGT
SINE3-1_AO            ---------------------------------
```

Sequence alignment of *AfuSINE3-3c#1_AO* with Afu5SrRNA1_1= 60 % similarity

```
AfuSINE3-3c#1_AO      ATTCAGTTATCTTTCTCCCAGCACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGTAGAACGTGGTCAATCCC
Afu5SrRNA1_1          ----------------------ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGCCGGCCGGTTAGTAGTTG
                                             ************************************************************  *  **      *

AfuSINE3-3c#1_AO      TAGAAAGCCTTTTTGCTTCTAAAGCAGCGGCGTAATTCTTTTAGTATGAGTGCTGGTGATACGGGTGAGGACCCCGTCACCCCTGGTGATGATGGCTTACTGGT
Afu5SrRNA1_1          GGTGGGTGA------CCACCAGCGAATCCCGGCTGTTGT------ATGT-----------------------------------------------------------
                                   *  * *  * * *   *   ** *      ***

AfuSINE3-3c#1_AO      GGTGACGAGGGCTCCAGGGGGGTGACGAAGGCGTACTCAGAAGTCAACCACAGAAGATACAGTGATTGAGCTACTTTTAATGGGCTCTGCGCCCCGCAGAAGTA
Afu5SrRNA1_1          ------------------------------------------------------------------------------------------------------------
```

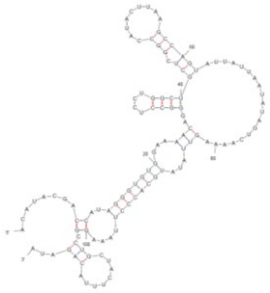
4	***>AfuSINE3-4a#1_AO*** [735462 to 735720, plus strand] *CGTCA*GATTGCACAATGACCATAGGGTGTGGAAAACAGGGCTTCCTGTCCGCT CAGCCGTATTCAAACCAGTATTATTAATTTGGTCGGAAGTTATATGCGTCCCT AAAGGCCTGCCGCTCTACGGATGAGAGACACTGAACCATAGGGTGAGTGTGGT GAGTGAAGGTGAGCAGAACGATGCTATCTATTGAATTATTCACATCGGCGATC CGTAGAACAGGTTGGTCCCACGTCCTGGACACGCCCATACTC*CGTCA*	5S rRNA/ *SINE3-1_AO*		**259**	**NA**	**SINE3F** GACCATAGGGTGTGGAAAAC Tm=52 C GC=50 % **SINE3R4a** CCCTATGGTTCAGTGTCTCTCA Tm=55 C GC=50 % Product size = 133 bp

Sequence alignment of *AfuSINE3-4a#1_AO* with *SINE3-1_AO* from *A. oryzae* = 65 % similarity

```
AfuSINE3-4a#1_AO      AGGACGTCAGATTGCACAATGACCATAGGGTGTGGAAAACAGGGCTTCCTGTCCGCTCAGCCGTATTCAAACCAGTATTATTAATTTGGTCGGAAGTTATATGC
SINE3-1_AO            -----------AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGCGGCCAATTATCAATGTG-
                                 * ***************** ************ *************** * ** ***    *   *    *      *   ** **

AfuSINE3-4a#1_AO      GTCCCTAAAGGCCTGCCGCTCTACGGATGAGAGACACTGAACCATAGGGTGAGTGTGGTGAGTGAAGGTGAGCAGAACGATGCTATCTATTGAATTATTCACAT
SINE3-1_AO            -CCCTTA----------GCTCTNAAAGAGCGRTGGCTTCGRTNTTAGGGTGAGTGCGGGTGATTTGGGTGAGAABTCTGGAGACCGGGGTGAGAGAGG-----T
                        ** **          *****      * *      *      *********** **    *   ****** *    *  *       *   *         *

AfuSINE3-4a#1_AO      CGGCGATCCGTAGAACAGGTTGGTCCCACGTCCTGGACACGCCCATACTCCGTCA
SINE3-1_AO            GAGAACTGGGTG---------------------------------------------
                        *   *  **
```

Sequence alignment of *AfuSINE3-4a#1_AO* with Afu5SrRNA1_1 = 50 % similarity

```
AfuSINE3-4a#1_AO      AGGACGTCAGATTGCACAATGACCATAGGGTGTGGAAAACAGGGCTTCCTGTCCGCTCAGCCGTATTCAAACCAGTATTATTAATTTGGTCGGAAGTTATATGC
Afu5SrRNA1_1          --------------ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGCCGGCCGGTTAGTAG---TTGGGTGG
                                   * **************************** *************** * ** ***           *     *   **   **

AfuSINE3-4a#1_AO      GTCCCTAAAGGCC--TGCCGCTCTACGGATGAGAGACACTGAACCATAGGGTGAGTGTGGTGAGTGAAGGTGAGCAGAACGATGCTATCTATTGAATTATTCAC
Afu5SrRNA1_1          GTGACCACCAGCGAATCCCGGCTGTTGTATGT-------------------------------------------------------------------------
                      **  * *   **   ****       * ***

AfuSINE3-4a#1_AO      ATCGGCGATCCGTAGAACAGGTTGGTCCCACGTCCTGGACACGCCCATACTCCGTCA
Afu5SrRNA1_1          ---------------------------------------------------------
```

5	***>AfuSINE3-5c#1_AO*** [251535 to 251831, plus strand] *AGCTTC*T*ATAT*GAATTTCTTGCACGAATATCCTAGCACATACGACCATAGGGT GTGGAAAACAGGGCCTCCTGGCTGCTCGGCCATACTTAAGCCAGTATTATTAA TATAGTCAAAAGTTATATGCACCCTTAAAGGCCTGCTACTTTACAGATAAGAG ATACTAAACCATAGGGTGAGTATAGGTGAGTAAAGGTGAGCTAAGTAATGCTA TCTAACAGCATACTTGAGCATGGTGATTAGTAGATAACAGAGTGATTCTATCA TGACTAGACTCCTCTATCACT*AGCTTC*A*ATAT*	5S rRNA/ *SINE3-1_AO*		**297**	**NA**	**SINE3F** GACCATAGGGTGTGGAAAAC Tm=52 C GC=50 % **SINE3R5c** ATCTGTAAAGTAGCAGGCCTTT Tm=51 C GC=41 % Product size = 112 bp

Sequence alignment of *AfuSINE3-5c#1_AO* with *SINE3-1_AO* from *A. oryzae* = 60 % similarity

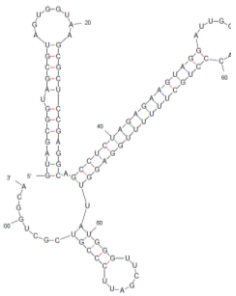
```
AfuSINE3-5c#1_AO      AGCTTCTATATGAATTTCTTGCACGAATATCCTAGCACATACGACCATAGGGTGTGGAAAACAGGGCCTCCTGGCTGCTCGGCCATACTTAAGCCAGTATTATTAATA
SINE3-1_AO            ---------------------------------AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGC
                                                       ** ********************** ******** *** * * **** *** ***********    *   *

AfuSINE3-5c#1_AO      TAGTCAAAAGTTATATGCACCCTTAAAGGCCTGCTACTTTACAGATAAGAGATACTAAACCATAGGGTGAGTATAGGTGAGTAAAGGTGAGCTAAGTAATGCTATCTAA
SINE3-1_AO            GGCCAATTATCAATGTG--CCCTTAGCTCTNAAAGA----------GCGRTGGCTTCGRTNTTAGGGTGAGTGCGGGTGATTTG----------GGTGAGA-ABTCTGG
                           *  *   ** **  ******          *           *      *      **********   ***** *            ** *     ***

AfuSINE3-5c#1_AO      CAGCATACTTGAGCATGGTGATTAGTAGATAACAGAGTGATTCTATCATGACTAGACTCCTCTATCACTAGCTTCAATAT
SINE3-1_AO            AGACCGGGGTGAGAGAGGTGAGAACTGGGTGAGAGTGGGTGAGAA----------------------------------
                         *     ****   *****  * * * * * ** * *     *
```

Sequence alignment of *AfuSINE3-5c#1_AO* with Afu5SrRNA1_1 = 60 % similarity

```
AfuSINE3-5c#1_AO      AGCTTCTATATGAATTTCTTGCACGAATATCCTAGCACATACGACCATAGGGTGTGGAAAACAGGGCCTCCTGGCTGCTCGGCCATACTTAAGCCAGTATTATTAATAT
Afu5SrRNA1_1          ----------------------------------ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGCCGGCCGGT
                                                         ****************************** *** * * **** *** ***********            *

AfuSINE3-5c#1_AO      AGTCAAAAGTTATATGCACCC---TTAAAGGCCTGCTACTTTACAGATAAGAGATACTAAACCATAGGGTGAGTATAGGTGAGTAAAGGTGAGCTAAGTAATGCTATCT
Afu5SrRNA1_1          TAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-------------------------------------------------------------
                          *   *        ***       **  ** ***  * **

AfuSINE3-5c#1_AO      AACAGCATACTTGAGCATGGTGATTAGTAGATAACAGAGTGATTCTATCATGACTAGACTCCTCTATCACTAGCTTCAATAT
Afu5SrRNA1_1          ---------------------------------------------------------------------------------
```

Note: NNNN = 5S rRNA-related regions *NNNN* = TSD, target site duplication NNNN = oligo T tract / poly A NNNN = oligonucleotide primer sequences

Table 19. List of **8** putative *A. fumigatus* Af293 tRNA-related *SINE* (*AfuSINE2).* Sequence alignment was performed using (CLUSTAL Omega).

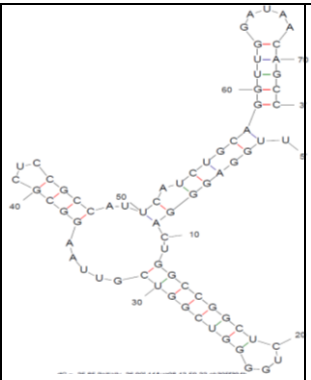
Chr. #	SINEs	Origin / Activity	2nd structure (mfold)	Length (bp)	Copy no.	Primers
	tRNA-related *SINEs* (*SINE2*)					
1	**>*AfuSINE2-1a*#arg** [336387 to 336756, plus strand] *AAGTGTACATAGAG*CCGGTAGCGGTAGCGTAGTGGTAAGCGCTCCGAGGCAGCCTCTAGAGAAGTAGGATTGGTGACCCTGCTTTTTTGGAGGTTATGGGTTCGATTCCCGTCGCTGGCACAACATTTACCACCACAATGGAAGATCACTTCCCACAATGGTATCAAGGCCACTCCCTTATCGCAAGGTGGTGGGGAAGTTGGAACAATCACAGGCCTGTAAGGCGAGGCTCTAAATTCGCCCTCATGTAATGGAACAAAATGTAACTAGACACACAAGGATTAGCTATAGTCGATACCTGCATATCGCCCAAGGCGAGGGGTCAGCGTATGAGTA*CTACTACTACTACTACTACT**AAGTGTACATAGAG*** **Note:** Sequence highlighted in grey is similar to *I-4_AO# LINE*/Tad1 of *A. oryzae* (Galagan *et. al.*, 2005; Kapitonov and Jurka, 2006) and *I-1_AF A. fumigatus* Af293 *LINE*.	**Arg**		**370** (Active)	**4**	**SINE2F1a** TAGTGGTAAGCGCTCCGAG Tm=59 C GC=58 % **SINE2R1a** GGTAAATGTTGTGCCAGCGA Tm=59 C GC=50 % Product size = 102 bp

Sequence alignment of *AfuSINE2-1a* with tRNA#Arg

```
AfuSINE2-1a#arg   GTAGCGGTAGCGTAGTGGTAAGCGCTCCGAGGCAGCCTCTAGAGAAGTAGGATTGGTGACCC-
tRNA2-Arg         GCCCTGCTGGCGCAATGGTAGCGCGTCAGACTTCTAATCTGAAGGTTGTGGGTTCGACCCCCA
                  *.  .* *.***.*.*****.    ** **  .    ***..**.    **.** *   ***
```

Sequence alignment of the 3'-termini of *AfuSINE2-1a* with *A. fumigatus* Af293 *I-1_AF LINE-like element* (3' –UTR of *LINE*-like element)

```
3'AfuSINE2-1a      ---CACAGGCCTGTA-AGGCGAGGCTCTAAATTCGCCCTCATGTAATGGAACAAAATGTAACTAG----------ACACACAAGGATTAGCTATAGTCGATACCTGCA
I_AF LINE          CGGCCCCTGCCTGAGCACGCCTGAGACAATACAATCCTTC-TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAGGAGGGGGGAAAGCAGAGACGCGGA
                      * *  *****   * **  *   * * *    ** ** * ***   ** ****   **  *          * * *  ****   *  * **  ** **  * *

3'AfuSINE2-1a      TATCGCCCAAGGCGAGGGGTCAGCGTATGAGT
I_AF LINE          GAATTGGC--GGATCATAGTGCGGAGATGGAT
                    *     *  **      **  *   ***  *
```

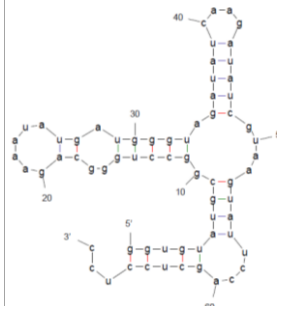
3	**>*AfuSINE2-3a#LFSINE*** [3928834 to 3929326, minus strand] *AGGGAGT*TTGATATCCCATTTTGGAGGGACTGGCCGGCTCTGGGGTCGGTCGTTAA GGCGCTCCGCCATTCATCTGCAGGTTGGATAACAGCCGGGCAATTTTCGTGACTGA AAATCCCATCCATCTGAAACTCTTGTGAGTATGTTGAAAGCCGACTTTTCAAGACG CCAATCTCAGGGCCATCCGCAGCTGAAAGATAACTCTCGAAGCAATGCCGGCACTG TAGAAAGCAGCAACGGAGAAGAGGCGATTGTTGGTGTCCGGCCAACTTGACATAAC GATCGAATAGATCGAACATTCGTTGCGAGGTTTCGACGGCCAGATCAATCATATCG GCACTGGGAGACCGACAGTGGCGGACAAGTGCGAGAATCACAAAGACGGCTGTTAG ATGCAAGGTAAGAAGAGGCGGCGTAGTAGAGTCGGACCTGGGATTTGAGTATTGTG CTAGAAGAATCGCCTGTTGTTCCTGGATCTGGAGCAGG*AGGGAGT* **Note:** The 5' tRNA-related region is similar to a 481 bp *LFSINE_Vert* (*SINE2*/tRNA#His) identified in *Latimeria menadoensis* (Bejerano *et al.*, 2006).	**Ser**		**493** (Active)	**1**	**SINE2F3a** CTCTGGGGTCGGTCGTTAAG Tm=60 C GC=60 % **SINE2R3a** GTCGGCTTTCAACATACTCACA Tm=59 C GC=46 % Product size = 120 bp

Sequence alignment of *AfuSINE2-3a* with tRNA#Ser

```
AfuSINE2-3a#LFS   TTTGGAGGGACTGGCCGGCTCTGGGGTCGGTCGTTAAGGCGCTCCGCCATTCATCTGCAGGTTGGATAACAGCCGGGCA
tRNA30-Ser        GTCAGTGTGCCCGAGTGGTTAAGGGGAGTGATTAGAAATCACTTGGCTTCGGCCGCATAGGTTCGAGTCCTATCGCTGA
                   *..* * * *.*. .**.*  ****   * .   **. *.**. **. .   . ...***** **   * ..**   *
```

Sequence alignment of *AfuSINE2-3a* with *LFSINE_Vert*

```
AfuSINE2-3a#LFS   TTGGAGGGACTGGCCGGCTCTGGGGTCGGTCGTTAAGGCGCTCCGCCATTCATCTGCAGGTTGGATAACAGCCGGGCAATTTTCGTGACTGAAAATCCCATCCATCTGAA
LFSINE_Vert       CTGGATGGCTCAGTGGAATTGGTAATGGGATATGGAGCCTTTCACCTCTAGGTCACTGGGTTCAAATCCAGCCCAG-GTCAGTAGTGACCGAAAGTCATTACCATCTGAT
                  .**** ** ...*. *. *. * ..* ** ..* .** * .**  *. *  .**  ..**** .*   ***** .* . .  * *****.****.** .  ********
```

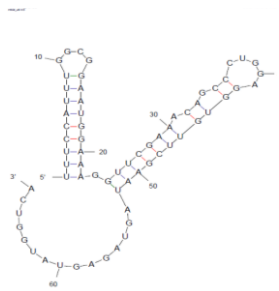
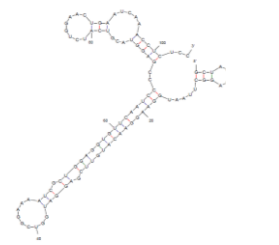
4	**>*AfuSINE2-4a#MIRc*** [3406775 to 3407123, plus strand] *ATGTACAT*CAGGTGTATGCGGCCTGGGCAGAAATATGATGGGTAGATATCAAGATA TCGTAAGTATTCCAGCTCCTCCACTGATTGACGATCCTTTTGTTATCTTGAGCCAA GGAATTTGGTGCCTTTGGGCCTGAGGTTCCAGATCCTCCCCCACATCCGCGAACCT CATTTCGTGGTTAAGGACAACATTAAATAACCGTGAATTTACTCCGTAGTGTCTCG ACCCTTTGGGGGGTGATCTTGTAAAGATTAGGAAAGAATGAGAGAAATAAATCACA GATTTTTTTTTTTTTTTTTCGTAAGAAAGAAAAGGCAACTCCCTCGTCACAGCAG GATTT*ATGTACAT* **5' tRNA-related CCA end **Note:** The 5' tRNA-related region is similar to a 268 bp *MIRc* (*SINE2*/tRNA #Pro) identified in mammals (Smit and Riggs, 1995).	**Ser**		**349** (Active)	**1**	**SINE2F4a** CCTGGGCAGAAATATGATGGG Tm=58 C GC=52 % **SINE2R4a** CCCAAAGGCACCAAATTCCT Tm=59 C GC=50 % Product size = 110 bp

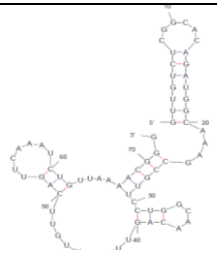
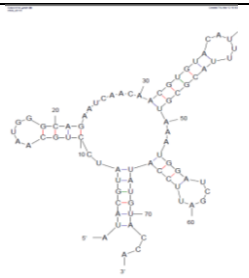
Sequence alignment of *AfuSINE2-4a* with tRNA#Ser

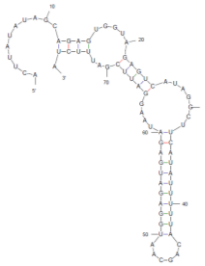
```
AfuSINE2-4a#MIR   A--GGTGTATGCGGCCTGGGCAGAAATATGATGGGTAGATATCAAGAT---ATCGTAAGTATTCCAGCTCCTCCACTGATTG
tRNA30-Ser        GTCAGTGTGCCCGAGTGGTTAAGGGGAGTGATTAGAAATCACTTGGCTTCGGCCGCATAGGTTCGAGTCCTATCGCTGAC-G
                  .  .****.. **. . *   **... .**** .* *. .*.. .* *   ..**.* . .*** **..*. .*.****. *
```

Sequence alignment of *AfuSINE2-4a* vs.*MIRc*

```
AfuSINE2-4a#MIR   GGT-GTATGCGGCCTGGGCAGAAATATGATGGGTAGATATCAAGATATC---GTAAGTATTCCAGCTCCTCCACTGATTGACGATCCTTTTGTTATCTTGAGCCAAGGAA
MIRc              GGCAGTGTGGTGCAGTGGAAAGAGCACTGGACTTGGAGTCAGGAAGACCTGGGTTCGAGTCCTGGCTCTGCCACTTACTAGC------TGTGTGACCTTGGG-CAAGTCA
                  **. **.**  **   ** *..*..*. . .  *.**  . ...* *.*   **  * .*.*..****. ***** *.*..*      * *** *.****.* ****  *

AfuSINE2-4a#MIR   TTTGGTGCCTTTGGGCCTGAGGTTCCAGATC
MIRc              CTTAACCTCTCTGAGCCTCAGTTTCCTCAT-
                  .**... .**.**.**** ** ****  **
```

7	**>*AfuSINE2-7a*#gly** [136570 to 136970, plus strand] *TCAAA*ACGCTGGAACTGGGTTTATTTCCATTTGGCGGAATGGAAAGGTTCGAAACA GCCCTGGAGGTGTTCGAATAGTAGAGTATGGTCAGCTTCGCTTCCTGCTCGTGTGA CTATGATGTAAAGATGTTGGGTTATTTTATTTCAATCAATCAGTGGGCCGCATTCT ATTCTTACAAGTCCCTTCAATGGCAATCATGGATATCAGATGGAGAGATTTGCGAT GCTTCTGTGTCTCTAATAATAAGCAGTCTTCAGATGAGTGGAGGCGGCAGAGAAGG GGAGCAATGGGGACCAATGGGGAGCAATGGGGAGTAATAAGGAGCAACAAGGATGA AGAGGATTGGTGTGACGACCCCTGCATACAACACTTGTAGGAAACGACATCTTATC AATC*TCAAA*	**Gly**		**401**	**NA**	**SINE2F7a** GCGGAATGGAAAGGTTCGAA Tm=59 C GC=50 % **SINE2R7a** GAATAGAATGCGGCCCACTG Tm=59 C GC=55 % Product size = 139 bp
	Sequence alignment of *AfuSINE2-7a* with tRNA#Gly AfuSINE2-7a#gly TTCCATTTGGCGGAATGGAAAGGTTCGAAACAGCCCTGGA-------GGTGTTCGAATAGTAGAGTATG tRNA23-Gly GCATCATTGGTCTAGTGGTAGAATTCGTCGTTGCCATCGACGAGGCCCGTGTTCGATTCACGGATGATG . . ****. *.*** *...**** .. *** * ** ******** * ...** ***					
5	**>*AfuSINE2-5d*#ser** [2378573 to 2378802, plus strand] *TTTTTAGT*TGCTACGTAGGCTTAATGGAAGGAACATGTTCGAGGATGGTCGGAAAA TCGCTGGAGGTGTTCAATCCCCGAGGTACGTCATCTGGAACTGAATCAAACCTCTC CACTCGAGGGAGAGCTGAGAGTATTAGGTGGTTTTCGTACTTATATCCAAGGCTCA TTATCACGAATGCCGTACAAACAAACACAATCGAGATACTCCATACTTGAATTA*TT* *TT*C*AGT* **5' tRNA-related CCA end	**Ser**		**230** (Active)	**2**	**SINE2F5d** GAGGATGGTCGGAAAATCGC Tm=59 C GC=55 % **SINE2R5d** ACCTAATACTCTCAGCTCTCCC Tm=58 C GC=50 % Product size = 101 bp
	Sequence alignment of *AfuSINE2-5d* with tRNA#Ser AfuSINE2-5d#ser GCTACGTAGGCTTAATGG--AAGGAACATGTTCGAGGATGGTCGGAAAATCGCTGGAGGTGTTCAATCCCCGAGGTACGTCATCTGGA tRNA30-Ser GTCAGTGTGCCCGAGTGGTTAAGGGGAGTGATTAGAAATCACTTGGCTTCGGCCGCATAGGTTCGAGTCCTA-------TCGCTGACG *..* * *. *.*** ****.. .** *.....** ... *. . **.* * . ****.* .**.. **... . .					

3	**>*AfuSINE2-3c*#ala** [3105817 to 3106017, minus strand] *TGTG*ATG*C**CTTTC**CA*CACTGCATGGTTGTCTCGGCACAGATGGCAAAGCCGTTCCT GGCAACAGTTAATTGTTCAGTTCAAATCTGTTAAAACGGGGTCGTACTTACCATAG ACTGTCGTCGAGCATCTGCCGCAGCTCTAGCCAAAAGACCCAGTGATTTCTGCCTA TGAATGACAGGTCAGCTT*TGTG*GGG*GCTTTGCA*	**Ala**		**201** (Active)	**5**	**SINE2F3c** GTTGTCTCGGCACAGATGG Tm=58 C GC=58 % **SINE2R3c** TCACTGGGTCTTTTGGCTAGA Tm=59 C GC=48 % Product size = 134 bp
	Sequence alignment of *AfuSINE2-3c* with tRNA#Ala AfuSINE2-3c#ala GTTGTCTCGGCACAGATGGCAAAGCCGTTCCTGGCAACAGTTAATTGTTCAGTTCAAATCTGTTAAAACGGG tRNA27-Ala GGGCTGTTAGCACAGTGGTAGTGCGCCCGCTTCGCATGCGGGAGGTCGTGGGTTCGATTCCCACATAGTCCA * * *..****** * . . * . *.* *** * *. * * .****.* **. .* *.. .					
4	**>*AfuSINE2-4c*#thr** [722885 to 723137, minus strand] *CTTGGAGA*TTACATACGTATCCTGCAATGGGCAGAATCAACAACGTGTACATTTTA CGCGTAAATGGATCGATTCCATATGTACCACTCTCTAAGCCAACGCAACCATGAAA GGGTGAGAAGACCCTGACGGCCACCGAAACACTTACCAACCAGTTCTTGGTTGTGA ACATAGTATGGATAATTTGGTAGCTATAGTTGATTTGAATAATGGTGCTGTCGCAC CTCGCAATGTCTTTTTAGCAT*CTTGGAGA* **5' tRNA-related CCA end	**Thr**		**253**	**NA**	**SINE2F4c** GGGCAGAATCAACAACGTGT Tm=59 C GC=50 % **SINE2R4c** TTGGTAAGTGTTTCGGTGGC Tm=59 C GC=50 % Product size = 123 bp
	Sequence alignment of *AfuSINE2-4c* with tRNA#Thr AfuSINE2-4c#thr ATACGTATCCTGCAATGGGCAGAATCAACAACGTGTACATTTTACGCGTAAATGGATCGATTCCATATGTACCAC tRNA11-Thr GCATCTATGGCTCAGT-GGTAGAGC--GTTCCACTCGTAATGGAAAGGTCCATGGTTCAATCCCGTGTAGATGCA ..*. *** . **.* **.***.. .. *.. ...* * * . ** **** **.**.**.*.*. *.					

7	**>*AfuSINE2-7e*#met** [35586 to 35725, plus strand] *AATT*ACTTATATAGCAGAGTGGTAGAGTCATAGGCTTCATATTTTTACAGCAATGG AGATGAGATAAGGATTCGATTTCTATATATTTTAGTACTAGATAATATTAACTTTA AGGGAAGCTATAGGATTTTCCCCT*AATT*	**Met**		**140**	**NA**	**SINE2F7e** ACTTATATAGCAGAGTGGTA Tm=45 C GC=35 % **SINE2R7e** TCCTATAGCTTCCCTTAAAG Tm=48 C GC=40 % Product size = 123 bp

Sequence alignment of *AfuSINE2-7e* with tRNA#Met

```
AfuSINE2-7e#met    ACTTATATAGCAGAGTGGTAGAGTCATAGGCTTCATATTTTTACAGCAATGGAGATGAGATAAGGATTCGATTTCTAT
tRNA31-Met         AGCATGTTAGCTCAGGGGAAGAG-CGCCGGGCTCATAACCCGGAGGTCCCTGGATCGAAACCAGGACATGCT-----A
                   * .    ****  ** ** **** *.. ** .***** ... . .*.  . *.. .**.*. ****. .* *
```

Note: NNNN = 5'tRNA-related regions NNNN = region similar to reference database NNNN = A and B boxes *NNNN* = TSD, target site duplication *NNNN* = short repetitive sequence NNNN = oligo T tract/poly A NNNN = oligonucleotide primer sequences

>Afu5SrRNA1_1 [119 bp]

ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGCCGGCCGGTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT

>SINE3-1_AO [206 bp] (Galagan *et al.*, 2005; Kapitonov and Jurka, 2006)

AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTATTAAACGCGGCCAATTATCAATGTGCCCTTAGCTCTNAAAGAGCGRTGGCTTCGRTNTTAGGGTGAGTG
CGGGTGATTTGGGTGAGAABTCTGGAGACCGGGGTGAGAGAGGTGAGAACTGGGTGAGAGTGGGTGAGAA

>LFSINE_Vert [481 bp]

TCGGGGGGACTGGATGGCTCAGTGGAATTGGTAATGGGATATGGAGCCTTTCACCTCTAGGTCACTGGGTTCAAATCCAGCCCAGGTCAGTAGTGACCGAAAGTCATTACCATCTGATGGCTGTTCAGTGGCCTAT
GTGAAATGAGTTGGTGGTCTCAGTCCAGTTCCTAGTGGACAGGTGTCCACATCACAAAACCACCATCACAATTGGCACTAATTGGCACCCTTGTTGGCAGTCTCAGCAGAGAGGCCAAGGATTGAATGGGCATGGA
GACTGAACTACCCTCTCAACCCTGTAGAGGTGGTCCCTCCAGGGCAGGGTTGAGGCACATTGGCAGGGCAATGTGGGGAAGCCTGCACTGCTGCTGCCCATGCTGTACCTGTTCTGTGGATAAATAGAGGACTTCA
GTCTCTGGTGCTATCAATCTAGCACCTTTCACGAGCACTAAATTCACACAAAAAAATTTAAAAAAAAAAAAAAAAAA

>MIRc [268 bp]

CGAGGCAGTGTGGTGCAGTGGAAAGAGCACTGGACTTGGAGTCAGGAAGACCTGGGTTCGAGTCCTGGCTCTGCCACTTACTAGCTGTGTGACCTTGGGCAAGTCACTTAACCTCTCTGAGCCTCAGTTTCCTCAT
CTGTAAAATGGGGATAATAATACCTGCCCTGCCTACCTCACAGGGTTGTTGTGAGGATCAAATGAGATAATGTATGTGAAAGCGCTTTGTAAACTGTAAAGTGCTATACAAATGTAAGGGGTTATTATTATT

Additionally, the CENSOR and RepeatMasker programmes revealed the sequences of a number of transposons (*Mariner* class II DNA transposon), LTR-retrotransposons (*Copia*, *Gypsy* and *Afut1*) and non-LTR retrotransposons (*I-1_AF LINEs*) in the *A. fumigatus* Af293 genome, including some 5S rRNA sequences. Several *I-1_AF LINEs* and 5S rRNAs sequences were predicted through the *in silico* directed searches against the genomic DNA yielding comparable results to the numbers of *Copia*, *I-1_AF LINEs* and 5S rRNAs described from southern blot and quantitative RT-PCR analyses representing copy numbers of 29 *Copia*, 24 *LINEs* and 34 5S rRNAs, respectively (Galagan *et al*., 2005; Kapitonov and Jurka, 2006; Huber and Bignell, 2014). A list and tentative numbers of retrieved sequences predicted by direct search strategy are shown in Table 20.

Table 20. Lists and tentative numbers of transposons found in the *A. fumigatus* Af293 genome predicted using direct search strategy.

Chro.#	Numbers of Hits						
	Total hits	***Mariner***	***Copia***	***Gypsy***	***Afut1***	***I-1_AF***	**5SrRNA**
1	3191	16	5	3	-	-	4
2	3351	15	5	8	-	-	4
3	2140	18	5	8	6	2	3
4	1331	8	1	5	7	6	4
5	2538	13	3	1	-	2	5
6	2499	18	4	10	11	9	7
7	1118	8	1	7	2	3	1
8	1102	4	3	2	2	3	5
total	17269	100	27	39	28	25	33

5.2 Sequence analysis and classification of *AfuSINEs*

5.2.1 5S rRNA-related *SINEs* (*AfuSINE3*)

In this study, 5S rRNA-related *SINE*-like elements (designated as *AfuSINE3*) which are derived from 5S rRNA were identified from the *A. fumigatus* Af293 genome based on the comparison of the sequences against the Repbase Update database. The elements are present in approximately five related-copies (named *AfuSINE3-1a, AfuSINE3-3a, AfuSINE3-3c, AfuSINE3-*

4a and *AfuSINE3-5c* from chromosomes 1, 3, 4 and 5) by *in silico* analysis, with an average length of 259-343 bp. They are found to be related to *SINE3-1_OA*, which is a 5S-derived *SINE* identified from *A. oryzae.*

Comparative analysis of the sequences was performed using Clustal Omega and MAFFT alignment. Phylogenetic analyses were performed using the neighbor-joining (NJ) method in MAFFT alignment. Pair-wise alignment of the *AfuSINE3* against 5S rRNAs revealed that the 5' termini of the sequences (positions 1 to 64) are well-preserved and are very similar to the sequences of *A. fumigatus* 5S rRNA. In addition, high similarity was also observed when the 5' terminal sequences of *AfuSINE3* candidates were aligned against *A. oryzae SINE3-1_AO* (Table 18). Excluding the 3'-termini, the 5S rRNA-related regions of *AfuSINE3* (position 1 to 119) showed 50-65% similarity to 5S rRNA genes and 60-75% similarity to *SINE3-1_OA*. With at least 60% similarity to the 5S rRNA species and in at least 60 nucleotide overlap, it confirms that all *AfuSINE3* sequences are derived from 5S rRNA.

Multiple sequence alignment of the 5S rRNA-related regions of *AfuSINE3* (position 1 to 119) showed that the 5' termini are highly conserved among the elements with 56-90 % similarity (Fig. 52). Phylogenetic analysis showed that *AfuSINE3-1a* and *AfuSINE3-3c* are grouped in the same clade as *SINE3-1_OA* (Fig. 53). Besides, significant similarity has been observed only at the initial sequences of 3'- termini among the five *AfuSINE3* candidates (Fig. 54). Sequence analysis of *AfuSINE3* showed the absence of apparent ORFs and poorly conserved type I RNA pol III internal promoter sequences comprising an A box an Intermediate element (IE) and a C box. Additionally they lacked the termination signal for RNA pol III transcription (GCTTTTCG), resulting in continuous transcription activity extending through the 5S rRNA-related region to the 3' terminus of *AfuSINE3*. However, this exception was found in *AfuSINE3-1a#1_AO* since the element contains a termination signal.

Two main considerations used to define an *AfuSINE* family in this study are 1) the same common origin and 2) the same sequence module/structure. However, the terminal sequences of some familial *SINEs* were variable thus these were omitted from the *AfuSINE* family tree constructions. From these criteria and the multiple sequence alignment of *AfuSINE3* candidates, the sequences were considered as copies of one element since the alignment showed similarities of the 5' terminal head regions of up to 90%. In addition, phylogenetic analysis of the 5' termini showed similar results, revealing all five predicted *AfuSINE3s* were grouped into two closely related, distinct lineages. Thus, it can be inferred that *AfuSINE3s*-derived from 5S rRNA genes originate from the same family. A multiple sequence alignment of

the *AfuSINE3* candidates against the *A. fumigatus* Af293 5S rRNAs and *SINE3-1_AO* from *A. oryzae* is shown in Fig. 55.

```
AfuSINE3-1a#1_AO ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAACCGTACTTAAGCCAGGTGTA-GAA
AfuSINE3-3c#1_AO ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGTA-GAA
AfuSINE3-3a#2_AO ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACATGTATCAA
AfuSINE3-4a#1_AO CACAATGACCATAGGGTGTGGAAAACAGGGCTTCCTGTCCGCTCAGCCGTATTCAAACCAGTATTATTAA
AfuSINE3-5c#1_AO ACATACGACCATAGGGTGTGGAAAACAGGGCCTCCTGGCTGCTCGGCCATACTTAAGCCAGTATTATTAA
                     *.************************.***.* *.****..**.**.*.**.***   **   **

AfuSINE3-1a#1_AO TGTGGTCAATCCATAGAAAGCCCTTT---TGCCTCCAAAGCAGCGGCGGGATG
AfuSINE3-3c#1_AO CGTGGTCAATCCCTAGAAAGCCTTTT---TGCTTCTAAAGCAGCGGCGTAATT
AfuSINE3-3a#2_AO CGTGGTCGGAAGTTATATGCGCCCTTAATGGCCTGCTGCTTTACGG----GTG
AfuSINE3-4a#1_AO TTTGGTCGGAAGTTATATGCGTCCCTAAAGGCCTGCCGCTCTACGG----ATG
AfuSINE3-5c#1_AO TATAGTCAAAAGTTATATGCACCCTTAAAGGCCTGCTACTTTACAG----ATA
                 . *.***..    ** * .  ....*    **.* . .  . .*.*    .*
```

Fig. 52 Multiple sequence alignment of the *AfuSINE3* sequences. Only the 5S rRNA-related regions (nt 1-119) of each sequence were selected for the alignment which was performed using the Clustal Omega programme available at the EMBL-EBI website.

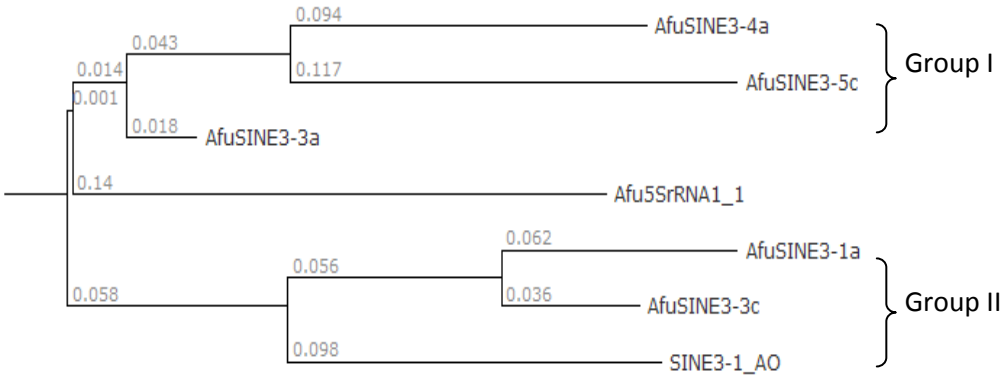


Fig. 53 Phylogram of the *AfuSINE3, A. fumigatus* Af293 5S rRNA sequences and *SINE3-1_AO* sequence from *A. oryzae*. Only the 5S rRNA-related regions (nt 1-119) of each sequence were selected for alignment which was performed using the MAFFT alignment programme.

```
AfuSINE3-1a#1_AO -----------CTTTTCGTATGAGTGCTGGTGA-TGTGGGTGAGAACTCCAGGGGGCCCCA---------
AfuSINE3-3c#1_AO -----------CTTTTAGTATGAGTGCTGGTGA-TACGGGTGAGGACCCCGTCACCCCTGGTGATGATGG
AfuSINE3-3a#2_AO AGAGACGCTAAACCATAGGGTGAGTGTGGGTGAGTGAAGGTGAGCTGAGTGATGCTATCCAACGCCATGG
AfuSINE3-5c#1_AO AGAGATACTAAACCATAGGGTGAGTATAGGTGAGTAAAGGTGAGCTAAGTAATGCTATCTAACAGCATAC
AfuSINE3-4a#1_AO AGAGACACTGAACCATAGGGTGAGTGT-GGTGAGTGAAGGTGAGCAGAACGATGCTATCTATTGAATTAT
                          .. * * .*****.. ***** *. .******     ..  .   .. .

AfuSINE3-1a#1_AO -TCACCCCTGGTGATGATGGCTCTAGGGGTTTGATGAAGGCTAACTTAACAGCTGACTGCAGAAGAAAAC
AfuSINE3-3c#1_AO CTTACTGGTGGTGACGAGGGCTCCAGGGGGGTGACGAAGGCGTACTCAGAAGTCAACCACAGAAG-ATAC
AfuSINE3-3a#2_AO GTGAGCAAGGGTGA------------------------------TCAGTGGGTGACAGTGTGATTCTAT
AfuSINE3-5c#1_AO TTGAGC-ATGGTGA------------------------------TTAGTAGATAACAGAGTGATTCTAT
AfuSINE3-4a#1_AO TCACAT--CGGCGA------------------------------TCCGTAG--AACAGGTTGGTCCCAC
                  .   .   **.**                                *. . .*  .** .   ..    *.

AfuSINE3-1a#1_AO AGTGATTCAGCTGCTTTTAA-------------------------------------------------
AfuSINE3-3c#1_AO AGTGATTGAGCTACTTTTAATGGGCTCTGCGCCCCGCAGAAGTAATGCACTTGATTATTCCTGAGTCC
AfuSINE3-3a#2_AO AATGACTAGACTCCGTCTATCG-----------------------------------------------CT
AfuSINE3-5c#1_AO CATGACTAGACTCC-TCTATCA-----------------------------------------------CT
AfuSINE3-4a#1_AO GTCCTGGACACGCC-CATACTC-------------------------------------------------
                  .      .*  * . **
```

Fig. 54 Multiple sequence alignment of *AfuSINE3* sequences. Only the 3S rRNA-unrelated regions of each sequence were selected for the alignment which was performed using the Clustal Omega programme available at the EMBL-EBI website.

```
Afu5SrRNA1_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA1_2       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA2_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA2_3       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA3_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA5_3       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA5_4       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA6_3       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA7_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA8_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA8_3       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA2_2       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA6_1       ---ACGTACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA8_2       ---ATATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA5_1       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA6_4       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA6_5       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCATACTTAAGCCACACGC-
Afu5SrRNA3_2       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA3_3       ---ACATACGACCATAGGGTGTGGAAAGCAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA5_2       ---ACATACGACCACAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA6_6       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGC-
Afu5SrRNA1_3       ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACGCGC-
Afu5SrRNA1_4       ---ACATACGACCACAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCATACTTAAGCTACACGC-
SINE3-1_AO         AGTACATACGACCATAGGGTGTGGAGAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCAACTAT-
AfuSINE3-1a#1_AO   ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAACCGTACTTAAGCCAGGTGT-
AfuSINE3-3c#1_AO   ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACACGT-
AfuSINE3-3a#2_AO   ---ACATACGACCATAGGGTGTGGAAAACAGGGCTTCCCGTCCGCTCAGCCGTACTTAAGCCACATGTA
AfuSINE3-4a#1_AO   --CACA-ATGACCATAGGGTGTGGAAAACAGGGCTTCCTGTCCGCTCAGCCGTATTCAAACCAGTATTA
AfuSINE3-5c#1_AO   ---ACATACGACCATAGGGTGTGGAAAACAGGGCCTCCTGGCTGCTCGGCCATACTTAAGCCAGTATTA
                      *.. *.*****.******.***.*.******.***.* *.****..**.**.*.**.*.*    .

Afu5SrRNA1_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA1_2       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA2_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA2_3       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA3_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA5_3       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA5_4       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA6_3       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA7_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA8_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA8_3       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA2_2       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
Afu5SrRNA6_1       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA8_2       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA5_1       CGGCTG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCGGCTGTTGTATGT-----------
Afu5SrRNA6_4       CGGCTG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCAGCTGTTGTATGT-----------
Afu5SrRNA6_5       CGGCTG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCAGCTGTTGTATGT-----------
Afu5SrRNA3_2       CGGCCA--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
Afu5SrRNA3_3       CGGCCA--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
Afu5SrRNA5_2       CGGCCA--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
Afu5SrRNA6_6       CGGCTA--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
Afu5SrRNA1_3       CGGCCG--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCCAGCTGTTGTATGT-----------
Afu5SrRNA1_4       CGGCCA--GTTAGTAGTTGGGTGGGTGACCACCAGCGAATCCTGGCTGTTGTATGT-----------
SINE3-1_AO         TAAACGCGGCCAATTATCA-ATGTGCCCTTAG--------------CTCTNAAAGAGCGRTGGCTT-
AfuSINE3-1a#1_AO   AGAATGTGGTCAATCCATA-GAAAGCCCT------------TTTGCCTCCAAAGCAGCGGCGGGATG
AfuSINE3-3c#1_AO   AGAACGTGGTCAATCCCTA-GAAAGCCTT------------TTTGCTTCTAAAGCAGCGGCGTAATT
AfuSINE3-3a#2_AO   TCAACGTGGTCGGAAGTTATATGCGCCCTTAATGG------CCTGCTGCTTTACGGGTG---------
AfuSINE3-4a#1_AO   TTAATTTGGTCGGAAGTTATATGCGTCCCTAAAGG------CCTGCCGCTCTACGGATG---------
AfuSINE3-5c#1_AO   TTAATATAGTCAAAAGTTATATGCACCCTTAAAGG------CCTGCTACTTTACAGATA---------
                    . .   *....    .. . . ..  .                . ..   *
```

Fig. 55 Multiple sequence alignment of the *AfuSINE3* sequences against the *A. fumigatus* Af293 5S rRNA sequences and the *SINE3-1_AO* sequences from *A. oryzae*. Only the 5S rRNA-related regions (nt 1-119) of each sequence were selected for the alignment which was performed using the Clustal Omega programme available at the EMBL-EBI website.

5.2.2 tRNA-related *SINEs* (*AfuSINE2*)

Eight sequences of tRNA-related *SINEs* (designated as *AfuSINE2*) were identified in the genome of *A. fumigatus* Af293, whose sizes ranged from 140-493 bp.

Sequence analysis showed that *AfuSINE2* lacked any discernible ORFs and possessed characteristics typical of tRNA-derived *SINEs* (Table 19). The sequences containing tRNA-derived regions were aligned against *A. fumigatus* Af293 tRNA sequences to identify internal RNA polymerase III promoter sequences (A and B boxes), which identified a consensus sequence of 11 bp, separated by a 22-46 bp sequence. However, the *AfuSINE* RNA pol III promoter sequences were degenerate and the A box sequences were discovered to be more degenerate than that B box sequences. Possible consensus sequences of RNA polymerase III promoter A box and B box of tRNA-derived *AfuSINEs* are shown in Fig. 56. It was discovered that the predicted tRNA-related *SINEs* do not have an oligo T sequence immediately at the 3'-terminus of the tRNA structure as observed for the tRNA genes. Instead they have an extended tRNA-unrelated sequence at the 3'-terminus that may (or may not) be A-rich (but not as a polyadenylation signal), followed by an oligo T tract which acts as a termination signal. Additionally, repeats of the target site duplication of *AfuSINE2* sequences varied in size from 4-14 bp which is a characteristic of *SINE* insertions into the genome.

RNA pol III	A Box	B Box
	TGGCNNAGTNGG	GGTTCGANNCC
	TRKYNNAR-NGG	RGTTCRADYYY
	TRKYNNARWNGG	RGTTCRANHYY
	TRGYNBAR-YGG	
	TRGYNBARTYGG	
Consensus	TGGCGCAR-TGG	GGTTCGATTCC

Fig. 56 Possible consensus A and B box promoter sequences of RNA polymerase III in tRNA-derived *AfuSINEs* (*SINE2*).

By analyzing the genomic sequences using the CENSOR programme, the results revealed that two *AfuSINE2* sequences contained masked sequence similar to some tRNA-derived *SINEs* published in the Repbase database. The 5' tRNA-related region of *AfuSINE2-4a* appeared to be similar to a 268 bp *MIRc* (*SINE2*/tRNA#Pro) identified in mammals (Smit and Riggs, 1995) whereas the 5' tRNA-related region of *AfuSINE2-3a* is similar to a 481 bp *LFSINE-Vert* (*SINE2*/tRNA#His) identified in *Latimeria menadoensis* (Bejerano *et al*., 2006). However

phylogenetic analysis revealed that *AfuSINE2-1a* appears to be closely related to the *MIRc SINE.*

Interestingly multiple sequence alignment showed that most of the candidate *AfuSINE2* sequences appeared to be degenerate (Fig. 57). No significant similarity was observed among the predicted *AfuSINEs* and they seem to originate from different tRNA-related sequences. The ancestral tRNA of each *AfuSINE2* was identified using the tRNA-SE Scan and multiple sequence alignment of the predicted tRNA sequences. The results indicated that *AfuSINE2* shows sequence similarity with the 5'-tRNA-related regions of $tRNA^{Ala}$, $tRNA^{Arg}$, $tRNA^{Gly}$, $tRNA^{Met}$, $tRNA^{Thr}$and $tRNA^{Ser}$. Phylogenetic analysis revealed that all predicted *AfuSINE2* sequences group into five distinct lineages with a 25-35% divergency between them (Fig. 58).

```
AfuSINE2-3c#ala          -GTTGTCTCGGCACAGATGGCAAAGCCGT-----------------TCCT----------
AfuSINE2-4a#MIRc         GGTGTATGCGGCCTGGGCAGAAATATGATGGGTAGATATCAAGATATCGTAAGTATTCCA
AfuSINE2-7e#met          ----------ACTTATATAGCAGAGTGGTAGAGT---CA-------TAGGCTTCATATTT
AfuSINE2-4c#thr          ------------------------------------ATA---------CGTATCCT----
AfuSINE2-5d#ser          ----------GCTACGTAGGCTTAATGG---------AA--G--------GAACATGTTC
AfuSINE2-3a#LFSINE       ---TTGGAGGGACTGGCCGGCTCTGGGGTCGGTCGTTAA--GGCGCTCCGCCATTCAT--
LFSINE_Vert              ----------GACTGGATGGCTCAGTGG---------AA----------------------
AfuSINE2-7a#gly          ----------TTCCATTTGGCGGAATGG---------AA--A--GGTTCGAAACAGCCCT
MIRc                     ----------GGCAGTGTGGTGCAGTGG---------AA----------------------
AfuSINE2-1a#arg          ----------GTAGCGGTAGCGTAGTGG---------TA--AGCGCTCCGAGGCAGCCTC

AfuSINE2-3c#ala          -GGCAACAGTT-----------------AATTGT---------TCAGTTCAAATCTGTTA
AfuSINE2-4a#MIRc         GCTCCTCCACTGATTGACGATCCTTTTGTTATCTTGAGC-CAAGGAATTTGGTGCCTTTG
AfuSINE2-7e#met          TTACAGCAATGG----------------AGATG-----AGATAAGGATTCGATTTCTAT-
AfuSINE2-4c#thr          -GCAATGGGCAGAATCAACAACGTGT--ACATTTTACGCGTAAATGGATCGATTCCAT-A
AfuSINE2-5d#ser          GAGGATGGTCGGAAAATCGCTGG-----AGGTGTTCAATCCCCGAGGTACGTCATCTG--
AfuSINE2-3a#LFSINE       -CTGCAGGTTGGATAACAGCCGG--G--CAATTT------TCGTGACTGAAAATCCCATC
LFSINE_Vert              -TTGGTAATGGGATATGGAGCCTT-T--CACCTCTAGGT-CACTGGGTTCAAATCCAGCC
AfuSINE2-7a#gly          GGAGGTGTTCGAATAGTAGAGTAT----------------GGTCAGCTTCGCTTCCTGCT
MIRc                     -AGAGCACTGGACTT----------G--GAGTCAGGAAG-ACCTGGGTTCGAGTCCTGGC
AfuSINE2-1a#arg          TAGAGAAGTAGGATTGGTGACCCTGC--TTTTTTGGAGG-TTATGGGTTCGATTCCCGTC
```

Fig. 57 Multiple sequence alignment of the *AfuSINE2* sequences. Only the tRNA-related region (nt 1-72) of each sequence was selected for alignment which was performed using the Clustal Omega programme available at the EMBL-EBI website. Potential A and B boxes are highlighted in pink.

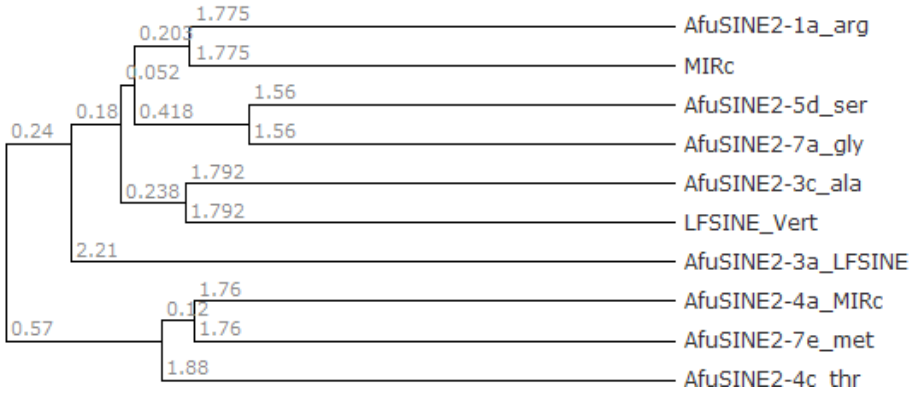


Fig. 58 Phylogram of the *AfuSINE2* candidates. Only the tRNA-related region (nt 1-72) of each sequence was selected for alignment which was performed using the MAFFT alignment programme.

5.3 Secondary structure of tRNA-derived *AfuSINE2s*

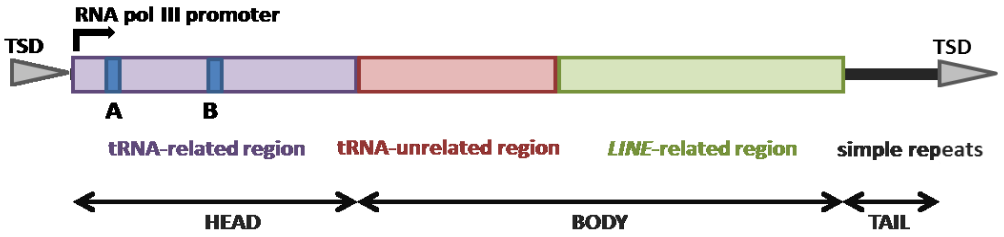
Possible secondary structures of each *AfuSINE2* were predicted using the mfold programme (Table 19). The results showed that the 5'-tRNA-related region (nt 1-73) of each *AfuSINE2* form potential loop structures but do not fold into a tRNA cloverleaf structure. However, some *AfuSINE2s* resembled a cloverleaf structure where the 3'-ends of the 5'-tRNA-related regions terminate with CCA, a characteristic of tRNA. This includes *AfuSINE2-4c, AfuSINE2-5d* and A*fuSINE2-4a.*

5.4 Similarity of *AfuSINEs* to retrotransposons

Interestingly the sequence of the *AfuSINE2-1a* 3'-terminus was found to be similar to *LINE* element, *I-4_AO#LINE/Tad1,* identified in *A. oryzae* (Fig. 59; Galagan *et al*., 2005; Kapitonov and Jurka, 2006) when the sequence was analyzed using CENSOR. Subsequently, the element was aligned manually against retrotransposon found in the *A. fumigatus* Af293 genomic sequence such as a retrotransposon-like element (*Afut1-LTR*; Neuveglise *et al*., 1996) and a non-LTR retrotransposon (*I-1_AF*; Galagan *et al*., 2005; Kapitonov and Jurka, 2006). Further analysis revealed that the sequence of the *AfuSINE2-1a* 3'-terminus at positions 210-335 appears similar to the 3'-UTR of the *I-1_AF LINE*-like sequence which terminates the reverse transcriptase gene in ORF2. It has been proposed that sequence and structural similarity of the 3'-terminal regions of tRNA-derived *SINE* sequences to the corresponding *LINE* sequences is crucial for *SINE* transcription and retrotransposition. Thus, this region is required for the recognition by *LINE* encoded reverse transcriptase (Fantaccione *et al*., 2008). Additionally, it was noted that the 3'-terminus of *AfuSINE2-1a* terminates with a series of short repetitive ACT trinucleotide sequences.

A diagram of the structure of *AfuSINE2-1a*, sequence and sequence alignment between *AfuSINE2-1a* and *I-4_AO* are shown in Fig. 59 which illustrates that the element represents a typical *SINE* structure, comprising head (tRNA-related region of RNA pol III A and B boxes), body (tRNA-unrelated region and *LINE*-related region) and tail (simple repeats).

AfuSINE2-1a



>*AfuSINE2-1a*

*AAGTGTACATAGAG*CCGGTAGCGGTAGCGTAGTGGTAAGCGCTCCGAGGCAGCCTCTAGAGAAGTAGGATTGGTGACC
CTGCTTTTTTGGAGGTTATGGGTTCGATTCCCGTCGCTGGCACAACATTTACCACCACAATGGAAGATCACTTCCCAC
AATGGTATCAAGGCCACTCCCTTATCGCAAGGTGGTGGGGAAGTTGGAACAATCACAGGCCTGTAAGGCGAGGCTCTA
AATTCGCCCTCATGTAATGGAACAAAATGTAACTAGACACACAAGGATTAGCTATAGTCGATACCTGCATATCGCCCA
AGGCGAGGGGTCAGCGTATGAGT*ACTACTACTACTACTACTACT**AAGTGTACATAGAG*

>Sequence alignment between *AfuSINE2-1a* and *I-4_AO LINE*

```
AfuSINE2-1a      CACAGGCCTGTAAGGCGAGGCTCTAAATTCGCC-CTCATGTAATGGAACAAAATGTAACT
I-4_AO           CACAGGCCT---TGACGAGGCTCCAAGTTCGTCAGGAAAGTAATAGTTTACTGTAAATAG
                 *********    *.********.**.****.*    * *****.*  .*  .*. *

AfuSINE2-1a      AGACACACAAGGATTAGCTATAGTCG-ATACCTGCATATCGCCC-AAGGCGAGGGGTCAG
I-4_AO           ACGCACAGATAGAGTACAGATAGTGGCGGTCCCGCATATCGCCCTGAGGCGAGGGGTCGG
                 * .**** * .** **   ***** * .  **.*********** .************.*

AfuSINE2-1a      CGTATGAGT
I-4_AO           CGTATGAAT
                 *******.*
```

Fig. 59 Structure and sequence of the *AfuSINE2-1a* tRNA-derived *SINE* found in *A. fumigatus* Af293 and the alignment of its 3'-terminus which is related to *I-4_AO LINE*. Target site duplication (TSD) is underlined; 5'tRNA-related regions is hilighted in blue; RNA polymerase III promoter sequences A and B boxes are hilighted in pink; short repetitive sequence (ACT) is shown in green; oligo T tract/poly A is shown in red; and the region similar to *I-4_AO*# *LINE*/Tad1 of *A. oryzae* and *I-1_AF A. fumigatus* Af293 *LINE* is highlighted in grey.

5.5 Distribution and location of *AfuSINEs*

The distribution and location of *AfuSINEs* on 8 chromosomes of the *A. fumigatus* Af293 genomic DNA were studied by comparative analysis of the sequences. Bioinformatic analyses revealed that *AfuSINEs* are not abundant in the *A. fumigatus* Af293 and they are absent from chromosomes 2 and 6. Examination of the other six chromosomes revealed that pericentromeric and subtelomeric regions are enriched with *AfuSINEs*. *AfuSINE3s* representing 5S rRNA-related *SINEs* are found on chromosomes 1, 3, 4 and 5 whereas the *AfuSINE2s* representing tRNA-related *SINEs* are found on chromosomes 1, 3, 4, 5 and 7. *AfuSINEs* are most abundant on chromosomes 3 and 4 and an overview of the genomic locations of the predicted *SINEs* is shown in Fig. 60.

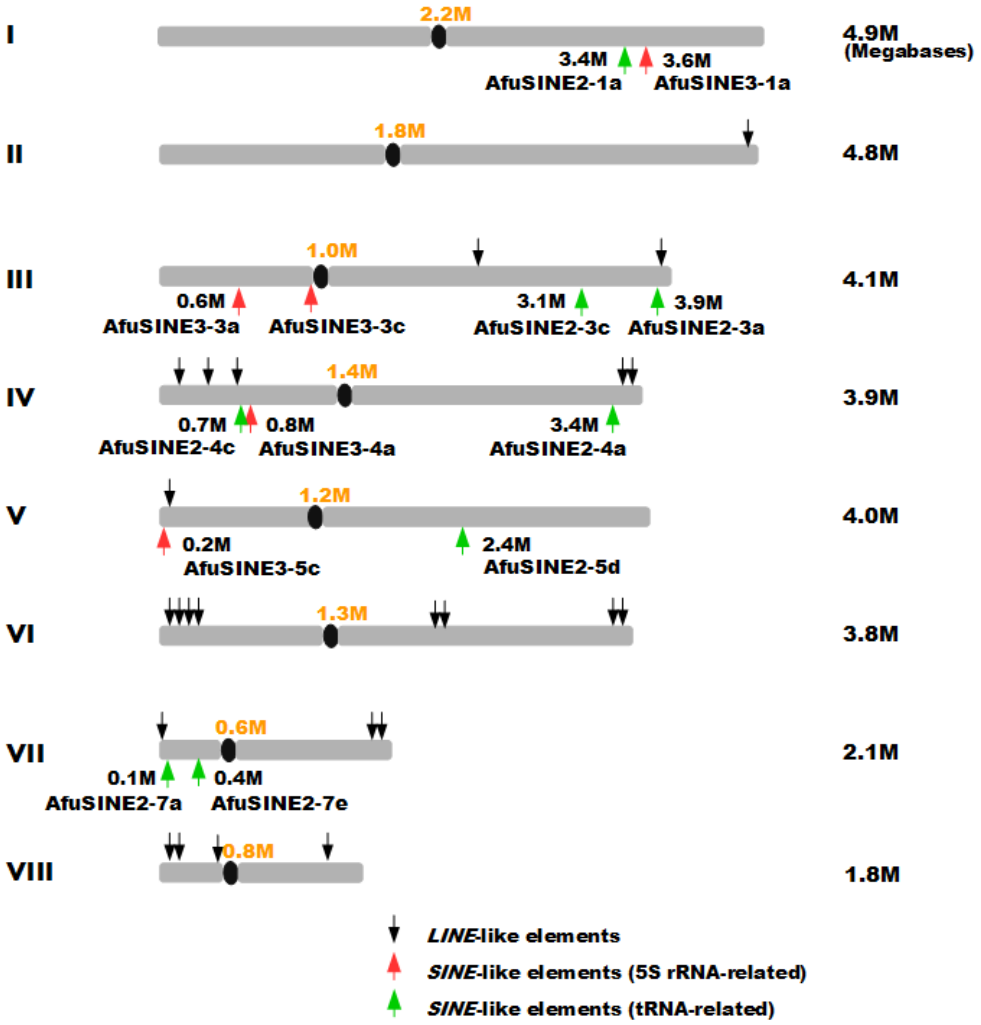


Fig. 60 Mapping of *SINE*-like sequences on eight chromosomes of the *A. fumigatus* Af293 genome including previously described *LINE*-like elements (LLEs) (Huber and Bignell, 2014). Chromosome numbers are shown on the left and the size of each chromosome is shown on the right.

Identification of the insertion of *AfuSINEs* revealed that the elements are randomly dispersed within the gene-rich regions-normally in the intergenic regions of the genes or close to the sequence of coding genes. Interestingly *AfuSINE2_4c* on chromosome 4 appears to be located very close to a retrotransposon region of the *LINE*-like sequences (LLE#4_3.0), a subclass previously identified in the *A. fumigatus* Af293 genome (Fig 61). The LLE#4_3.0 sequence is a non-LTR retrotransposon from I clade which is flanked by 13-bp target site duplications (TSDs) and contains two overlapping ORFs, ORF1 and ORF2. ORF1 encodes the 413-aa DNA/RNA-binding protein (pos. 175-1669) and ORF2 encodes 1273-aa polyprotein (pos. 1666-5487) comprising an endonuclease and a reverse transcriptase. Additionally comparative analysis revealed that *AfuSINE2_4c* is located next to the 3' TSD and close to the RT of LLE#4_3.0 (Fig. 61).

Fig. 61 Structure of *LINE*-like element (LLE#4_3.0) on *A. fumigatus* Af293 chromosome 4 showing insertion of the *AfuSINE2_4c* sequence.

5.6 Transcription activity of *AfuSINEs* in the genome by RT-PCR

Oligonucleotide primers designed to bind to regions of the A and B boxes of individual *SINEs* were used for PCR amplification. It is very important that primers used for PCR amplification of *AfuSINE* sequences were carefully designed to discriminate between *AfuSINEs* and tRNA sequences as their sequences are often very similar. Initially, PCR amplification was performed using genomic DNA as a template under high stringency PCR conditions. The results of these experiments showed the expected clear amplicons for the 13 candidate *AfuSINEs* (Fig. 62, lane 1) which represented confirmatory evidence of the existence of the sequences on the *A. fumigatus* Af293 genome and also proved that genomic information on the database is correct. Additionally a ladder of DNA amplicons of increasing size was not found following PCR amplification indicating that *AfuSINE* sequences are not present as an array on the genome.

Subsequently, RT-PCR analysis was performed using total RNA as template with the same sets of primer pairs to investigate the production of *AfuSINE* transcripts by genomic DNA. The results showed that 6 out of 13 *AfuSINEs* are transcriptionally active (*viz*. *AfuSINE3-1a, AfuSINE3-3c, AfuSINE2-1a, AfuSINE2-3a, AfuSINE2-4a* and *AfuSINE2-5d*) since the amplicons of the correct anticipated sizes were generated following RT-PCR amplification (Fig. 62, lane 2). No amplicons were observed from the cDNA(-RT) samples, indicating that contaminating gDNA was absent from the reaction mixture for cDNA synthesis (Fig. 62, lane 3). All RT-PCR products were sequenced to confirm their identity. As illustrated by these results the levels of transcription of the *AfuSINEs* are very low as compared to the β-tubulin gene which is a constitutetively expressed house-keeping gene.

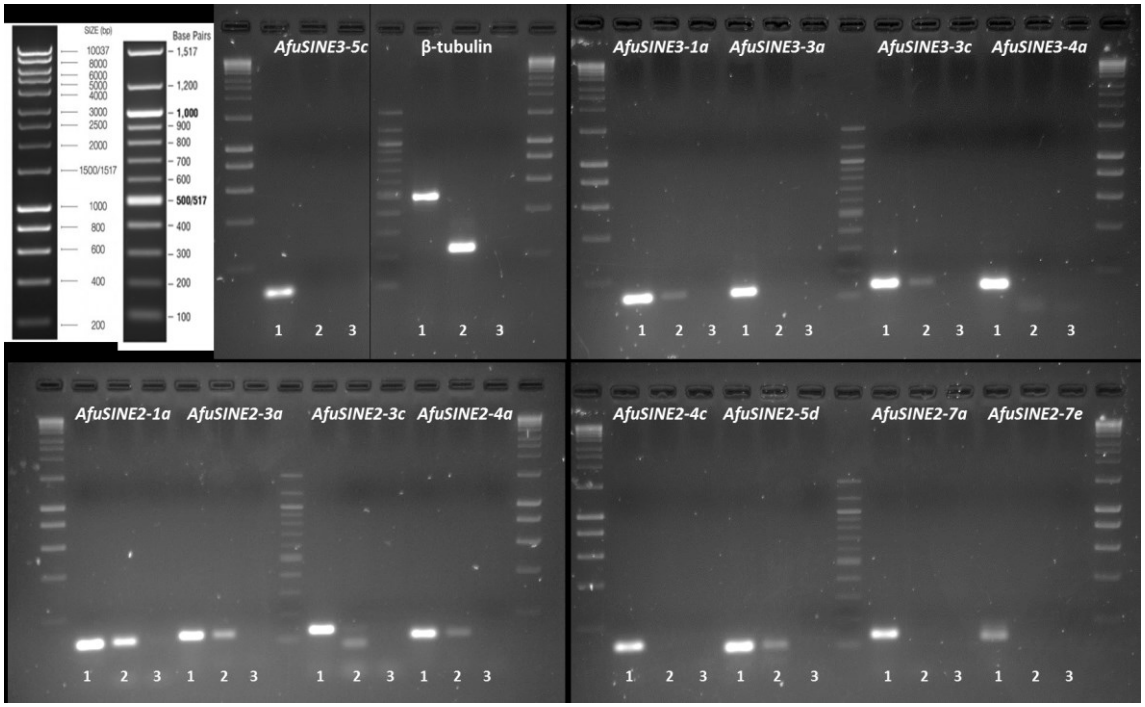


Fig. 62 Agarose gel electrophoresis showing the PCR and RT-PCR products of the 13 candidate *AfuSINEs*; Lane 1 for each *AfuSINE* shows PCR amplicons generated from genomic DNA; Lane 2 for each *AfuSINE* represents amplicons generated following RT-PCR amplification; Lane 3 for each *AfuSINE* represents amplicons generated from (-RT) negative controls) RT-PCR. Hyperladder 1 (M; 10 kbp; Bioline) and Quick-Load® 100 bp DNA Ladder (NEB) were used as markers. Electrophoretic analysis was performed in 2.5 % agarose gel for 3 h at 80 V.

5.7 Estimation of *AfuSINE* copy number in the *A. fumigatus* Af293 genome

In this study only *AfuSINEs* (*AfuSINE3-1a, AfuSINE3-3c, AfuSINE2-1a, AfuSINE2-3a, AfuSINE2-4a* and *AfuSINE2-5d*) previously shown to be actively transcribed were investigated. The copy numbers and transposition of the six *AfuSINEs* above per genome of two isogenic lines of *A. fumigatus* Af293, one infected with AfuTmV-1-and one virus-free (NK125) were examined by southern blot hybridization. For southern blot hybridization, *A. fumigatus* Af293 genomic DNA was digested with *Hind*III a restriction enzyme that does not restrict the *AfuSINE* sequences, prior to hybridisation with DNA probes specific for each individual *AfuSINE*. Therefore, it can be assumed that any single hybridising band observed correspond to one copy of each *AfuSINE.*

The hybridization analysis of tRNA-derived *AfuSINE2s* showed a strong, single hybridization signal for *AfuSINE2-3a, AfuSINE2-4a* and *AfuSINE2-5d* indicating that these elements are present as single copy in the genome (Fig. 63). Four or more strong hybridization signals were detected for *AfuSINE2-1a* and *AfuSINE2-3c*, indicating that at least four copies of these

elements are present in the genome. At least five hybridization signals were detected for both *AfuSINE3-1a* and *AfuSINE3-3c* indicating multiple copies of each *AfuSINE3* dispersed in the genome. The *A. fumigatus fks* gene, which is present as a single copy in the genome, was used as a positive control in these experiments. Both virus-free and virus-infected isogenic lines showed identical hybridization patterns indicating that virus infection does not apparently have any effect on *SINE* activity.

In contrast to other reports on *SINE* copy number, *AfuSINEs* from *A. fumigatus* Af293 *SINEs* are present in low copy number. Because of known and significant divergences in the Af293 *SINE* sequences and potential insertions and/or deletions within the *SINEs* it is likely that the copy numbers estimated here are an under-estimate and represent the minimum copy numbers for each element. It is likely that *A. fumigatus SINE* RNAs are silenced since aspergilli encode all of the enzymes required for the process and are active *in vivo* (Hammond *et al*., 2008b).

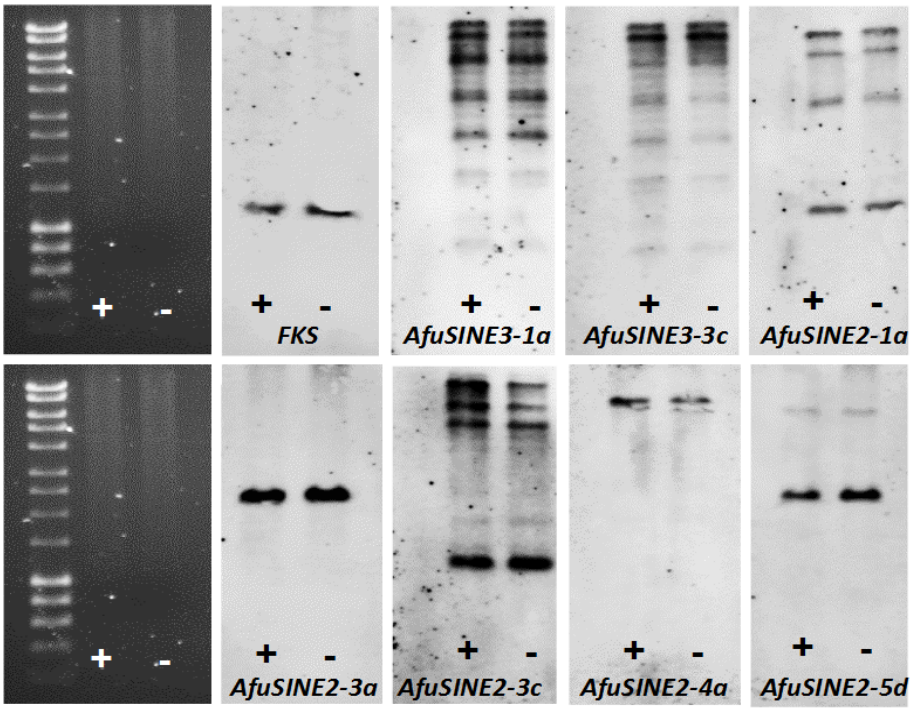


Fig. 63 Southern blot hybridisation of *AfuSINEs* in the genome of *A. fumigatus* Af293 (virus-infected strain; indicated by plus sign) and *A. fumigatus* NK125 (virus-free strain; indicated by minus sign). *Hind*III-digested DNA of each strain was separated in 1% agarose gel in 1xTAE, denatured and blotted onto nylon membrane. Hyperladder 1 (M; 10 kbp; Bioline) was used as marker.

5.8 Detection of small RNA molecules homologous to *AfuSINEs*

In gaining and understanding of *AfuSINE* regulation and gene silencing in *A. fumigatus* Af293, small RNAs homologous to *AfuSINE* sequences were identified using northern blot hybridization. Antibody digoxigenin-labelled probes of *AfuSINEs* were *in vitro* transcribed in both sense and antisense orientation to detect strand-specific small RNAs. In this study, only the representative actively transcribed *AfuSINEs* (A*fuSINE3-1a, AfuSINE2-1a, AfuSINE2-3a* and *AfuSINE2-4a*) were selected for detection of small RNAs.

Total RNA, high molecular weight RNAs (HMW RNAs) and low molecular weight RNAs (LMW RNAs) fractions were analyzed on 1.5% agarose and 15% (w/v) polyacrylamide Tris-borate-EDTA-urea gels (Fig. 64a-b). Autoradiographic analysis showed a weak hybridization signal of a band corresponding to <40 nt was detected only for the antisense (-) strand of the A*fuSINE3-1a* and *AfuSINE2-1a* small RNA samples (Fig. 64d). However, small RNAs homologous to *AfuSINE2-3a* and *AfuSINE2-4a* were not detected. Constitutively expressed control β-tubulin probes were also used to detect small RNAs homologous to the β-tubulin gene. No hybridization signal was found from the control, suggesting that hybridization of small RNAs by *AfuSINE* probes may not attribute to mRNA degradation.

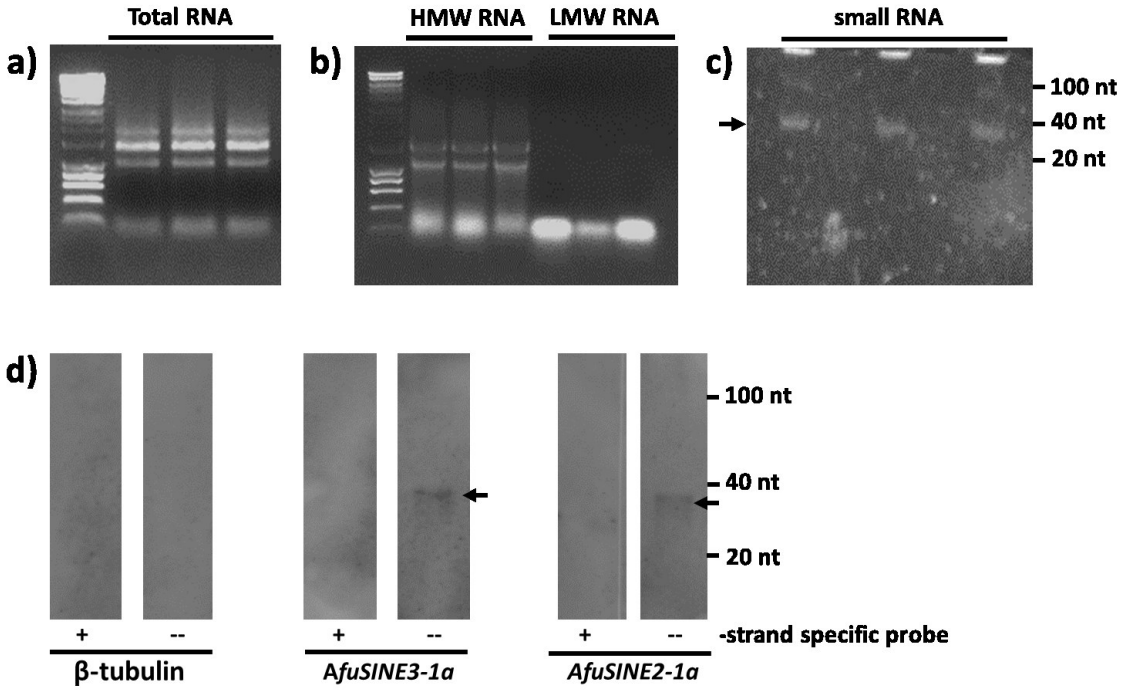


Fig. 64 Detection of small RNA molecules homologous to *AfuSINEs.* Total RNAs isolated with Trizol were run on 1.5% agarose gel (a). After separation by PEG precipitation, high molecular weight RNAs (HMW RNAs) and low molecular weight RNAs (LMW RNAs) were run on 1.5% agarose gel (b). The LMW fractions from 400 ng/µl of total RNAs were resolved on a 15% (w/v) polyacrylamide Tris-borate-EDTA-

urea gels and stained with SYBR Gold Nucleic Acid Gel Stain (c). Northern blot hybridization of small RNAs homologous to *AfuSINEs* in *A. fumigatus* af293 wild type (d). LMW RNA fractions were isolated using TRIzol and 5 µg RNA sample was loaded into each well for Northern blot analysis. Only small RNAs (<40 nt; arrowed) from the antisense strand were detected.

Discussion (Part III)

Here I report for the first time the presence and distribution *SINE* families the *A. fumigatus* Af293 genome using bioinformatics tools to interrogate the *A. fumigatus* Af293 genome sequence. Two main bioinformatic tools were used in this study the CENSOR and RepeatMasker programmes. These programmes also revealed the sequences of several transposons (*Mariner* class II DNA transposon), LTR-retrotransposons (*Copia*, *Gypsy* and *Afut1*) and non-LTR retrotransposons (*I-1_AF LINEs*) in the fungal genome, including 5S rRNA sequences. The numbers of *I-1_AF LINEs* and 5S rRNAs predicted from the *in silico* analysis of genomic DNA shown here are comparable to the numbers of these elements described in several publications (Galagan *et al*., 2005; Kapitonov and Jurka, 2006; Novikova *et al*., 2009 and Huber and Bignell, 2014) and confirm the usefulness of the CENSOR and RepeatMasker programmes for such investigations. However since these programmes identify exact sequence matches any retrieved data requires further screening with the tRNAscan-SE programme to eliminate real tRNA and 5S rRNA sequences.

Computational analysis of the sequence of the eight chromosomes of the *A. fumigatus* Af293 genome identified *AfuSINEs* with predicted repetitive sequences based on the available sequence information on the Repbase database. Further analysis showed that all of the *AfuSINEs* possessed features in common with generic *SINE* sequences. However in some *AfuSINEs* target site duplication and intact RNA polymerase III internal sequences are missing. The *SINEs* described can be divided into two *AfuSINE*-like families, tRNA-related (*AfuSINE2*) and 5S rRNA-related (*AfuSINE3*) *SINEs,* which are *ca.* 259-343 bp in size for the *AfuSINE3* sequences and 140-493 bp for the *AfuSINE2* sequences.

A family of five *AfuSINE3* sequences was discovered in the *A. fumigatus* Af293 genome which are homologous to the 5' terminal region of fungal 5S rRNA sequences. Generally 5S rRNAs play an important role in stabilizing ribosome structure during protein synthesis. Elements are present as tandem repeated sequences in multiple copies in the genome and contain GCTTTTCG RNA polymerase III termination signals (Paule and White, 2000; Barciszewska *et al*., 2001). The 5'-terminal 120 bp of 5S rRNA genes are highly conserved among different species and found to be *ca*. 99% identical in any particular genome. A comparison of the five *AfuSINE3* sequences revealed similar conservation at the 5'-terminus and a significant similarity (65%) to the 5' terminal sequences of 5S rRNA genes. However, regions of three RNA polymerase III

domains (A, IE and B boxes) are not well-conserved in *AfuSINE3* sequences as compared to typical 5S rRNAs.

The *AfuSINE3* sequences are significantly homologous (75%) to *SINE3-1_OA*, a 5S-derived *SINE* present in *A. oryzae.* Additionally the *AfuSINE3* sequences lack an RNA pol III transcription termination signal, except for the *AfuSINE3-1a.* This feature results in continuous transcription activity extending the 5S rRNA-related region to the 3' termini of the *AfuSINE3s*. The *AfuSINE3s* are present in very low copy numbers of *ca.* five copies of each candidate in the genome following *in silico* analysis. This is in contrast to other eukaryotic *SINE3* families where *ca.* 10^4 copies of *SINE3* are present in the zebrafish genome with an average length of 600 bp and a 5'-terminal sequence which is 75% identical to similar regions on 5S rRNA genes (Kapitonov and Jurka, 2003).

A family of tRNA-derived *SINE*-like elements (*AfuSINE2*) was also identified in the *A. fumigatus* Af293 genome consisting of eight representative candidates that have apparently originated from different tRNAs. Sequence comparisons of individual *AfuSINE2s* revealed that the sequences of the *AfuSINE2s* are not conserved with little sequence similarity. Additionally the *AfuSINE2* A-and B-box RNA pol III promoter sequences are both degenerate, the former more so than the latter. Repeats of the target site duplication (TSD) sequence, a characteristic feature of *SINE* genomic insertion, were present and varied in size from 4-14 bp. However TSD sequences were not found on all *AfuSINE2s*. For instance the *AfuSINE2-3c* and *AfuSINE2-5d* TSD sequences appeared to be partially degraded with 1 or 2 bases missing. Such degradation phenomenon might reflect the ages of a *SINE* and concern nucleotide substitution over many generations over an extended period of time (Whisson *et al*., 2005a). As a consequence, it can be inferred that *AfuSINE2-1a* might have emerged more recently that other *AfuSINE2s* since intact direct repeats were observed in its sequence. It has been reported that *SINE* insertion into the genome usually produces a short TSD of 5-8 bp at the 5' and 3' ends of the sequence. However, if *SINE* insertion is very ancient any identification of the TSD might be difficult due to mutations in the sequence. Also a TC motif located immediately upstream of the A/T-rich tail and an additional B- box downstream of the pol III promoter were absent from the *AfuSINE2* sequences although this feature has been reported in some tRNA-related *VES SINEs* and *P.k. SINEs* in bats (Borodulina and Kramerov, 1999; Fantaccione *et al*., 2008). Nearly all *AfuSINE2s* possess a 3' tRNA-unrelated sequence of variable length upstream of a poly-T tract of 4 Ts or greater length TTTNTTT. This is in contrast to genuine tRNAs where the sequence terminates in poly-T (usually 4 or more) immediately after the main coding region. The poly T tract has been

described as a transcriptional terminator sequence for RNA polymerase III. None of the 5'-terminal regions of the *AfuSINE2s* could be folded into tRNA-like cloverleaf structures which is in contrast to the *MgSINE* and the *MIR SINE* (Smit and Riggs, 1995; Kachroo *et al*., 1995).

Among eight recognized *AfuSINE2s, AfuSINE2-1a* appears to be the most interesting element since it possesses several unique features as compared to the other *AfuSINE2s* and was found to be closely related to a *MIRc SINE* identified in mammals (Smit and Riggs, 1995). The element demonstrated active transcription, intact target site 14 bp duplications and sequence similarity of the 5'-tRNA-related region to the $tRNA^{Arg}$ sequence, as reported for most tRNA-derived *SINEs* (Daniels and Deininger, 1983; Fantaccione *et al*., 2004; Fantaccione *et al*., 2008). Additionally the sequence of its 3' tRNA-UTR was similar to the *A. oryzae I-4_AO SINE* sequence and the *I-1_AF LINE*-like element identified in the *A. fumigatus* Af293 genome. Also unique short repetitive sequences of ACT were present.

Since 3'-terminal sequence of *AfuSINE2-1a* is similar to the sequence of the 3' UTR of the *I-1_AF LINE* recently identified in the *A. fumigatus* Af293 genome and the *I-1_AF LINE* 3' UTR is downstream of the RT gene in *LINE* ORF2, it is possible that the element might be recognised by the RT for retrotransposition and amplification. Summarizing the information presented, *AfuSINE2-1a* could be a functional *SINE* and that its transposition might rely on the enzymes encoded by *I-F_AF LINE*. Since *AfuSINEs* are non-autonomous retrotransposons with no gene coding capability they need to exploit the action of other retrotransposons (often *LINEs* and possibly also *Copia* and *Gypsy* types) for their replication and insertion into the genome. Thus, the presence and transcription activity of retrotransposons such as *I-1_AF LINEs* in the genome could signify the presence and activity of some *AfuSINEs*.

A. fumigatus SINEs are present in very low copy numbers in the fungus genome as only single copy were detected for *AfuSINE2s* and single few copy numbers for *AfuSINE3s.* Since numbers of repetitive elements in *A. fumigatus* are limited, this could contribute to the small size of the *A. fumigatus* genome (29.4 megabases). These observations are in contrast to the situation with the *Phytophthora infestans* genome which consists of a large proportion of TEs and repeats contributing to a genome size of 240 megabases (Whisson *et al*., 2005a). In the rice blast fungus *Magnaporthe grisea* the copy number of the *Mg-SINE* was predicted to be approximately 100 (Kachroo *et al*., 1995). *SINE* abundance is clearly diverse and variable between different fungal species. Low copy number non-LTR retrotransposons with degenerate sequences are likely to be lost from the genome as a result of genetic drift and natural selection (Brookfield and Badge, 1997). However, complete non-LTR retrotransposons

with no degenerate sequences even as single copy can retrotranspose and invade a population to compensate for any lost retrotransposons (Le Rouzic and Capy, 2005). In *A. fumigatus* strain A56 it has been shown recently that chrysovirus infection stimulates mobilization of LLEs (Huber and Bignell, 2014). However infection of *A. fumigatus* strain Af293 with AfuTmV-1 has no effect on *AfuSINE* transposition or copy number.

From RT-PCR analysis it is clear that not all of the *AfuSINEs* in the *A. fumigatus* Af293 genome are transcribed and only 6 out of 13 *AfuSINEs* are transcriptionally active (*AfuSINE3-1a, AfuSINE3-3c, AfuSINE2-1a, AfuSINE2-3a, AfuSINE2-4a* and *AfuSINE2-5d*). This observation might can be explained by the fact that:- 1) either the RT synthesized by the partner *LINE* or *LINE* itself is inactive; 2) the internal promoter sequences of some *AfuSINEs* for RNA polymerase III recognition are not optimal; or 3) regulatory signals or cellular factors necessary for *AfuSINE* amplification have been lost or cannot function. Much still needs to be elucidated concerning *AfuSINE* transcription. Poly adenylation is important for efficient *SINE* transcription (Deininger *et al*., 2003) but can occur in its absence or with a short 3'-terminal tail of 0 to 5 adenylate residues for some *AfuSINEs.* It is suspected but unproven that trans-mobilization by other active retrotransposons might be involved in the transcription activity of the non-poly (A) tract associated *SINEs* (Whisson *et al*., 2005a). Besides the poly (A) tail, intact internal RNA pol III promoter sequences and similarities between the 3'-terminal region of a *SINE* and a partner *LINE* are also important factors in active *SINE* transcription (Weiner, 2002).

AfuSINE elements are dispersed on chromosomes 1, 3, 4, 5 and 7 of the *A. fumigatus* Af293 genome but they are highly abundant on chromosomes 3 and 4. The elements are randomly dispersed in pericentromeric and subtelomeric regions on the chromosomes and inserted within the gene-rich regions, normally in intergenic regions or close to coding regions as reported for LLEs (Huber and Bignell, 2014). These observations suggest that the insertion of *SINEs* and *LLEs* in *A. fumigatus* chromosomes may be linked. The insertion of *AfuSINE2_4c* on chromosome 4 next to the LLE#4_3.0 *LINE*-like element and close to the LLE#4_3.0 RT region suggests that activity of *AfuSINE2_4c* possibly relies on this *LINE*. However *AfuSINE2_4c* is transcritionally inactive possibly caused by inactivation of of the LLE#4_3.0 RT domain by mutation. Since most of the LLE sequences identified in the genome of *A. fumigatus* Af293 are not intact, this could contribute to loss in transcription and retrotransposition activities for some *AfuSINEs*.

The presence of small RNAs (of <40 nt) homologous to *AfuSINE3-1a* and *AfuSINE2-1a* suggested that the elements may be targeted for degradation and silencing in the fungus.

However, small RNAs were not detected in the case of *AfuSINE2-3a* and *AfuSINE2-4a*. This indicated that not all *AfuSINEs* are targeted by host RNA silencing mechanisms. In addition, it can be observed that hybridization signals observed from the northern blot analysis are very weak. This could attribute to low transcription activity of the elements resulting in low abundance of *AfuSINE* siRNAs.

SINEs are believed to be involved in genome organisation and pathogenicity in some organisms. For example the human *Alu* elements have been describing as a human endogenous mutagens and *de novo* insertions of their sequences in the genome have been linked with disease (Deininger and Batzer, 2002). Additionally abundant *SINE* transcripts has been described in many malignant tumors (Feber *et al*., 2011; Gualtieri *et al*., 2013). However it is unlikely that *LLEs* in *A. fumigatus* Af293 or other clinical isolates play a crucial role or are associated with initiation of respiratory infection in human since *LLEs* are absent from some known pathogenic isolates (Huber and Bignell, 2014). According to the previous findings, it would be intriguing to investigate the activity of *AfuSINEs* on adaptability and pathogenicity of *A. fumigatus* Af293 and other clinical and environmental *A. fumigatus* isolates.

Conclusion (Part III)

1) The presence and distribution of *SINE* family was identified through directed search against the *A. fumigatus* Af293 genomic DNA and the predicted elements are categorized into two *AfuSINE*-like families, tRNA-related (*AfuSINE2*) and 5S rRNA-related (*AfuSINE3*) *SINEs*.

2) Five candidate *AfuSINE3s* related to the 5'-terminal region of 5S rRNA sequences were identified in the genome ranging in size from 259-343 bp. Sequence comparison of five individual *AfuSINE3s* demonstrates the conserved nature among the sequences (56-90%). *AfuSINE3s* are similar to *SINE3-1_OA* identified in *A. oryzae* (75% similarity). However, sequence conservation of the regions of three RNA polymerase III domains (A, IE and B boxes) is indiscernible. The elements are present in a very low copy number approximately five copies of each candidate in the genome when identified *in silico*.

3) *AfuSINE2s* is related to 5'-terminal region of tRNA sequences were identified in the genome ranging in size from 140-493 bp. The elements are found to have arisen from different tRNA-related origins thus any sequence similarity between the predicted elements was not distinct. RNA pol III promoter sequences are found to be more degenerate with A box sequences being more degenerate than B boxes and repeats of the target site duplications (TSD) vary from 4-14 bp.

4) *AfuSINE2-1a* from a family of $tRNA^{Arg}$-derived *SINE*-like elements demonstrates 3' tRNA-unrelate sequence similarity to the *I-4_AO* identified in *A. oryzae* and *I-1_AF LINE*-like element identified in *A. fumigatus* Af293 with a unique short repetitive sequence (ACT).

5) *A. fumigatus SINE*-like elements exist in a very modest copy numbers in the fungus genome and some of them are actively transcribed (6 out of 13). They are dispersed on chromosomes 1, 3, 4, 5 and 7 in pericentromeric and subtelomeric regions and are highly abundant on chromosome 3 and 4. Distributions of the elements are usually found within the gene-rich regions-normally in the intergenic regions of genes or close to the sequence of the coding genes. The *AfuSINE2_4c* from a family of $tRNA^{Thr}$-derived *SINE*-like elements is found to be inserted next to the LLE#4_3.0 *LINE*-like element and locates close to the RT ORF.

6) Comparative studies between isogenic virus-infected and virus-free *A. fumigatus* isolates revealed that AfuTmV-1 infection has no effect on *AfuSINE* activity and retrotransposition of the elements.

Chapter 3

Results and Discussion (Part IV)

(Silencing Vector Construction)

Construction of silencing vectors in the human pathogenic fungus *Aspergillus fumigatus* based on mycovirus and *SINE* element

Aspergillus fumigatus is an opportunistic fungal pathogen which is a cause of fatal invasive aspergillosis among immunocompromised patients. To identify key elements involved in *A. fumigatus* virulence, RNA silencing which is a process associated with specific degradation of a target gene *via* sequence-homology-dependent mechanism has been widely employed. Hairpin RNA expressing plasmid is a versatile tool commonly used to induce RNA silencing in such fungus, however, the approach shows some drawbacks ranging from difficulty in plasmid preparation and instability of hairpin constructs in the fungal cells. This study report the use of alternative silencing vectors by exploiting a mycovirus sequence and short interspersed nuclear elements (*SINEs*) for vector construction. A recombinant virus vector derived from the Aspergillus fumigatus tetramycovirus 1 (AfuTmV-1) was developed to trigger RNA silencing mechanism in the fungal cells by using direct transfection method. Transcriptional fusion *SINE*-derived vector was also developed to silence the *ALB1/PKSP* gene based on the hypothesis stated that silencing of target *SINEs* leads to the degradation of transcriptionally fused endogenous gene transcripts (Anbar *et al*., 2005, Whisson *et al*., 2005a, Lerat and Semon, 2007). The approaches primarily developed in the study would provide a wider application of utilizing mycovirus-based and *SINE*-based vectors as tools in functional genomic studies of filamentous fungi and would allow easier identification of phenotypic consequences resulting from the silencing effect to the target genes.

RNA silencing | virus vector | *SINEs* | *Aspergillus fumigatu*s.

Chapter 3: Results and Discussion (Part IV)
(Silencing Vector Construction)

6. Silencing of *A. fumigatus* Af293 using virus-based and *SINE*-based silencing vectors

Essential tools which provide a simple and rapid deletion or down-regulation of specific genes for functional genetic analyses in the human pathogenic fungus *A. fumigatus* are needed. The development of vectors allows easy investigation on genes involved in fungal pathogenicity and virulence mechanisms which lead to the identification of novel fungal drug targets. In recent years, homologous recombination system is often used in filamentous fungi to generate gene-knockout strains through gene replacement, gene disruption or plasmid integration. However, the recombination occurs at low frequencies, resulting in lower transformation efficiencies. In addition, the approach needs permanent deletion of the target genes which may be lethal to the organisms. Thus, RNA silencing technique which induce transiently expression of the specific gene has become a method of choice to silence gene of interest in several fungi.

Here, the silencing vector system for *A. fumigatus* was constructed based on either mycovirus sequences or fragment of short intersperse nuclear element (*SINEs*). The host endogenous gene to be used as a silencing target is a gene involved in morphological development that shows clear phenotype if the approach was successful. In this case, *ALB1/PKSP* gene encoding polyketide synthase which are responsible for melanin biosynthesis necessary for conidial pigmentation was chosen. Since the *A. fumigatus* Af293 virus-infected (wild type) isolate and virus-free isolates showed undistinguishable colony morphology, it indicates that the virus does not change morphology of the fungus. Therefore, the *ALB1/PKSP* gene is ideally suited for RNA silencing screen as readily discernible phenotype and difference in colony pigmentation can be easily observed.

Construction of gene silencing vectors

6.1 Construction of gene silencing vector based on AfuTmV-1

As previously described in the characterization and infectivity studies of (AfuTmV-1 chapter 3, part II), its unusual unencapsidated state and replication properties make it suitable for exploitation as an RNA-silenicng vector.

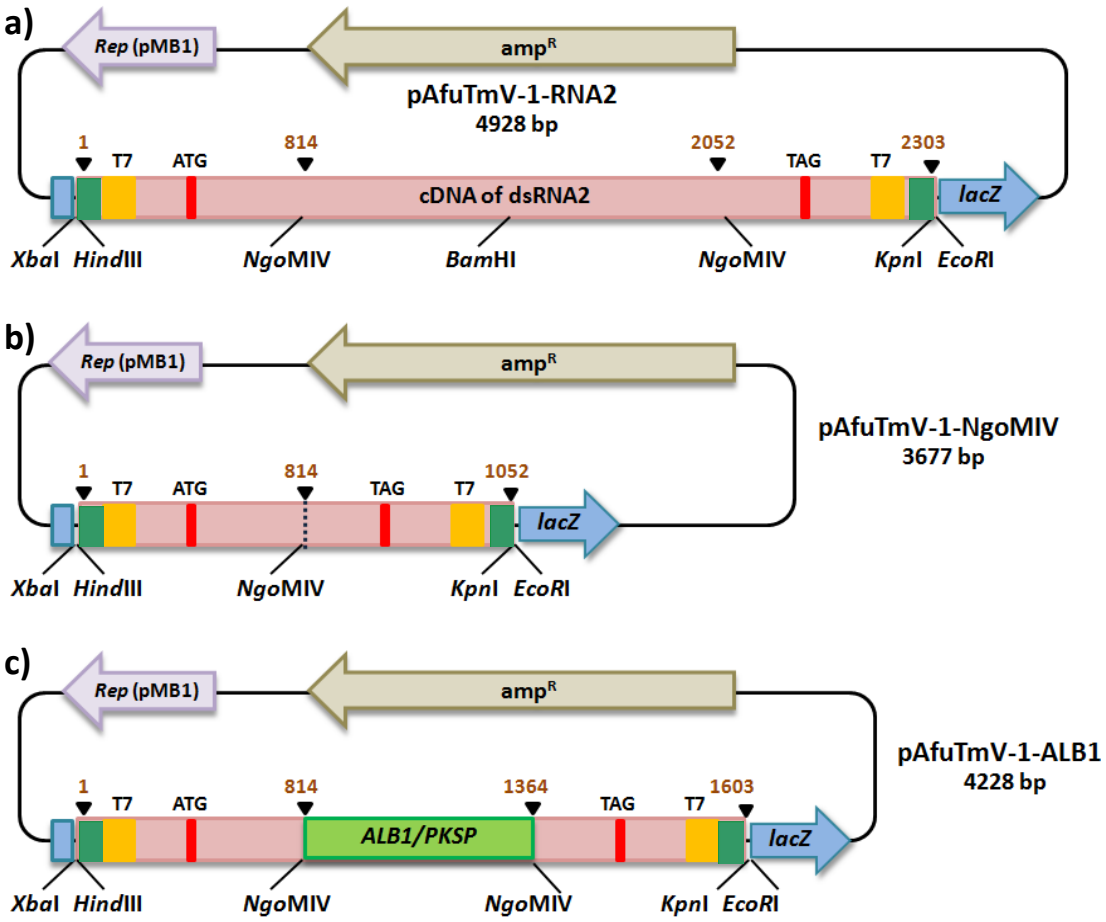


Fig. 65 Schematic representation of the mycovirus-based silencing vector constructs. A modified cDNA clone of the AfuTmV-1 dsRNA2 carring T7 RNA promoter sequences and restriction enzyme cloning sites at both 5'- and 3- termini, designated as pAfuTmV-1-RNA2 (a). A truncated version of cDNA clone, designated as pAfuTmV-1-NgoMIV (b). A modified pAfuTmV-1-ALB1 vector carrying untranslatable sense and anti-sense *ALB1/PKSP* fragments (c). All constructs were modified using pUC18 as a background plasmid vector. Constructs are not drawn to scale.

Oligonucleotides were designed to generate an amplicon by RT-PCR comprising a full-length cDNA clone of AfuTmV-1 dsRNA2 containing a T7 RNA promoter sequence and restriction enzyme sites on both 5'- and 3'-ends. This amplicon was purified and then subjected to a three-way ligation into pUC18, yielding the **pAfuTmV-1-RNA2** construction (Fig. 65). A

truncated clone of this full-length clone was subsequently constructed from the pAfuTmV-1-RNA2 using two restriction sites for *Ngo*MIV to delete an internal fragment of the dsRNA2 ORF2, resulting in an identifiable **pAfuTmV-1-NgoMIV** (Fig. 65). The construct contains the AfuTmV-1 5'- and 3'- UTRs which are believed to be important sites for virus replication, and includes the N-terminal portion of the ORF. The resulting pAfuTmV-1-NgoMIV construct is a prototype vector to facilitate insertion of non-viral or host endogenous gene fragments. Initial attempts to use pAfuTmV-1-NgoMIV as a gene vector, the construct was replaced the truncated ORF with an untranslatable fragment upstream of the *ALB1/PKSP* host gene nominated **pAfuTmV-1-ALB1** (Fig. 65).

To test the silencing efficiency of the modified AfuTmV-1 mycovirus vector, two plasmids (pAfuTmV-1-NgoMIV and pAfuTmV-1-ALB1), which are both derivatives of the full length pAfuTmV-1 dsRNA2 construct, were transcribed *in vitro via* T7 polymerase promoter sequences at both ends of the plasmid DNAs. Sense and anti-sense transcripts of each construct were synthesized together and allowed to form dsRNA molecules *in vitro*. The synthesized dsRNAs derived from plasmids pAfuTmV-1-NgoMIV and pAfuTmV-1-ALB1 were designated **AfuTmV-1-dsRNA2ΔORF** (1056 bp) and **AfuTmV-1-dsRNA2::*ALB1*** (1562 bp) respectively.

6.2 Construction of silencing vector using transcriptional fusion to a *SINE* element

Fusion PCR was used to generate linear sequence consisting of two fragments of a tRNA-related *AfuSINE2-1α* element and *ALB1/PKSP* sequence encoding gene associated with conidial pigmentation. *AfuSINE2-1α* was selected for transcriptional fusion with the *ALB1/PKSP* sequence because of its active transcription activity and intact RNA polymerase III internal sequence.

The *AfuSINE2-1α* and *ALB1/PKSP* sequences from *A. fumigatus* Af293 genomic DNA were amplified in two separate PCR reactions using oligonucleotide primers that produce overlapping ends. A second PCR was performed using nested primers to fuse the two amplicons together in the correct orientation. Long oligonucleotide primers used in these initial PCR amplification reactions (29-42 mers) generated several unspecific products necessitating gel purification of the correctly sized amplicon prior to cloning into the *Pme*I site of linearized pSK379 plasmid vector containing a pyrithiamine (PT) resistance gene (*ptrA*) and the *E. coli* β-lactamase (Amp^R) gene cassette as selectable markers (Integrative *A. fumigatus* expression plasmid carrying *ᵽgpdA-his2A^t* cassette and *ptrA,* a kind gift from Dr. Elaine Bignell).

The background plasmid vector pSK379 was used for all constructs. Inserts were placed between the constitutive glyceraldehyde-3-phosphate dehydrogenase (*gpdA)* promoter from *A. nidulans* and the *his2A* terminator from the region (his2At) of *A. fumigatus.* More detail on the construction of each *SINE*-based vectors were described in Chapter 2, Materials and Method (12.2). A schematic representation of the *SINE*-based silencing vector constructs for silencing of *ALB1/PKSP* is shown in Fig. 66.

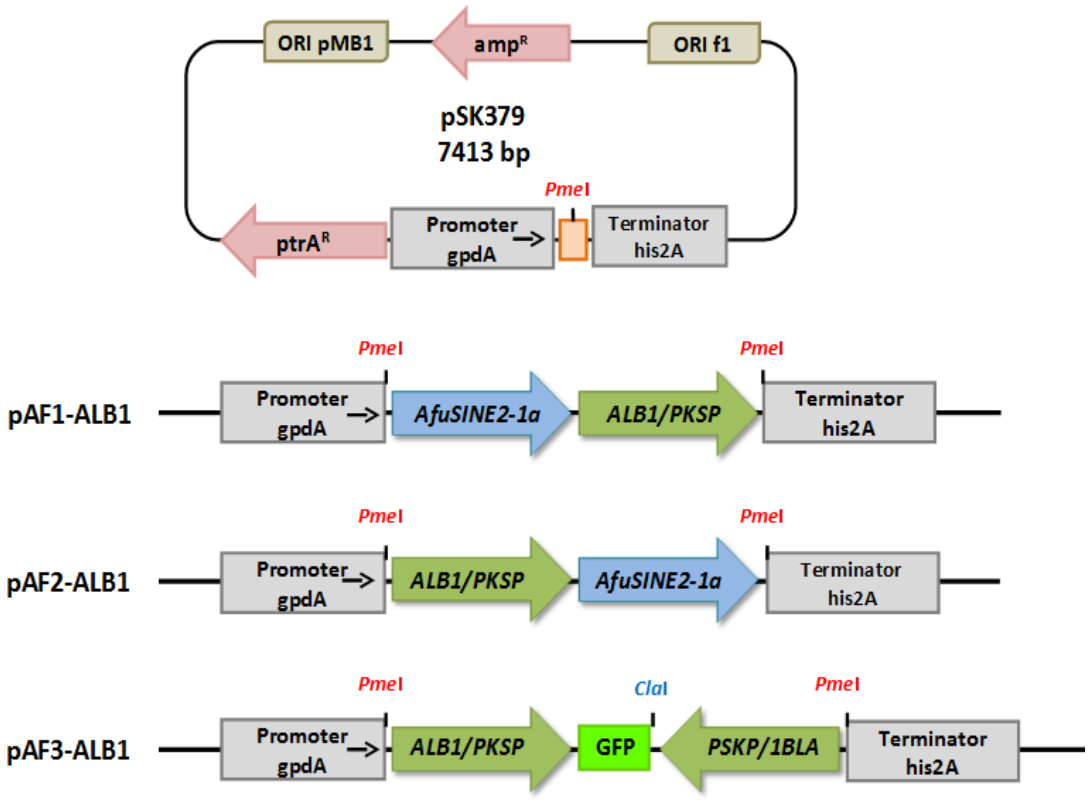


Fig. 66 Schematic representation of the *SINE*-based silencing vector constructs for silencing of *ALB1/PKSP*. **pAF1-ALB1:** 5'-*AfuSINE2-1α::ALB1/PKSP*-3' fusion in sense orientation; **pAF2-ALB1:** 5'-*ALB1/PKSP::AfuSINE2-1α*-3' fusion in sense orientation; **pAF3-ALB1**: *ALB1/PKSP* designed with inverted repeats of 500 bp of *AfuSINE2-1α-ALB1* separated by a 250 bp spacer of the green fluorescent protein (GFP). All constructs were modified using integrative *A. fumigatus* expression plasmid carrying *pgpdA-his2At* cassette and *ptrA,* pSK379, as a background plasmid vector. Constructs are not drawn to scale.

7. Silencing of *A. fumigatus* Af293 with gene silencing vectors

7.1 Silencing of *ALB1/PKSP* encoding gene from *A. fumigatus* Af293 by mycovirus-based vector

7.1.1 *In vitro* translation of pAfuTmV-1-RNA2 and pAfuTmV-1-NgoMIV vectors

The translation potential of AfuTmV-1 derived silencing vectors (pAfuTmV-1-RNA2 and pAfuTmV-1-NgoMIV) were investigated by *in vitro* translation in the TNT® coupled rabbit reticulocyte lysate system. A prominent translation product at 76 kDa corresponding to the dsRNA2 cDNA clone of pAfuTmV-1-RNA2 and translation product at 27.6 kDa corresponding the truncated dsRNA cDNA clone of pAfuTmV-1-NgoMIV were synthesized (Fig. 67), indicating that plasmid vectors were functionally transcribed and translated.

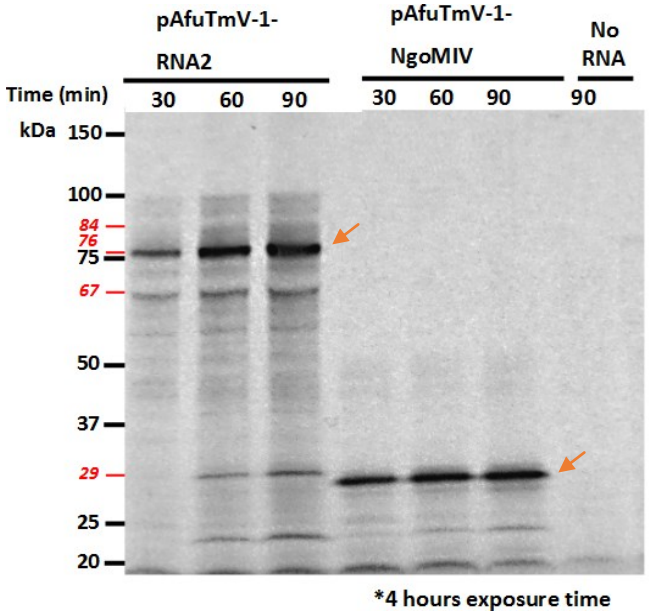


Fig. 67 Autoradiograms of *in vitro* translation products of AfuTmV-1 derived silencing vectors. Total *in vitro* translation products from pAfuTmV-1-RNA2 and pAfuTmV-1-NgoMIV were analyzed on NuPAGE® Novex® 4-12% Bis-Tris protein gel (Invitrogen). A time course over a period of 90 min was investigated taking samples at 30, 60 and 90 min. A negative control sample (no RNA) was also included in the experiment. Precision plus protein™ all blue standard (BIO-RAD) were used as marker proteins. Films were exposed in the dark for 4 h. Arrows indicate major translation products from each sample.

7.1.2 Silencing efficiency of AfuTmV-1-based silencing vector

The study was aimed at modification of an infectious cDNA clone derived from the AfuTmV-1 mycovirus to facilitate the insertion of non-viral fragments for RNA silencing study in fungi. Since several fungi including *A. fumigatus* have been reported to possess the RNA-mediated

defense response against infecting viruses (Hammond *et al*., 2008b), it is possible that introducing virus vectors carrying endogenous fragment derived from host gene would trigger silencing of both the viral genome and host genes that are partially homologous to viral genomic sequences. Thus, mycovirus genome can thus be potentially exploited as a source of RNA silencing.

Both dsRNA virus constructs, **AfuTmV-1-dsRNA2ΔORF** and **AfuTmV-1-dsRNA2::*ALB1***, were individually introduced into AfuTmV-1-infected *A. fumigatus* Af293 protoplasts. Analysis of dsRNAs from transfected protoplasts showed that none of the modified virus fragments were detectable by gel electrophoretic analysis or by RT-PCR amplification of predictable amplicons from pooled regenerants. These results suggest replication incompetence of both potential vectors. However several regenerants lost conidial pigmentation either partially or completely following transfection and these need to be investigated further especially to exclude the possibility of spontaneous non-pigmented mutants.

A mean of *ca*. 110 regenerants were regained from a single reaction (calculated from each serial dilution of 200 µl sample/plate; 10^{-3}-10^{-5}, x3 replicates). The absence or a reduction of *ALB1/PKSP* expression was found among the AfuTmV-1-dsRNA2::*ALB1* transfectants which included white colonies (designated as AfM-W_N; N= 4) indicating a possibly complete silenced phenotype, or light green colonies (designated as AfM-LG_N; N= 32) possibly indicating an intermediate silenced phenotype (Fig. 68a). The results suggested that the endogenous *ALB1/PKSP* gene was partially or completely silenced. However following a second passage of these regenerants in both solid and liquid ACM medium all of the potential silenced transfectants reverted to the wild type, dark green phenotype (Fig. 68b-g) perhaps because of transient silencing and an unstable phenotype. The identities of the transfectants and the parental strain were confirmed following successful amplification of the internal transcribed spacer (ITS) region and the dicer gene.

The instability of the silenced phenotype might be explained by the vector losing replication competence, due to deletion of the insert fragments. Nevertheless, it cannot be excluded that the silenced phenotypes observed could just be artifactual.

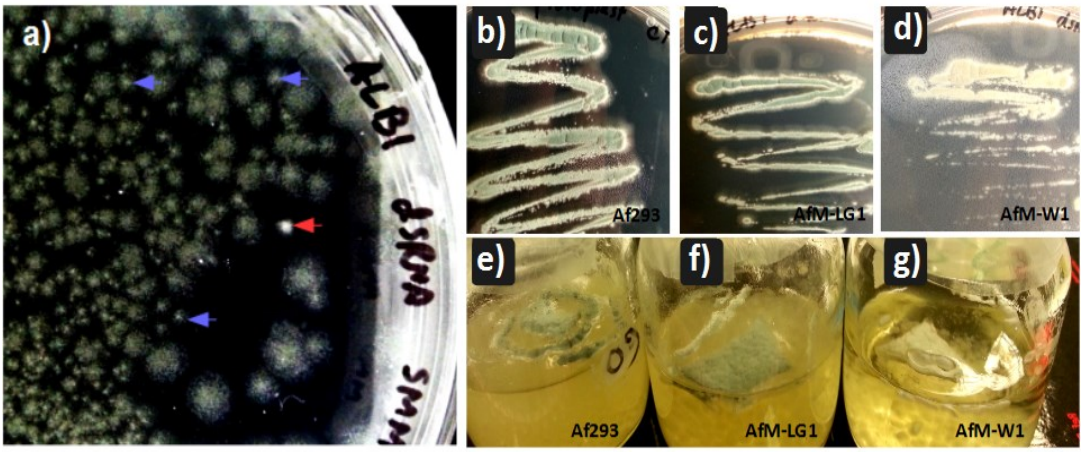


Fig. 68 Phenotype of transfectants obtained after *ALB1/PKSP* silencing. Regeneration plate (at 10^{-2} serial dilution of sample) of the transfectants transfected with AfuTmV-1-dsRNA2::*ALB* showing white colonies (arrowed in red), transfectants transfected with AfuTmV-1-dsRNA2::*ALB1* showing light green colonies (arrowed in blue) and non-transfectected colonies (dark green) (a). Colony morphology of the wild-type *A. fumigatus* Af293 (green; b); intermediate phenotype (AfM-LG1; light green; c); and complete phenotype (AfM-W1; white; d) on ACM regeneration plates. Mycelial mass of the wild-type *A. fumigatus* Af293 (green; e); intermediate phenotype (AfM-LG1; light green; f); and complete phenotype (AfM-W1; white; g) in ACM liquid medium.

7.2 Silencing of *ALB1/PKSP* encoding gene from *A. fumigatus* Af293 by a transcriptional fusion *SINE*-based vector

Wild type *A. fumigatus* Af293 protoplasts (virus-infected strain) were transformed with a control integrative plasmid pSK379 and three types of integrative silencing plasmids containing a PT resistance gene (*ptrA*). PT-resistant transformants were successfully regenerated on selection medium containing 500 ng/ml of the antibiotic. Several untransformed colonies, which were slow growing and small as compared to transformants (which were larger, grew faster and produced spores) were found in both control experiments (no plasmid) and experimental samples (with plasmid) (Fig. 69a). Additionally, pUC18 plasmid containing β-galactosidase gene was introduced into *A. fumigatus* Af293 protoplasts and transformants were subsequently regenerated on Amp^R-X-gal selection plates. All regenerated pUC18-incorporated transformant colonies turned blue, indicating development of a successful and high efficiency transformation procedure (Fig. 69b).

Transformant selection using PT was unreliable particularly for the *ALB1/PKSP* hpRNA-expressing plasmid where few apparently silenced lines were recovered. Thereafter ampicillin resistance was used to identify transformed colonies. No colonies grew on negative control plates but some amp-resistant, pSK379-transformed colonies were obtained which were

pigmented with a dark green colony colour as compared to the Af293 wild type. Several regenerants transformed with pAF1-ALB1, pAF2-ALB1 and pAF3-ALB1 were also recovered on following ampicillin selection. However, none of the transformants transfected with either *SINE*-based silencing vectors pAF1-ALB1 or pAF2-ALB1 demonstrated any alterations in colony pigmentation or morphology, indicating no silencing effect. However colonies regenerated from pAF3-ALB1-transformed protoplasts demonstrated a range of pigmentation types including dark green, light green and some white types, potentially indicating some silencing (Fig. 69c). Representative colonies of these variously pigmented isolates were grown on ACM agar plates for further investigation.

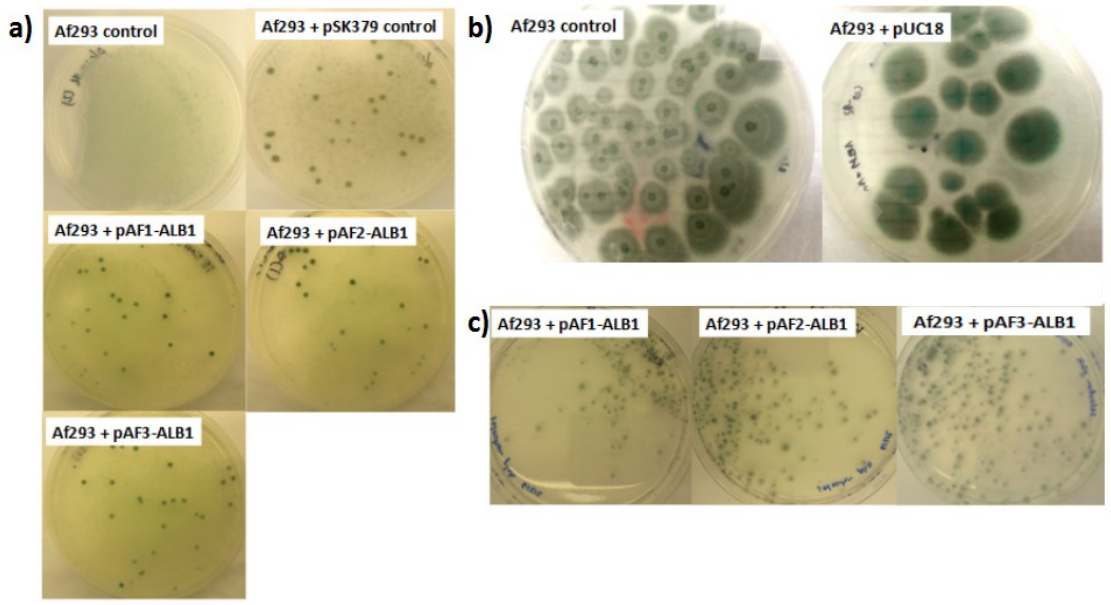


Fig. 69 Regeneration plates of *A. fumigatus* transformants containing various types of silencing vectors. Pyrithiamine (PT)-resistant transformants regenerated on selection medium containing 500 ng/ml of PT which were transformed with pAF1-ALB1, pAF2-ALB1, pAF3-ALB1 or pSK379 control plasmid (a). Ampicilin (Amp)- resistant transformatants regenerated on selection medium containing 0.1 mg/ml of Amp and X-gal solution which were transformed with pUC18. Non-transformed *A. fumigatus* was used as a control (b). Ampicilin (Amp)- resistant transformatants regenerated on selection medium containing 0.1 mg/ml of Amp which were transformed with pAF1-ALB1, pAF2-ALB1, pAF3-ALB1 or pSK379 control plasmid (c).

Following this passage only *A. fumigatus* regenerants transformed with pAF3-ALB1 produced an altered pigmentation phenotype as compared to the *A. fumigatus* Af293 parental strain with some colonies appearing a lighter green or white in colour (Fig. 70). Regenerants produced following transformation with the pSK379 control plasmid and the two *SINE*-based silencing plasmid constructs pAF1-ALB1 and pAF2-ALB1 were identical in colour to the *A. fumigatus* Af293 parental strain. These results suggest that the *SINE*-based constructs

investigated here do not induce silencing in *A. fumigatus* but that pAF3-ALB1, which contains an *ALB1/PKSP*-based hairpin structure, elicited a mild silencing effect. However further passaging of these isolates revealed that the silencing effect observed with the *ALB1/PKSP* construct was transient

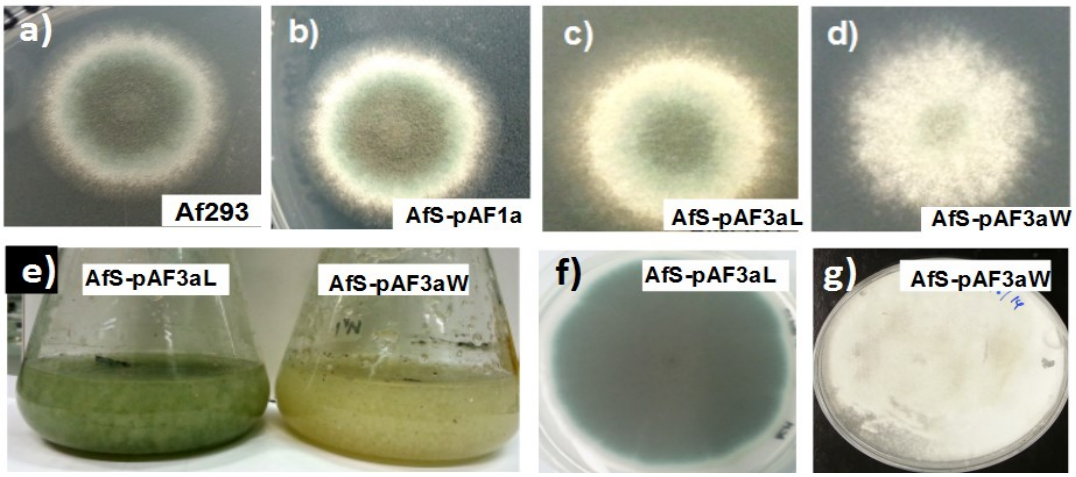


Fig. 70 Phenotype of transformants regenerated after *ALB1/PKSP* silencing. Colony morphology of the wild-type *A. fumigatus* Af293 (green; a); transformant (AfS-pAF1a) containing pAF1-ALB1 (green; b); and two transformants (AfS-pAF3aL and AfS-pAF3aW) containing pAF3-ALB1 (light green; c; white; d). Mycelial mass of the AfS-pAF3aL and AfS-pAF3aW) transformants containing pAF3-ALB1 in the ACM liquid medium (e). Colony of the AfS-pAF3aL and AfS-pAF3aW transformants containing pAF3-ALB1 on the ACM agar plate (f and g).

7.2.1 Investigation on integration of the plasmid in the transformed lines

Mycelia from the colonies transformed with pSK379, pAF1-ALB1, pAF2-ALB1, pAF3-ALB1 and non-transformed wild-type strain were further analyzed by southern blot hybridisation to examine whether the introduced plasmids integrated into the fungal genome. Integrity of the plasmid constructs was confirmed by PCR using primer pairs designed to amplify the insert fragment within *SINE*-based silencing plasmids. Transformants containing pSK379, pAF1-ALB1, pAF2-ALB1 and pAF3-ALB1 designated as AfS-pSKa, AfS-pAF1a, AfS-pAF2a and AfS-pAF3aW were selected for the analyses.

Southern blot analysis using a PCR probe specific to the pSK379-*ptrA* fragment demonstrated integration event of the silencing plasmid since hybridization signals corresponding to the *ptrA* gene were detected (Fig. 71a). It can be observed that all transformants showed the same hybridization pattern, suggesting that *SINE*-based silencing plasmids integrated into the genome. This integration event possibly result from the existence of a highly recombinogenic chromosomal site which could offer a preferential site for integration of the plasmid on the fungal genome (Herrera-Estrella *et al.*, 1990). However, determination of the sequence of the

junction between integrated plasmid and fungal genome would provide the clarification of mechanism associated with plasmid integration. Furthermore, the presence of multiple signals on the samples could not be detected, indicating integration of a single copy of *SINE*-based silencing plasmid into the fungal genome. In addition, probing of the digested transformant DNA with *ptrA* probe did not give different sized hybridizing fragments for all transformed lines tested, suggesting no insertion at different genomic sites of each transformant. No hybridization signal was detected in the non-transformed line.

Through PCR analysis, a fragment covering *pgpdA*-insert-*his2At* cassette of *ca*. 1028 bp from pAF1-ALB1 and pAF2-ALB1, and *ca*. 1359 bp from pAF3-ALB1 were amplified, suggesting successful integration of the silencing vectors into the fungal genome and indicating that endogenous *AfuSINE2-1a-ALB1/PKSP* gene was intact in the transgenic lines (Fig. 71b). Wild-type strain showed no amplification of the product from each primer pair. Taken together the results from southern blot and PCR analysis, it can be inferred that transformation of *SINE*-based silencing plasmids into the fungal protoplasts was achieved and silencing plasmid vectors were able to integrate into the fungal genome, although *SINE*-based silencing plasmid constructs (pAF1-ALB1 and pAF2-ALB1) did not show silencing effect of the *ALB1/PKSP* target gene.

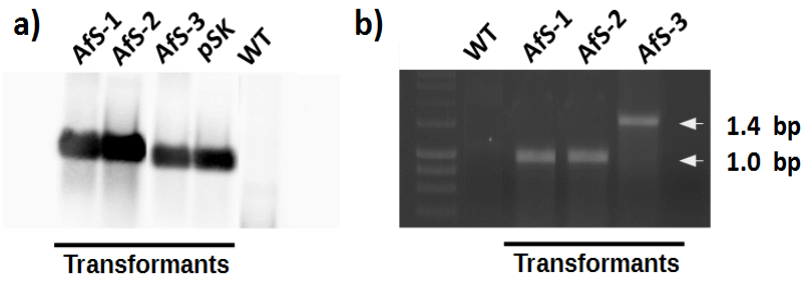


Fig. 71 The analysis on integration and integrity of the silencing plasmid construct into the *A. fumigatus* Af293 genome. Southern blot analysis of *A. fumigatus* transformants and *A. fumigatus* Af293 parental strain (a). Genomic DNA of protoplast transformants and the parental strain of *A. fumigatus* Af293 was digested with *Hind*III, and probed with the *ptrA* gene. PCR confirmation of silencing vector integrated into the genome of the transformants (b). WT, *A. fumigatus* Af293 wild-type; pSK, transformant AfS-pSKa containing pSK379 vector control; AfS1-3, transformant, AfS-pAF1a, AfS-pAF2a and AfS-pAF3aW containing pAF1-ALB1, pAF2-ALB1 pAF3-ALB1, respectively.

Discussion (Part IV)

1) Mycovirus-based silencing vectors

This investigation indicates that AfuTmV-1 and other tetraviruses possess obvious potential for development as tools for gene expression and gene silencing in fungi. AfuTmV-1 is unique in that is infectious both as a purified unconventional virus and as a dsRNA. Furthermore, AfuTmV-1 is unencapsidated and has a small genome (1.2 to 2.4 kbp in size) eliminating any constraints of inserting novel nucleic acid sequences. Additionally mycoviruses naturally initiate RNA silencing and are targeted by their fungal hosts as a defence response. According to this notion RNA silencing would be activated in transfected fungi when fungal RNAs or transgenes are fused as part of the viral genome. Additionally AfuTmV-1 infection in *A. fumigatus* is essentially cryptic, a useful feature in that alterations to phenotype caused by inserted genetic information in any virus-based constructs would not be confused with symptoms elicited by the virus itself.

In this investigation an AfuTmV-1-based silencing vector was constructed and tested. An *A. fumigatus* Af293 endogenous gene, *ALB1/PKSP*, which has been reported to be successfully silenced previously (Mouyna *et al*., 2004; Henry *et al*., 2007), was selected for the study. Introduction of the engineered AfuTmV-1-dsRNA2::*ALB1* caused loss of conidial pigmentation in the fungal host but to different levels. Silencing was demonstrated by the generation of white colonies in completely silenced types and light green colonies in the intermediate silenced phenotypes where the silencing event was transient. However, transfection occurred in at a very low frequency with only 4 fully silenced isogenic lines and 32 partially silenced isogenic lines being generated. Fully silenced (white) and partially silenced (light green) transfectants showed recovery of the wild type phenotype (dark green) in some lines after second subculturing, suggesting transient silencing and an unstable phenotype. These results suggest that the virus vector lost replication capacity or the insert was unstable in the construct. Since fragments of the virus vector sequence were not found following RT-PCR amplification this suggests that the virus vectors might not be replication competent. Nevertheless it could be possible that silenced phenotypes observed in this experiment could only be a transient silencing resulting from silencing mechanisms triggered by non-replicated dsRNAs. This needs to be confirmed and further investigated. A possible experiment could be performed by introducing the AfuTmV-1-dsRNA2::*ALB1* vector into the virus-free protoplasts

to confirm that previously observed silencing effect was not triggered by non-replicated dsRNAs.

If the virus contributes to the silencing effect, it is unclear how a range of silencing events occurred when virus-based silencing vector or dsRNA fragments were used. One explanation for this phenomenon could be the difference in replication rate between the recombinant virus vector which leads to uneven distribution, variation in silencing level and silencing stability in fungal cells. The ability of the host to mediate RNA silencing mechanisms could also be considered as a factor for variation in silencing efficacy which has a potential effect on degradation of the mRNA target. However, the presence of partially silenced and fully silenced phenotypes among isogenic lines has been observed in other microorganisms, for instance *Cryptococcus neoformans* (Liu *et al*., 2002); *Neurospora crassa* (Goldoni *et al*., 2004); *Magnaporthe oryzae* (Nakayashiki *et al*., 2005a); *Aspergillus fumigatus* (Mouyna *et al*., 2004; Henry *et al*., 2007) and *Trypanosoma brucei* (Ngo *et al*., 1998).

This study indicates the first demonstration of virus-induced gene silencing in the human pathogenic fungus *A. fumigatus* and shows that AfuTmV-1 might be used for functional genomics and studies of RNA silencing in *Aspergillus* sp. and other filamentous fungi since the vectors showed identifiable silencing effects when the host endogenous gene (*ALB1/PKSP)* fragment was inserted into the engineered viral genome and introduced into fungal protoplasts. However, the use of AfuTmV-1 as a virus-based silencing needs much refinement and development.

Even though most VIGS systems were developed for gene silencing in agriculturally important plants, application of VIGS for functional genomic studies in fungi has been currently investigated and several protocols for the effective transfection have been developed. In addition to the development of the AfuTmV-1-based vector, the *Barley stripe mosaic virus* (BSMV)-based vector expressing dsRNA fragments derived from the fungus *Puccinia triticina* genes has been successfully used to silence genes involved in fungal pathogenicity (Panwar *et al*., 2013). Mascia *et al*. (2014) reported successful expression of RNA silencing in the fungus *Collectotrichum acutatum* (strain C71) using a *Tobacco mosaic virus* (TMV)-based vector carrying a gene for the expression of GFP by direct transfection to the transgenic fungal protoplasts. The result showed stable silencing of the GFP in the fungal transformants expressing GFP derived from C71 wild type of up to six subcultures. This intriguing finding of utilizing a TMV-based vector for gene silencing in *C. acutatum* has offers strong evidence that this approach could be exploited as a reverse genetics tool to transiently silence specific genes

without complex transformation technology. In addition, stability of the modified TMV-based vector has revealed an advantage of using VIGS strategy to substitute the conventional gene silencing approaches. However, instability of the chimeric virus vector resulting from antiviral defense response of the fungal host was described by Suzuki *et al*., (2000) which has become an obstacle that hampers the use of recombinant virus as an expression vector and silencing vectors to induce gene silencing in fungi. Besides the stability of the silencing vector, replication competency of the virus-based vector is one of the important factors to certify silencing efficiency.

In contrast to other virus-induced gene silencing vectors previously described for gene silencing in fungi the novelty of the silencing vector developed in this study is the use of mycovirus genome which naturally infects fungi as a background for vector construction instead of using plant or animal virus genomes. AfuTmV-1 has co-evolutionally evolved in *A. fumigatus* without causing any harm to the fungus. The symbiotic relationship has a mutual benefit to virus survival by providing genetic machinery necessary for virus replication, and the fungal host likewise may receive some benefit from the virus. Therefore, AfuTmV-1 may have the ability to systemically infect and replicate in host cells by the viral RdRP and generate dsRNA during the replication cycle. This is a key factor for adopting the virus as a gene silencing vector. Since AfuTmV-1 infection is stably maintained in the host where there is little or no erosion to its genetic elements, this indicates that the virus is replication competent and in this regard, virus-based silencing vectors were developed to imitate a typical virus genome structure (as described previously). In this study, the development of mycoviruses as silencing vectors offers some advantages over other virus silencing vectors. Some conventional virus vectors are replication defective which leads to transient silencing because the virus vectors are usually diluted and may be lost during host cell division whereas mycovirus vectors, which are able to spread through fungal mycelial network by viral replication, are suitable for long-term silencing. In addition, virus vectors should just retain genomic regions important for its replication such as 5'- and 3'- UTRs for RdRP recognition or N-terminal portion (codon 1 to 24) of the 5' proximal ORF upstream and downstream of the start codon as reported for the hypovirus expression vector where these regions are indispensable for virus replication (Suzuki *et al*., 2000). In this study, dsRNA2 was selected for cDNA clone synthesis because its ORF2 encodes a putative protein with unknown function that seems to be less necessary for virus replication as compared to the others. In addition, accumulation of dsRNA2 in the fungal cells is the highest, representing high replication rate of the element. For virus-based vector

construction in this approach, the majority of the coding region of dsRNA2 was internally deleted and replaced by a target gene.

Introduction of *in vitro* synthesized dsRNAs, siRNAs or dsRNA-/shRNA-expression vector cassettes have been commonly used for gene silencing in fungi. In *A. fumigatus*, integrative vector constructs (Mouyna *et al*., 2004; Henry *et al*., 2007) and self-replicating episomal vector constructs (Khalaj *et al*., 2007) which express shRNA structures controlled by specific inducible promoter have been developed for silencing *ALB1* and *FKS1* genes. However, instability of the vectors and transient silencing were common eventually resulting in restoration of wild type gene expression (Mouyna *et al*., 2004; Henry *et al*., 2007). Another limitation that hinders the use of integrative/shRNA-expressing plasmid vectors in fungi is the occurrence of genome rearrangement. The event usually occurs after fungal transformation with silencing plasmids which affect silencing efficiency (Oliveira *et al.*, 2008). Therefore, the use of mycoviruses could overcome this difficulty since they constitutively replicate in host cells over many generations. As a consequence, silencing activity is continuous as the virus-based vector is replication competent, resulting in a long-term down-regulation of potential target genes.

Since RNA silencing is a natural defense response of fungi against viral infection, another consideration which influences silencing efficiency is that the fungal host should possess the basic components of the RNA silencing pathway. Three key silencing components were identified in *A. fumigatus* using the TIGR *A. fumigatus* database (http://tigrblast.tigr.org/ufmg/) as being homologs of *Neurospora crassa* genes, QDE1, QDE2 and QDE3, which suggests that RNA silencing is functional in *A. fumigatus* (Mouyna *et al*., 2007). Evidence demonstrating the presence of mycovirus-derived siRNA supports the presence of silencing mechanisms in *Aspergillus* that mycoviruses are targeted for degradation by an *Aspergillus* silencing pathway (Hammond *et al.*, 2008b). Thus, this significance has been exploited to use mycovirus genomes as a source for RNA silencing vector development. Although the virus genomic components are targeted *via* silencing pathways, the virus titre remains stable in the fungal host. It could be possible that viruses are targeted when the accumulation of dsRNA is beyond a threshold level. The ability of AfuTmV-1 to overcome antiviral host defense responses *via* RNA silencing is not well-understood. It is possible that the virus might possess a functional gene encoding silencing suppressors or exploit alternative pathways to suppress host silencing which might operate in a novel fashion. Avoiding sequences with strong suppressor activity in any virus-based vector is paramount.

AfuTmV-1 appears to be a good candidate as an alternative source for the development of virus-induced gene silencing vectors. However there maybe unforeseen drawbacks in utilisation of the virus as a vector. It is assumed but not proven that AfuTmV-1 replication requires the cooperation of all four proteins encoded by the four genomic dsRNAs. It is therefore unlikely that for instance a recombinant virus derived from dsRNA2 alone would replicate and be active in silencing. In this study the AfuTmV-1 dsRNA2-derived silencing vector was transfected into *A. fumigatu*s Af293 which is residually infected with AfuTmV-1. As the consequence, replication of the virus vector should be activated and assisted by the resident virus. Of especial interest in such a system is the essential requirement for the viral RdRP. In order to ascertain the minimum requirements necessary to replicate AfuTmV-1 it might be worthwhile to concentrate on developing dsRNA1 based vectors for use in combination with wild type dsRNA1 alone since it encodes the essential RdRP RNA transcribing enzyme. Identification of essential and dispensable virus-encoded replication elements would also help to clarifying function of each protein that relates to virus replication. This information could facilitate the establishment of new strategies in order to develop silencing vectors.

Engineering the virus to drive expression of a silencing effector such as an inverted repeat sequence or sense/antisense sequences homologous to the target gene would eliminate the hurdle of RdRP dependency. Constitutive promoters such as the *Cauliflower mosaic virus* 35S promoter could be used to generate full-length RNA copies of an insert. Additionally 5'-capping of the genomic viral RNAs could enhance replication competency of the vectors.

2) *SINE*-based silencing vectors

The study involved the construction of silencing vectors based on *SINEs* identified in the *A. fumigatus* Af293 genome by transcriptional fusion to an endogenous *ALB1/PKSP* gene. Following transformation of wild type *A. fumigatus* Af293 protoplasts (AfuTmV-infected strain) with *SINE*-based silencing plasmids no phenotypic alterations to the regenerants were found following transformation with two *SINE*-based silencing plasmid constructs (pAF1-ALB1; 5'-*AfuSINE2-1a::ALB1/PKSP*-3' and pAF2-ALB1: 5'-*ALB1/PKSP::AfuSINE2-1a*-3') suggesting that *ALB1/PKSP* gene silencing was not initiated. The finding also indicated no spread of silencing from the *AfuSINE2-1a* to the endogenous gene. On the other hand, an inverted repeat construct of the pAF3-ALB1 expressing hairpin structure of *ALB1/PKSP* fragment showed some silencing effect on the *ALB1/PKSP* expression.

Amplification of ^{p}*gpdA*-insert-*his2A*^{t} cassette and southern blot analysis demonstrated successful integration events for all silencing constructs. Thus failures to silence the target gene with either construct above are unlikely to be a loss of integrity of the vector after transformation or a failure to integrate into the fungal genome. A possible explanation for these results might concern transcription activity and silencing of *AfuSINE2-1a*. Not withstanding these observations Vetukuri *et al*. (2011a) have demonstrated successful silencing of a *PiAvr3a* effector gene in *Phytophthora infestans* using transcriptional fusions of *infSINEm* in a similar experimental approach to the strategy used in this investigation. The *infSINEm* element was identified in the *P. infestans* genome as being an active *SINE* and present with a genomic copy number of 14 (Whisson *et al*., 2005) dissimilar to *AfuSINE2-1a* from *A. fumigatus* Af293 whilst weakly transcriptionally active was only present with a genomic copy number of 4. As a consequence of low copy number and low expression of *AfuSINE2-1a* this could result in negligible (if any) silencing in the fungal cells, resulting in non-silencing of the element and the adjacent gene fragment. It is still conceivable that the differences in silencing efficiency between *A. fumigatus* and *P. infestans SINEs* reflect basic differences in the silencing mechanisms used in both fungi. The *P. infestans* silencing components are composed of single copies of the Dicer and RdRP genes and four Argonaute genes (Vetukuri *et al*., 2011b). In contrast *A. fumigatus* contains two Dicers (*dclA* and *dclB*) and two Argonautes (*ppdA*/*rsdA* and *ppdB*) but lacks an RdRP (*RrpC*) (Hammond and Keller, 2005; Hammond *et al*., 2008b). Importantly, if *RrpC* plays a crucial role in silencing *A. fumigatus SINEs* (as proposed in schematic diagram Fig. 72) the absence of *RrpC* in the genome of the fungus activity might be responsible for unsuccessful silencing of transcriptionally fused *AfuSINE2-1a* constructs. Thus differences in the silencing machinery might reflect diverse mechanisms for complete gene silencing, including silencing of the retrotransposons. Additionally the integration site of the plasmid construct may be important affecting the activity of the transcriptional fused *SINE*. For instance integration of the vector in a region with low transcriptional activities e. g. in heterochromatin regions, might result in low silencing effect. Also, genome instability and genome rearrangement events in the *A. fumigatus* genome could possibly be responsible for low silencing efficiency of the integrated plasmid vectors. Variable silencing activity of *SINEs* in *A. fumigatus* requires further investigation.

A number of *SINEs* of different origin, sequence, transcription activity and copy number have been identified in the *A. fumigatus* genome in Chapter 3 (part III) in this investigation. It is possible that individual *AfuSINEs* possess subtle differences in terms of transcription and retrotransposition activity, resulting in mutability of the fungal hosts silencing pathways to

silence these parasitic elements. In this regard, other active *AfuSINEs* could be used to construct silencing plasmids and their silencing frequency determined in a similar fashion. However much remains to be done to identify suitable *SINE* candidates for development of *SINE*-based silencing vectors as potential reverse genetic tools for functional genomics in fungi.

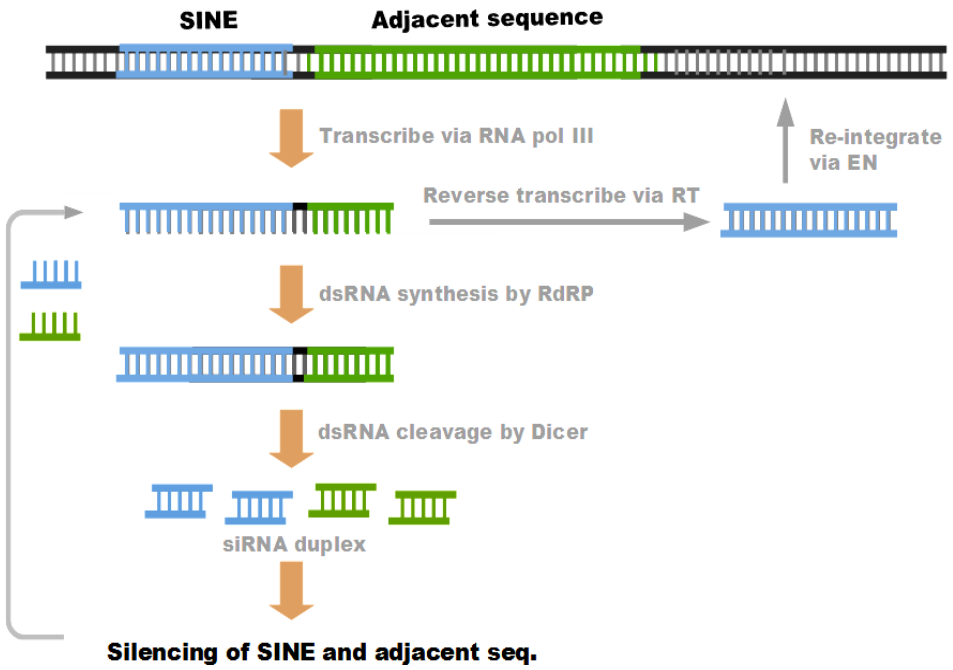


Fig. 72 Schematic diagram of the proposed *SINE* silencing pathway.

Conclusion (Part IV)

1) Gene silencing has been verified to be a valuable approach for determining function of genes in *Aspergillus* sp. and many fungi. Here, development of the first mycovirus-based silencing vector was described in the *A. fumigatus* by transfecting recombinant virus vector carrying target gene into fungal protoplasts to trigger gene silencing. An unencapsidated Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1), which naturally infects *A. fumigatus* Af293, was manipulated into virus-derived vectors and used as a tool to deliver fragment of endogenous gene (*ALB1/PKSP*) for silencing. The introduction of the engineered AfuTmV-1-dsRNA2::*ALB1* caused loss of conidial pigmentation to the fungal host in different levels in some lines. The silencing effect occurred in a very low frequency and demonstrated both white colonies of the complete silenced phenotype and light green colonies of the intermediate silenced phenotype where the silencing event was transient.

2) The construction of silencing vectors based on short interspersed nuclear element (*SINE*) identified from the *A. fumigatus* Af293 genome by transcriptional fusion of the element to an endogenous *ALB1/PKSP* gene was successfully established. Changing in colony colour was not observed in the transgenic lines individually transformed with two *SINE*-based silencing plasmid constructs (pAF1-ALB1 and pAF2-ALB1), suggesting that silencing of the *ALB1/PKSP* gene was not initiated by transcriptional fusion of the element. However, the inverted repeat construct of the pAF3-ALB1 expressing hairpin structure of *ALB1/PKSP* fragment showed some silencing effect on the *ALB1/PKSP* expression. General mechanisms of RNA silencing induced from mycovirus-based and *SINE*-based constructs are proposed in Fig. 73.

3) In conclusion, mycovirus-based vector seems to be applicable to be used as a silencing tool in order to study gene function in *A. fumigatus* as the approach requires only direct transfection and less time in preparation of the constructs. *SINE*-based silencing construct does not seem to be working in this fungus since the organism lacks some important silencing components such as the *RrpC*. Low copy number of *AfuSINE* in the genome cannot be excluded as one of the limitations to exploiting the elements as a silencing vector. However, in-depth study of each silencing constructs remains to be investigated in many aspects. Once promising evidence demonstrating a potential of the developing approaches is elucidated, the mycovirus-based silencing vector and *SINE*-based vector systems could be used and would provide a high-throughput tool for functional genomic study and elucidation of RNA-silencing mechanisms in many human pathogenic fungal species.

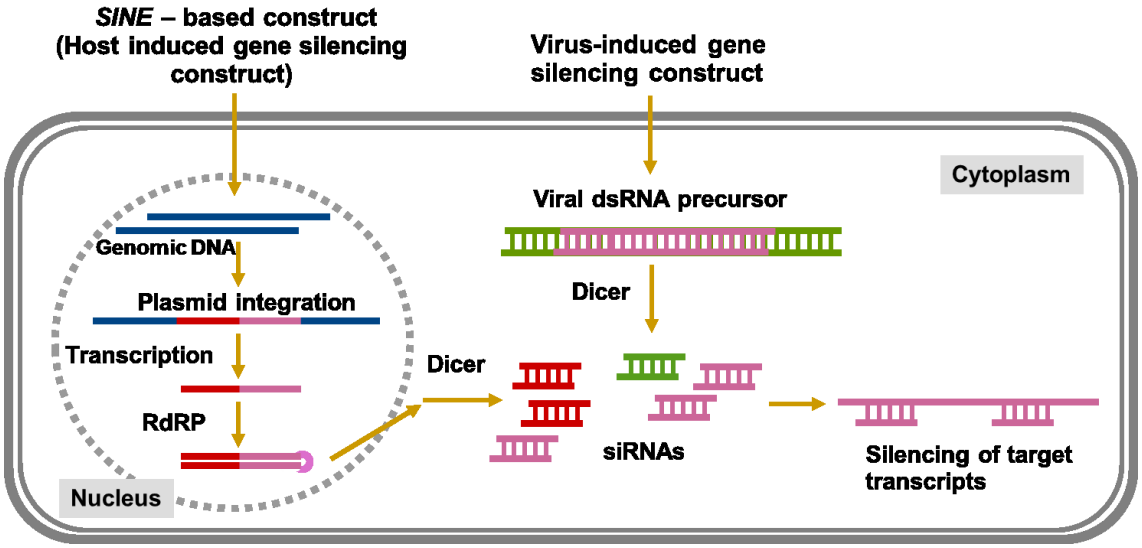


Fig. 73 General mechanisms of RNA silencing via mycovirus-based constructs and *SINE*-based constructs.

Chapter 4

Conclusion of the Study

Chapter 4: Conclusion of the Study

1) A novel double-stranded RNA (dsRNA) mycovirus named Aspergillus fumigatus tetramycovirus-1 (AfuTmV-1) from the human pathogenic fungus *Aspergillus fumigatus* was fully characterized. The AfuTmV-1 genome consists of four dsRNAs, the largest of which encodes an RNA dependent RNA polymerase (RdRP) containing a unique GDNQ motif. The third largest dsRNA encodes an S-adenosyl methionine-dependent methyltransferase (SAM) capping enzyme and the smallest dsRNA a PA-rich protein which apparently coats but does not encapsidate the viral genome. This virus shows several unique properties that it could represent a prototype member of a new virus family.

2) Novel families of short interspersed nuclear elements (*SINEs*) were identified in the human pathogenic fungus *Aspergillus fumigatus* clinical isolate Af293. The elements were categorized into tRNA-related and 5S rRNA-related *SINE*-like families which were nominated as *AfuSINE2* and *AfuSINE3*, respectively, and they are found in a low copy number. Eight predicted *AfuSINE2s* ranging in size from 140-493 bp show target site duplication of short direct repeat sequences (4-14 bp), extended tRNA-unrelated region, typical feature of RNA polymerase III promoter sequences and are found to originate from different tRNA origins. One of the elements in this family shows similarity at the 3'-end to the *I-4_AO* and *I-1_AF LINE*-like element. A family of *AfuSINE3s* ranging in size from 259-343 bp shows 5'-5S rRNA-related regions with 50-65% and 60-75% similarity to the *A. fumigatus* 5S rRNAs and *SINE3-1_AO*, respectively. RT-PCR analysis revealed that 6 out of 13 *AfuSINEs* are actively transcribed. In addition, investigations on *AfuSINE* distribution revealed that the elements are enriched in the pericentromeric and subtelomeric regions.

3) Silencing vectors based on mycovirus and *SINE* sequences were constructed. A recombinant virus vector derived from AfuTmV-1 was developed to trigger RNA silencing mechanism in the fungal cells by using direct transfection method which demonstrated transient silencing of an *ALB1/PKSP* gene whereas transcriptional fusion *SINE*-derived vector failed to silence the *ALB1/PKSP* gene in this study.

Chapter 5

Bibliography

Chapter 5: Bibliography

Abad, A., Molina, J. V. F., Bikandi, J., Ramırez, A., Margareto, J., Sendino, J., Hernando, F. L., Ponto´N, J., Garaizar, J. and Rementeria, A. (2010). What makes *Aspergillus fumigatus* a successful pathogen? Genes and molecules involved in invasive aspergillosis. *Revista Iberoamericana de Micologia*, **27**, 155–182.

Accotto, G.P., Brisco, M.J. and Hull R. (1987). *In vitro* translation of the double stranded RNA genome from beet cryptic virus 1. *Journal of General Virology*, **68**, 1417-1422.

Ah-Fong, A.M.V., Bormann-Chung, C. and Judelson, H.S. (2008). Optimization of transgene-mediated silencing in *Phytophthora infestans* and its association with small-interfering RNAs. *Fungal Genetics and Biology*, **45**, 1197-1205.

Ahlquist, P. (2006). Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nature Reviews Microbiology*, **4**, 371-382.

Akiyama, K., Katahami, H. and Takata, R. (2007). Mobilization of a retrotransposon in 5-azacytidine-treated fungus *Fusarium oxysporum*. *Plant Biotechnology*, **24(3)**, 345-348.

Amariglio, N. and Rechavi, G. (1993). Insertional mutagenesis by transposable element in the mammalian genome. *Environment and Molecular Mutagenesis*, **21**, 212-218.

Anagnostakis, S.L., Chen, B., Geletka, L.M. and Nuss, D.L. (1998). Hypovirus transmission to ascospore progeny by field-released transgenic hypovirulent strains of *Cryphonectria parasitica*. *Phytopathology*, **88**, 598-604.

Anbar, M., Bracha, R., Nuchamowitz, Y., Li, Y., Florentin, A. and Mirelman, D. (2005). Involvement of a short interspersed element in epigenetic transcriptional silencing of the amoebapore gene in *Entamoeba histolytica*. *Eukaryotic Cell*, **4**, 1775-1784.

Aoki, N., Moriyama, H., Kodama, M., Arie, T., Teraoka, T. and Fukuhara T. (2009). A novel mycovirus associated with four double-stranded RNAs affects host fungal growth in *Alternaria alternata*. *Virus Research*, **140**, 179-187.

Aravind, L., Watanabe, H., Lipman, D.J. and Koonin, E.V. (2000). Lineage-specific loss and divergence of functionally linked genes in eukaryotes. *Proceedings of the National Academy of Sciences, USA*, **97**, 11319-11324.

Arnaud, M.B., Chibucos, M.C., Costanzo, M.C., Crabtree, J., Inglis, D.O., Lotia, A., Orvis, J. Shah, P., Skrzypek, M.S., Binkley, G., Miyasato, S.R., Wortman, J.R. and Sherlock, G. (2010). The *Aspergillus* Genome Database, a curated comparative genomics resource for gene, protein and sequence information for the *Aspergillus* research community. *Nucleic Acids Research*, **38** (D420-7).

Balloy, V. and Chignard, M. (2009). The innate immune response *to Aspergillus fumigatus*. *Microbes and Infection*, **11**, 919-927.

Banks, G.T., Buck, K.W., Chain, E.B., Arbyshire, J.E., Himmelweit, F., Ratti, G., Sharpe, T.J., and Planterose, D.N. (1970). Antiviral activity of double stranded RNA from a virus isolated from *Aspergillus foetidus*. *Nature*, **227**, 505 -507.

Bao, W. and Jurka, J., (2011). *SINEs* from barley powdery mildew. *Repbase Reports*, **11 (9)**, 2582-2582.

Barciszewska, M.Z., Szymanski, M., Erdmann, V.A. and Barciszewski, J. (2001). Structure and functions of 5S rRNA. *Acta Biochimica Polonica*, **48 (1)**, 191-198.

Bejerano, G., Lowe, C.B., Ahituv, N., King, B., Siepel, A., Salama, S.R., Rubin, E.M., Kent, W.J. and Haussler, D. (2006). A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature*, **441(7089)**, 87-90.

Bergamini, G., Preiss, T. and Hentze, M.W. (2000). Picornavirus IRESes and the poly(A) tail jointly promote cap-independent translation in a mammalian cell-free system. *RNA*, **6**, 1781-1790.

Bhatti, F. (2011). Investigations on *Aspergillus fumigatus* double-stranded RNAs and their effects on the fungus. Ph. D. thesis, Imperial College.

Bhatti, F., Bignell, E. M. and Coutts R. H. A. (2011a). Complete nucleotide sequences of two dsRNAs associated with a new partitivirus infecting *Aspergillus fumigatus. Archives of Virology*, **156**, 1677-1680.

Bhatti, F., Jamal, A., Petrou, M. A., Bignell, E.M. and Coutts R.H.A. (2011b). The effects of dsRNA mycoviruses on growth and murine virulence of *Aspergillus fumigatus*. *Fungal Genetics and Biology*, **48**, 1071-1075.

Bhatti, F., Jamal, A., Bignell, E. M. Petrou, M. A. and Coutts R. H. A. (2012). Incidence of dsRNA mycoviruses in a collection of *Aspergillus fumigatus* isolates. *Mycopathologia*, **174**, 323-326.

Biswas, B., Adhya, S., Washart, P., Paul, B., Trostel, A.N., Powell, B., Carlton, R. and Merril, C. (2002). Bacteriophage therapy rescues mice bacteremic from a clinical isolate of vancomycin-resistant *Enterococcus faecium*. *Infection and Immunity*, **70**, 204-210.

Boeke, J. D. (1989). Transposable elements in *Saccharomyces cerevisiae*. In: Berg, D. E. and Howe, M. (eds) Mobile DNA. ASM, Washington DC, 335-374.

Bokar, J.A. and Rottman, F.M. (1999). Nucleoside methylation in eukaryotic mRNA: HeLa mRNA (N6-adenosine) methyltransferase. In. *S-Adenosylmethionine-dependent Methyl transferase: Structures and function*. (Cheng, X. and Blumental, R.M., eds.).World Scientific Publishing Company, Singapore, 227-253.

Borodulina, O.R. and Kramerov, D. A. (1999). Wide distribution of short interspersed elements among eukaryotic genomes. *FEBS Letters*, **457**, 409-413.

Borodulina, O.R. and Kramerov, D.A. (2005). PCR-based approach to SINE isolation: simple and complex *SINEs*. *Gene*, **349**, 197-205.

Boukhvalova, M.S., Prince, G.A. and Blanco, J.C. (2010). Inactivation of respiratory syncytial virus by zinc finger reactive compounds. *Virology Journal*, **7**, 20.

Bradbury, J. (2004). “My enemy's enemy is my friend.” Using phages to fight bacteria. *Lancet*, **363**, 624-625.

Bradshaw, V.A. and McEntee, K. (1989). DNA damage activates transcription and transposition of yeast *Ty* retrotransposons. *Molecular and General Genetics*, **218**, 465-474.

Britten, R.J. (1996). DNA sequence insertion and evolutionary variation in gene regulation. *Proceedings of the National Academy of Sciences*, USA, **93**, 9374-9377.

Brookfield, J.F., Badge, R.M. (1997). Population genetics models of transposable elements. *Genetica*, **100**, 281-294.

Brosius, J. (1999). RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene*, **238**, 115-134.

Bruenn, J.A. (1993). A closely related group of RNA-dependent RNA polymerases from double-stranded RNA viruses. *Nucleic Acids Research*, **21**, 5667-5669.

Buck, K.W., Girvan, R.F. and Ratti, G. (1973). Two serologically disticnt double-stranded ribonucleic acid viruses isolated from *Aspergillus niger. Biochemical Society Transaction*s, **1**, 1138-1140.

Buck, K.W. and Ratti, G. (1975). Biophysical and biochemical properties of two viruses isolated from *Aspergillus foetidus*. *Journal of General Virology*, **27**, 211-224.

Cai, G., Krychiw, J.F., Myers, K., Fry, W.E. and Hillman, B.I. (2013). A new virus from the plant pathogenic oomycete *Phytophthora infestans* with an 8 kb dsRNA genome: The sixth member of a proposed new virus genus. *Virology*, **435**, 341-349.

Campbell, C.K., Johnson, E.M., Philpot, C.M. and Warnock, D.W. (1996). Identification of pathogenic fungi. Public Health Laboratory Service, London, United Kingdom, p. 298.

Capy, P., Gasperi, G., Biemont, C. and Bazin, C. (2000). Stress and transposable elements: Co-evolution or useful parasites? *Heredity*, **85**, 101-106.

Caribe dos Santos, A.C., Sena, J.A.L., Santos, S.C., Dias, C.V., Pirovani, C.P., Pungarnik, C., Valle, R.R., Cascardo, J.C.M. and Vincentz, M. (2009). DsRNA-induced gene silencing in *Moniliophthora perniciosa*, the causal agent of witches' broom disease in cacao. *Fungal Genetics and Biology*, **46**, 825-836.

Castro, M., Kramer, K., Valdivia, L., Ortiz, S. and Castillo, A. (2003). A double-stranded RNA mycovirus confers hypovirulence-associated traits to *Botrytis cinerea*. *FEMS Microbiology Letters*, **228**, 87-91.

Cerqueira, G.C., Arnaud, M.B., Inglis, D.O., Skrzypek, M.S., Binkley, G., Simison, M., Miyasato, S.R., Binkley, J., Orvis, J., Shah, P., Wymore, F., Sherlock, G. and Wortman, J.R. (2013). The *Aspergillus* Genome Database: multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Research*, **42(Database issue)**, D705-D710.

Chen, B.J. and Lamb, R.A. (2008). Mechanisms for enveloped virus budding: can some viruses do without an ESCRT? *Virology*, **372**, 221-232.

Chiba, S., Salaipeth, L., Lin, Y.H., Sasaki, A., Kanematsu, S. and Suzuki, N. (2009). A novel bipartite double-stranded RNA mycovirus from the white root rot fungus *Rosellinia*

necatrix: molecular and biological characterization, taxonomic considerations, and potential for biological control. *Journal of Virology*, **83**, 12801-12812.

Chiba, S., Lin, Y.H., Kondo, H., Kenematsu, S. and Suzuki, N. (2013). Effects of defective interfering RNA on symptom induction by and replication of a novel partitivirus from a phytopathogenic fungus *Rosellinia necatrix. Journal of Virology*, **87(4)**, 2330-2341.

Choi, G.H. and Nuss, D.L. (1992). A viral gene confers hypovirulence associated traits to the chestnut blight fungus. *EMBO Journal*, **11**, 473-477.

Chu, Y.M., Jeon, J.J., Yea, S.J., Kim, Y.H., Yun, S.H., Lee, Y.W. and Kim, K.H. (2002). Double-stranded RNA mycovirus from *Fusarium graminearum*. *Applied and Environmental Microbiology*, **68**, 2529-2534.

Chu, Y.M., Lim, W.S., Yea, S.J., Cho, J.D., Lee, Y.W. and Kim. K.H. (2004). Complexity of dsRNA mycovirus isilated from *Fusarium graminearum*. *Virus Genes*, **28(1)**, 135-143.

Chun, S.J. and Lee, Y.H. (1997). Inheritance of dsRNAs in the rice blast fungus, *Magnaporthe grisea*. *FEMS Microbiology Letters*, **148(2)**, 159-162.

Clutterbuck, A., Kapitonov, V.V. and Jurka, J. (2007). Transposable elements and repeat-induced point mutation in *Aspergillus nidulans, Aspergillus fumigatus,* and *Aspergillus oryzae*. In: Goldman, G.H. and Osmani, S.A. (2007). The Aspergilli: Genomics, medical aspects, biotechnology, and research methods. *CRC Press*, **26**, 343-355.

Cogoni, C. And Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature*, **399**, 166-169.

Cogoni, C. (2001). Homology-dependent gene silencing mechanisms in fungi. *Annual Review of Microbiology*, **55**, 381-406.

Cole, T.E., Muller, B., Hong, Y., Brasier, C.M. and Buck, K.W. (1998). Complexity of virus-like double-stranded RNA elements in a diseased isolate of the Dutch elm disease fungus, *Ophiostoma novo-ulmi*. *Journal of Phytopathology*, **146**, 593–598.

Cost, G.J., Feng, Q., Jacquier, A. and Boeke, J.D. (2002). Human L1 element target-primed reverse transcription *in vitro*. *The EMBO Journal*, **21**, 5899-5910.

Coutts, R.H.A. and Livieratos, I. (2003). A rapid method for sequencing the 5' and 3'- termini of double-stranded RNA viral templates using RLM-RACE. *Journal of Phytopathology*, **151**, 525-527.

Dalton R.E., Podila, G.K., Flurkey, W.H. and Bozarth, R.F. (1985). *In vitro* translation of the major capsid polypeptide from *Ustilago maydis* virus strain P1. *Virus Research*, **3**, 153-163.

Dalzoto, P.R., Glienke-Blanco, C., Kava-Cordeiro, V., Ribeiro, J.Z., Kitajima, E.W. and Azevedo, J.L. (2006). Horizontal transfer and hypovirulence associated with double-stranded RNA in *Beauveria bassiana*. *Mycological Research*, **110**, 1475-1481.

Daniels, G.R. and Deininger, P.L. (1983). A second major class of *Alu* family repeated DNA sequences in a primate genome. *Nucleic Acids Research*, **11**, 7595-7610.

Darissa, O. Willingmann, P., Schafer, W. and Adam, G. (2011). A novel double-stranded RNA mycovirus from *Fusarium graminearum*: nucleic acid sequence and genomic structure, *Archieves of Virology*, **156**, 647-658.

Dawe, A.L. and Nuss, D.L. (2001). Hypoviruses and chestnut blight: exploiting viruses to understand and modulate fungal pathogenesis. *Annual Review of Genetics*, **35**, 1-29.

Decroly, E., Ferron, F., Lescar, J. and Canard, B. (2012). Conventional and unconventional mechanisms for capping viral mRNA. *Nature Reviews Microbiology*, **10**, 51-65.

Deininger, P.L. and Batzer, M.A. (2002). Mammalian retroelements. *Genomes Research*, **12**, 1455-1465.

Deininger, P.L., Moran, J.V, Batzer, M.A. and Kazazian, H.H. (2003). Mobile elements and mammalian genome evolution. *Current Opinion in Genetics & Development*, **13**, 651-658.

Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990). An attempt to unify the structure of polymerases. *Protein Engineering,* **3**, 461–467.

Deng, W., Muast, B.S., Nickle, D.C., Learn, G.H., Liu, Y., Heath, L., Pond, S.L.K. and Mullins, J.I. (2010). DIVEN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques*, **48(5)**, 405-408.

Deragon, M.J. and Zhang, X. (2006). Short interspersed elements (SINEs) in plants: origin, classification and, use as phylogenetic markers. *Systems Biology*, **6**, 949-956.

Dewannieux, M. and Heidmann, T. (2005). *LINEs*, *SINEs* and processed pseudogenes: parasitic strategies for genome modeling. *Cytogenetic and Genome Research*, **110**, 35-48.

Diaz-Ruiz, J. and Kaper, J. (1978). Isolation of viral double-stranded RNAs using a LiCl fractionation procedure. *Preparative Biochemistry and Biotechnology*, **8**, 1-17.

Dickinson, M.J. and Pryor, A.J. (1989). Encapsidated and unencapsidated double-stranded RNAs in flax rust, *Melampsora lini*. *Canadian Journal of Botany*, **67**, 1137-1142.

Dickinson, M. J., Zhang, R. and Pryor, A. (1993). Nucleotide sequence relationships of double-stranded RNAs in flax rust, *Melampsora lini*. *Current Genetics*, **24**, 428-432.

Ding, S.W. (2010). RNA-based antiviral immunity. *Nature Reviews Immunology*, **10**, 632-644.

Edmondson, S.P., Lang, D. and Gray, D.M. (1984). Evidence for sequence heterogeneity among the double-stranded RNA segments of *Penicillium chrysogenum* mycovirus. *Journal of General Virology*, **65**, 1591-1599.

Eickbush, T.H. and Furano, A.V. (2002). Fruit flies and humans respond differently to retrotransposons. *Current Opinion in Genetics and Development*, **12**, 669-674.

Fahima, T., Kazmierczak, P., Hansen, D. R., Pfeiffer, P. and van Alfen, N. K. (1993). Membrane-associated replication of an unencapsidated double-strand RNA of the fungus, *Cryphonectria parasitica*. *Virology*, **195**, 81-89.

Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C.J., Hofmann, K. and Bairoch, A. (2002). The PROSITE database, its status in 2002. *Nucleic Acids Research*, **30(1)**, 235-238.

Fantaccione, S., Russo, C., Palomba, P., Rienzo, M. and Pontecorvo, G. (2004). A new pair of *CR1*-like *LINE* and tRNA-derived *SINE* in *Podarcis sicula* genome. *Gene*, **339**, 189-198.

Fantaccione, S., Woodrow, P. and Pontecorvo, G. (2008). Identification of a family of *SINEs* and *LINEs* in the *Pipistrellus kuhli* genome: A new structure and functional symbiotic relationship. *Genomics*, **91**, 178-185.

Feber, A., Wilson, G.A., Zhang, L., Presneau, N., Idowu, B. Down, T.A., Rakyan, V.K., Noon, L.A., Lloyd, A.C., Stupka, E., Schiza, V., Teschendorff, A.E., Schroth, G.P., Flanagan, A. and

Beck, S. (2011). Comparative methylome analysis of benign and malignant peripheral nerve sheath tumors. *Genome Research*, **21(4)**, 515- 524.

Fedorova, N.D., Khaldi, N., Joardar, V.S., Maiti, R. Amedeo, P., Anderson, M.J., Crabtree, J., Silva, J.C., Badger, J.H., Albarrag, A., Angiuoli, S., Bussey, H., Bowyer, P., Cotty, P.J., Dyer, P.S., Egan, A., Galens, K., Fraser-Liggett, C.M., Haas, B.J., Inman, J.M., Kent, R., Lemieux, S., Malavazi, I., Orvis, J., Roemer, T., Ronning, C.M., Sundaram, J.P., Sutton, G., Turner, G., Venter, J.C., White, O.R., Whitty, B.R., Youngman, P., Wolfe, K.H., Goldman, G.H, Wortman, J.R., Jiang, B., Denning, D.W. and Nierman, W.C. (2008). Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genetics*, **4(4)**, e1000046.

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L.L. and Bateman, A. (2008). The Pfam protein families database. *Nucleic Acids Research*, **36**(Database issue), D281-D288.

Finerty, P.J. Jr and Bass, B.L. (1997). A Xenopus zinc finger protein that specifically binds dsRNA and RNA-DNA hybrids. *Journal of Molecular Biology*, **271**, 195–208.

Finkler, A., Ben-Zvi, B.S., Koltin, Y. and Barash, I. (1988). Transcription and *in vitro* translation of the dsRNA virus isolated from *Rhizoctonia solani*. *Virus Genes*, **1(2)**, 205-209.

Finnegan, D.J. (1989). Eukaryotic transposable elements and genome evolution. *Trends in Genetics*, **5**, 103-107.

Fontaine, T., Simenel, C., Dubreucg, G., Adam, O., Delepierre, M., Lemoine, J., Vorgias, C.E., Diaquin, M. and Latge, J.P. (2000). Molecular organization of the alkali-insoluble fraction of *Aspergillus fumigatus* cell wall. *Journal of Biological Chemistry*, **275**, 27594-29607.

Froussard, P. (1992). A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Research*, **20**, 2900.

Fujimura, T. and Esteban, R. (2004). The bipartite 3'-cis-acting signal for replication is required for formation of a ribonucleoprotein complex *in vivo* between the viral genome and its RNA polymerase in yeast 23 S RNA virus. *Journal of Biological Chemistry*, **279**, 44219-44228.

Fujimura, T. and Estaban, R. (2011). Cap-snatching mechanism in yeast L-A double-stranded RNA virus. *Proceedings of the National Academy of Sciences USA*, **108**, 17667-17671.

Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., Lee, S.I., Spevak, C.C., Basturkmen, M., Clutterbuck, J., Kapitonov, V.V., Jurka J., Scazzocchio, C.,Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G.H., Draht, O., Busch S., D'Enfert, C., Bouchier, C., Goldman, G.H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J.H., Yu, J.,Vienken, K., Pain, A., Freitag, M., Selker, E.U., Archer, D.B.,Penalva, M.A., Oakley, B.R., Momany, M., Tanaka, T., Kumagai, T., Asai, K., Machida, M., Nierman, W.C., Denning, D.W., Caddick, M., Hynes, M., Paoletti, M., Fischer, R., Miller, B., Dyer, P., Sachs, M.S.,Osmani, S.A. and Birren, B.W. (2005). Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*, **438(7071)**, 1105-1115.

Garcia-Cuellar, M.P., Esteban, R. and Fujimura, T. (1997). RNA-dependent RNA polymerase activity associated with the yeast viral p91/20S RNA ribonucleoprotein complex. *RNA,* **3**, 27-36.

Ghabrial, S.A. (1998). Origin, adaptation and evolutionary pathways of fungal viruses. *Virus Genes*, 16(1), 119-131.

Ghabrial, S.A., Suzuki, N. (2008). Fungal Viruses. In B. W. J. Mahy and M. H. V. Van Regenmortel (ed.), Encyclopedia of Virology, 3rd ed., vol. 2. Elsevier, Oxford, United Kingdom, 284-291.

Gibbs, M., Pfeiffer, P. And Fukuhara, T. (2005). *Endornavirus*, in *Eighth Report of the International Committee on Taxonomy of Viruses*, ed. By Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U. and Ball, L.A., Elsevier Academic Press, London, 2005).

Gogolevsky, K.P., Vassetzky, N.S. and Kramerov, D.A. (2008). *Bov-B* mobilized *SINEs* in vertebrate genomes. *Gene*, **407**, 75-85.

Gogolevsky, K.P., Vassetzky, N.S. and Kramerov, D.A. (2009). 5S rRNA-derived and tRNA-derived *SINEs* in fruit bats. *Genomics*, **93**, 494-500.

Goldman, G.H. and Osmani, S.A. (2007). The Aspergilli: Genomics, medical aspects, biotechnology, and research methods. *CRC Press*, **26**, 576.

Goldoni, M., Azzalin, G., Macino, G. and Cogoni, C. (2004). Efficient gene silencing by expression of double stranded RNA in *Neurospora crassa*. *Fungal Genetics and Biology*, **41**, 1016-1024.

Goyon, C., Rossignol, J.L. and Faugeron, G. (1996). Native DNA Repeats and Methylation in *Ascobolus*. *Nucleic Acids Research*, **24(17)**, 3348-3356.

Grandbastien, M.A. (1998). Activation of plant retrotransposons under stress conditions. *Trends in Plant Science*, **3**, 181-187.

Gualtieri, A., Andreola, F., Sciamanna, I., Sinibaldi-Vallebona, P., Serafino, A. and Spadafora, C. (2013). Increased expression and copy number amplification *of LINE-1* and *SINE B1* retrotransposable elements in murine mammary carcinoma progression. *Oncotarget***, 4(11)**, 1882-1893.

Hammond, S.M., Caudy, A.A. and Hannon, G.J. (2001). Post-transcriptional gene silencing by double-stranded RNA. *Nature Reviews Genetics*, **2**, 110-119.

Hammond, S.M. (2005). Dicing and slicing the core machinery of the RNA interference pathway. *FEBS Letters*, **579**, 5822-5829.

Hammond, T.M. and Keller, N.P. (2005). RNA silencing in *Aspergillus nidulans* in independent of RNA-dependent RNA polymerase. *Genetics*, **169**, 607-617.

Hammond, T.M., Bok, J.W., Andrewski, M.D., Reyes-Dominguez, Y., Scazzocchio, C. and Keller, N.P. (2008a). RNA silencing gene truncation in the filamentous fungus *Aspergillus nidulans*. *Eukaryotic Cell*, **7**, 339-349.

Hammond, T.M., Andrewski, M.D., Roossinck, M.J. and Keller, N.P. (2008b). *Aspergillus* mycoviruses are targets and suppressors of RNA silencing. *Eukaryotic Cell*, **7(2)**, 350-357.

Hansen, D. R., van Alfen, N. K., Gillies, K. and Powell, W. A. (1985). Naked dsRNA associated with hypovirulence of *Endothia parasitica* is packaged in fungal vesicles. *Journal of General Virology*, **66**, 2605-2614.

He, C., Nourse, J.P., Kelemu, S., Irwin, J.A.G. and Manners, J.M. (1996). *CgT1*: a non-LTR retrotransposon with restricted distribution in the fungal phytopathogen *Colletotrichum gloeosporioides*. *Molecular and General Genetics*, **252**, 320-331.

Herrera-Estrella, A., Goldman, G.H. and Van Montagu, M. (1990). High-efficiency transformation system for the biocontrol agents, *Trichoderma* spp*. Molecular Microbiology,* **4***,* 839-843.

Henry, C., Mouyna, I. and Latge, J.P. (2007). Testing the efficacy of RNA interference constructs in *Aspergillus fumigatus*. *Current Genetics*, **51**, 277-284.

Hillman, B.I. and Cai, G. (2013). The family *Narnaviridae*: simplest of RNA viruses. *Advances in Virus Research*, 86, ed Ghabrial SA (Elsevier, New York), 149-176.

Himeno, M., Maejima, K., Komatsu, K., Ozeki, J., Hashimoto, M., Kagiwada, S., Yamaji, Y. and Namba, S. (2010). Significantly low level of small RNA accumulation derived from an encapsidated mycovirus with dsRNA genome. *Virology*, **396**, 69-75.

Hodgson, V.J., Button, D. and Walker, G.M. (1995). Anti-*Candida* activity of a novel killer toxin from the yeast *Williopsis mrakii. Microbiology*, **141**, 2003-2012.

Holzberg, S., Brosio, P., Gross, C. and Pogue, G.P. (2002). Barley stripe mosaic virus-induced gene silencing in a monocot plants. *The Plant Journal*, **30(3)**, 315-327.

Hong, Y., Dover, S.L., Cole, T.E., Brasier, C.M. and Buck, K.W. (1999). Multiple mitochondrial viruses in an isolate of the Dutch elm disease fungus *Ophiostoma novo-ulmi*. *Virology*, **258**, 188-127.

Howitt, R.L.J., Beever, R.E., Pearson, M.N. and Forster, R.L.S. (2001). Genome characterization of *Botrytis* virus F, a flexuous rod-shaped mycovirus resembling plant 'potex-like' viruses. *Journal of General Virology*, **82**, 67-78.

Huang, X. and Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Research*, **9**, 868-877.

Huber, F. and Bignell, E. (2014). Distribution, expression and expansion of *Aspergillus fumigatus LINE*-like retrotransposon populations in clinical and environmental isolates. *Fungal Genetics and Biology*, **64**, 36-44.

Hull, R. and Covey, S. (1993). Characterization of cauliflower mosaic virus DNA forms isolated from infected turnip leaves. *Nucleic Acids Research*, **11**, 1881-1895.

International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Jablonski, S.A. and Morrow, C.D. (1995). Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with alteres metal ion requirements for activity. *Journal of Virology*, **69**, 1532-1539.

Jamal, A., Bignell, E.M. and Coutts R.H.A. (2010). Complete nucleotide sequences of four dsRNAs associated with a new chrysovirus infecting *Aspergillus fumigatus. Virus Research*, **153**, 64-70.

Jiang, D. and Ghabrial, S.A. (2004). Molecular characterization of *Penicillium chrysogenum* virus: reconsideration of the taxonomy of the genus Chrysovirus. *Journal of General Virology*, **85**, 2111-2121.

Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. *Trends in Genetics*, **16(9)**, 418-420.

Kachroo, P., Leong, S.A., and Chatto, B.B. (1995). *Mg-SINE*: A short interspersed nuclear element from the rice blast fungus, *Magnaporthe grisea*. *Proceedings of the National Academy of Sciences*, USA, **92**, 11125-11129.

Kajikawa, M. and Okada, N. (2002). *LINEs* mobilize *SINEs* in the eel through a shared 3' sequence. *Cell*, **111**, 433-444.

Kamper, J., Kahmann, R., Bolker, M., Ma, L., Brefort, T., Saville, B.J., Banuett, F., Kronstad, J.W., Gold, S.E., Muller, O. (2006). Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, **444**, 97-101.

Kanematsu, S., Sasaki, A., Onoue, M., Oikawa, Y. and Ito, T. (2010). Extending the fungal host range of a partitivirus and a mycoreovirus from *Rosellinia necatrix* by inoculation of protoplasts with virus particles. *Phytopathology*, **100**, 922-930.

Kapitonov, V.V. and Jurka, J. (2003). A novel class of *SINE* elements derived from 5S RNA. *Molecular Biology and Evolution*, **20**, 694-702.

Kapitonov, V.V. and Jurka, J. (2006). *I-4_AO*, a family of I non-LTR retrotransposons in the *Aspergillus oryzae* genome. *Repbase Reports*, **6(1)**, 12-12.

Kapitonov, V.V. and Jurka, J. (2006). *SINE3-1_AO*, a family of 5S rRNA-derived nonautonomous non-LTR retrotransposons in the *Aspergillus oryzae* genome. *Repbase Reports*, **6(1)**, 45-45.

Kapitonov, V.V. and Jurka, J. (2006). SINE3-2_AO, a family of 5S rRNA-derived nonautonomous non-LTR retrotransposons in the *Aspergillus oryzae* genome. *Repbase Reports*, **6(1)**, 46-46.

Katoh, K., Misawa, K., Kuma, K. and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059-3066.

Kavanagh, K. (2007). Medical Mycology: Cellular and Molecular techniques, John Wiley & Sons, United Kingdom, p. 348.

Kazmierczak, P., McCabe, P., Turina, M., Jacob-Wilk, D. and van Alfen, N.K. (2012). The mycovirus CHV1 disrupts secretion of a developmentally regulated protein in *Cryphonectria parasitica*. *Journal of Virology*, **86(11)**, 6067-6074.

Khalaj, V., Eslami, H., Azizi, M., Rovira-Graells, N. and Bromley, M. (2007). Efficient downregulation of *alb1* gene using an AMA1-based episomal expression of RNAi construct in *Aspergillus fumigatus*. *FEMS Microbiology Letters*, **270**, 250-254.

Kim, J.W. and Bozarth, R.F. (1985). Intergeneric occurrence of related fungal viruses: the *Aspergillus ochraceous* virus complex and its relationship to the *Penicillium stoloniferum* virus S. *Journal of General Virology*, **66**, 1991-2002.

Kim, H.G, Meinhardt, L.W, Benny, U. and Kistler, H.C. (1995). *Nrs*I, a repetitive element linked to pisatin demethylase genes on a dispensable chromosome of *Nectria haematococca*. *Molecular Plant-Microbe Interactions*, **8**, 524-531.

Kinsey, J. A. and Helber, J. (1989). Isolation of a transposable element from *Neurospora crassa*. *Proceedings of the National Academy of Sciences*, **86**, 1929-1933.

Kohany, O., Gentles, A.J., Hankus, L. and Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase: Repbase Submittor and Censor. *BMC Bioinformatics*, **7**, 474.

Koltin, Y. (1988). The killer system of *Ustilago maydis*: secreted polypeptides encoded by viruses, p. 209-242. In Koltin, Y. and Leibowitz, M.J. (ed.), Viruses of fungi and simple eukaryotes. Marcel Dekker, Inc., New York.

Kondo, H., Chiba, S., Toyoda, K. and Suzuki, N. (2013). Evidence of negative-strand RNA virus infection in fungi. *Virology*, **435**, 201-209.

Koonin, E.V. and Dolja, V.V. (1993). Evolution and taxonomy of positive-stranded viruses: implications of comparative analysis of amino acid seqeunces. *Critical Review in Biochemistry and Molecular Biology*, **28**, 375-430.

Koonin, E.V., Wolf, Y.I., Nagasaki, K. and Dolja, V.V. (2008). The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nature Reviews Microbiology*, **6**, 925-939.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283-292.

Kozlakidis, Z., Hacker, C.V., Bradley, D., Jamal, A., Phoon, X., Webber, J., Brasier, C.M., Buck, K.W. and Coutts, R.H. (2009). Molecular characterisation of two novel double-stranded RNA elements from *Phlebiopsis gigantea*. *Virus Genes*, **39**, 132-136.

Kozlakidis, Z., Herrero, N., Ozkan, S., Bhatti, M.F. and Coutts, R.H.A. (2013). A novel dsRNA element isolated from the *Aspergillus foetidus* mycovirus complex. *Archives of Virology*, **15**, 2625-2628.

Kramerov, D., Vassetzky, N. and Serdobova, I. (1999). The evolutionary position of dormice (Gliridae) in *Rodentia* determined by a novel short retroposon. *Molecular Biology and Evolution*, **16**, 715–717.

Kramerov, D.A. and Vassetzky, N.S. (2011a). *Wiley Interdisciplinary Reviews: RNA*, **2**, 772-786.

Kramerov, D.A. and Vassetzky, N.S. (2011b). Origin and evolution of *SINEs* in eukaryotic genomes. *Heredity*, **107**, 487-495.

Kreuze, J.F., Savenkov, E.I., Cuellar, W., Li, X. and Valkonen, J.P.T. (2005). Viral class 1 RNase III involved in suppression of RNA silencing. *Journal of Virology*, **79**, 7227-7238.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J*ournal of Molecular Biology*, **305(3)**, 567-580.

Le Rouzic, A. and Capy, P. (2005). The first steps of transposable elements invasion: parasitic strategy *vs*. genetic drift. *Genetics*, **169**, 1033–1043.

Lamarre, C., Sokol, S., Debeaupuis, J.P., Henry, C., Lacroix, C., Glaser, P., Coppee, J.Y., Francois, J.M. and Latge, J.P. (2008). Transcriptomic analysis of the exit from dormancy of *Aspergillus fumigatus* conidia. *BMC Genomics*, **9**, 417.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M. and FitzHugh, W. (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Lawe, T.M. and Eddy, S.R. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, **25**, 955-964.

Lee, K.M., Yu, J., Son, M., Lee-Y.W. and Kim, K.H. (2011). Transmission of *Fusarium boothii* mycovirus *via* protoplast fusion causes hypovirulence in other phytopathogenic fungi. *PLoS ONE,* **6(6)**, e21629.

Lerat, E. and Semon, M. (2007). Influence of the transposable element neighborhood on human gene expression in normal and tumor tissues. *Gene*, **396**, 303-311.

Lima Favaro, L.C., Araujo, W.L., Azevedo, J.L. and Paccola-Meirelles, L. D. (2005). The biology and potential for genetic research of transposable elements in filamentous fungi. *Genetics and Molecular Biology*, **28(4)**, 804-813.

Lin, Y.H., Chiba, S., Tani, A., Kondo, H., Sasaki, A., Kanematsu, S. and Suzuki, N. (2012). A novel quadripartite dsRNA virus isolated from a phytopathogenic filamentous fungus, *Rosellinia necatrix*. *Virology*, **426(1)**, 42-50.

Liu, H., Cottrell, T.R., Pierini, L.M., Goldman, W.E. and Doering, T. (2002). RNA interference in the pathogenic fungus *Cryptococcus neoformans*. *Genetics*, **160**, 463-470.

Liu, W., Duns, G. and Chen, J. (2008). Genomic characterization of a novel partitivirus infecting *Aspergillus ochraceus*. *Virus Genes*, **37(3)**, 322-327.

Liu, L., Xie, J., Cheng, J., fu, Y., Li, G., Yi, X. and Jiang, D. (2014). Fungal negative-stranded RNA virus that is related to bornaviruses and nyaviruses. *Proceedings of the National Academy of Sciences USA*, **111(33)**, 12205-12210.

Lu, C., Meyers, B.C. and Green, P.J. (2007). Construction of small RNA cDNA libraries for deep sequencing. *Methods*, **43(2)**, 110-117.

Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987). Selection of AUG initiation codons differs in plants and animals. *EMBO Journal*, **6**, 43-48.

Mabey, G.J., Cooley, J. and Bowyer, P. (2012). CADRE: the central *Aspergillus* data Repository 2012. *Nucleic Acids Research*, **40**, D660-6.

Magae, Y. (2012). Molecular characterization of a novel mycovirus in the cultivated mushroom, *Lentinula edodes*. *Virology Journal*, **9**, 60.

Magliani, W., Conti, S., Gerloni, M., Bertolotti, D. and Polonelli, L. (1997). Yeast killer systems. *Clinical Microbiology Reviews*, **10**, 369-400.

Marchenko, A.N., Kozlov, D.G., Svirshchevskaya, E.V., Viskova, N.Y. and Benevolensky, S.V. (2003). The p1 protein of the yeast transposon *Ty1* can be used for the construction of bi-functional virus-like particles. *Journal of Molecular Microbiology and Biotechnology*, **5**, 97-104.

Marchler-Bauer, A. Zheng, C., Chitsaz, F., Derbyshire, M.K., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D. and Bryant, S.H. (2013). "CDD: conserved domains and protein three-dimensional structure", *Nucleic Acids Research*, **41(D1)**, D384-52.

Marquez, L.M., Redman, R.S., Rodriguez, R.J. and Roossinck, M.J. (2007). A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. *Science*, **315**, 515-523.

Mascia, T., Nigro, F., Abdallah, A., Ferrara, M., de Stradis, A., Faedda, R., Palukaitis, P. and Gallitelli, D. (2014). Gene silencing and gene expression in phytopathogenic fungi using a plant virus vector. *Proceedings of the National Academy of Sciences, USA*, **111(11)**, 4291-4296.

McDonald, T., Brown, D., Keller, N.P. and Hammond, T.M. (2005). RNA silencing of mycotoxin production in *Aspergillus* and *Fusarium* species. *Molecular Plant-Microbe Interactions*, **18**, 539-545.

McDonald, S.M., Tao, Y.J. and Patton, J.T. (2009). The ins and outs of four-tunneled Reoviridae RNA-dependent RNA polymerase. *Current Opinion in Structural Biology*, **19**, 775-782.

Melnyk, C.W., Molnar, A. and Baulcombe, D.C. (2011) Intercellular and systemic movement of RNA silencing signals. The *EMBO Journal.* **30**, 3553-3563.

Mes, J.J., Haring, M.A. and Cornelissen, B.J. (2000). *Foxy*: an active family of short interspersed nuclear elements from *Fusarium oxysporum*. *Molecular and General Genetics*, **263**, 271-280.

Milgroom, M.G. and Cortesi, P. (2004). Biological control of chestnut blight with hypovirulence: a critical analysis. *Annual Review of Phytopathology*, **42**, 311-338.

Moon, S., Byun, Y., Kim, H., Jeong, S. and Han, K. (2004). Predicting genes expressed *via* -1 and +1 frameshifts. *Nucleic Acids Research*, **32**, 4884-4892.

Morandini, P. (2010). Inactivation of allergens and toxins. *New Biotechnology*, **27(5)**, 482-493.

Mouyna, I., Henry, C., Doering, T.L. and Latge, J.P. (2004) Gene silencing with RNA interference in the human pathogenic fungus *Aspergillus fumigatus*. *FEMS Microbiology Letters*, **237**, 317-324.

Murata, H., Miyazaki, Y. and Babasaki, K. (2001). The long terminal repeat (LTR) sequence of *marY1*, a retroelement from the ectomycorrhizal homobasidiomycete *Tricholoma matsutake*, is highly conserved in various higher fungi. *Bioscience, Biotechnology and Biochemistry*, **65**, 2297-2300.

Nagawa, M., Nakatani, A., Gonda, K., Shimosaka, M. and Okazaki. (1996). Replication of double-stranded RNA in mycovirus from the plant pathogenic fungus, *Fusarium solani*. *FEMS Microbiology Letters*, **137**, 45-49.

Nakayashiki, H., Hanada, S., Bao-Quoc, N., Kadotani, N., Tosa, Y. and Mayama, S. (2005a). RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genetics and Biology*, **42,** 275-283.

Nakayashiki, H. (2005b). RNA silencing in fungi: Mechanisms and applications. *FEBS Letters,* **579**, 5950-5957.

Neuveglise, C., Sarfati, J., Latge, J.P. and Paris, S. (1996). *Afut1*, a retrotransposon-like element from *Aspergillus fumigatus*. *Nucleic Acids Research*, **24(8)**, 1428-1434.

Ngo, H., Tschudi, C., Gull, K. and Ullu, E. (1998). Doubles stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proceedings of National Academy of Sciences USA*, **95**, 14687-14692.

Nguyen, Q.B., Kadotani, N., Kasahara, S., Tosa, Y., Mayama, S. and Nakayashiki, H. (2008). Systematic functional analysis of calcium-signalling proteins in the genome of the rice-blast fungus, *Magnaporthe oryzae*, using a high-throughput RNA-silencing system. *Molecular Microbiology*, **68**, 1348-1365.

Nibert, M., Ghabrial, S.A., Maiss, E., Lesker, T., Vainio, E.J., Jiang, D. and Suzuki, N. (2014). Taxonomic reorganization of family *Partitiviridae* and other recent progress in partitivirus research. *Virus Research*, **188**, 128–141.

Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., Berriman, M., Abe, K., Archer, D.B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P.S., Farman, M., Fedorova, N., Fedorova, N., Feldblyum, T.V., Fischer, R., Fosker, N., Fraser, A., García, J.L., García, M.J., Goble, A., Goldman, G.H., Gomi, K., Griffith-Jones, S., Gwilliam, R., Haas, B., Haas, H., Harris, D., Horiuchi, H., Huang, J., Humphray, S., Jiménez, J., Keller, N., Khouri, H., Kitamoto, K., Kobayashi, T., Konzack, S., Kulkarni, R., Kumagai, T., Lafon, A., Latgé, J.P., Li, W., Lord, A., Lu, C., Majoros, W.H., May, G.S., Miller, B.L., Mohamoud, Y., Molina, M., Monod, M., Mouyna, I., Mulligan, S., Murphy, L., O'Neil, S., Paulsen, I., Peñalva, M.A., Pertea, M., Price, C., Pritchard, BL., Quail, MA., Rabbinowitsch, E., Rawlins, N., Rajandream, M.A., Reichard, U., Renauld, H., Robson, G.D., Rodriguez de Córdoba, S., Rodríguez-Peña, J.M., Ronning, C.M., Rutter, S., Salzberg, S.L., Sanchez, M., Sánchez-Ferrero, J.C., Saunders, D., Seeger, K., Squares, R., Squares, S., Takeuchi, M., Tekaia, F., Turner, G., Vazquez de Aldana, C.R., Weidman, J., White, O., Woodward, J., Yu, J.H., Fraser, C., Galagan, J.E., Asai, K., Machida, M., Hall, N., Barrell, B. and Denning, D.W. (2005). Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature*, **438**, 1151-1156.

Nishihara, H., Terai, Y., Okada, N. (2002). Characterization of novel Alu- and tRNA-related SINEs from the tree shrew and evolutionary implications of their origins. *Molecular Biology and Evolution*, **19**, 1964-1972.

Nishihara, H., Smit, A.F. and Okada, N. (2006). Functional noncoding sequences derived from *SINEs* in the mammalian genome. *Genome Research,* **16(7)**, 864-874.

Nishimura, M., Hayashi, N., Jwa, N.S., Lau, G.W., Hamer, J.E. and Hasebe, A (2000). Insertion of the *LINE* retrotransposon *MGL* causes a conidiophore pattern mutation in *Magnaporthe grisea*. *Molecular Plant- Microbe Interactions*, **13**, 892-894.

Novikova, O., Fet, V. and Blinov, A. (2009). Non-LTR retrotransposons in fungi. Functional and Integrative *Genomics*, **9(1)**, 27-42.

Nuss, D.L. and Koltin, Y. (1990). Significance of dsRNA genetic elements in plant pathogenic fungi. *Annual Review of Phytopathology*, **28**, 37-58.

Nuss, D.L. (2005a). Hypovirulence: mycoviruses at the fungal-plant interface. *Nature Reviews Microbiology*, **3**, 632-642.

Nuss, D.L., Hillman, B.I., Rigling, D. and Suzuki, N. (2005b). *Hypoviridae*, in *Eighth Report of the International Committee on Taxonomy of Viruses*, ed. By Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U. and Ball, L.A., Elsevier Academic Press, London, 2005).

Nuss, D.L. (2011). Mycoviruses, RNA silencing and viral RNA recombination. *Advances in Virus Research*, **80**, 25-48.

Ogino, T. and Banerjee, A.K. (2007). Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of Vesicular stomatitis virus. *Molecular Cell*, **25**, 85-97.

Ohshima, K., Hamada, M., Terai, Y. and Okada, N. (1996). The 3'-end of tRNA-derive short interspersed repetitive elements are derived from 3'-ends of long interspersed repetitive elements. *Molecular and Cell Biology*, **16**, 3756-3764.

Oliveira, J.M., van de Veen, D., de Graaff, L.H. and Ling, Q. (2008). Efficient cloning system for construction of gene silencing vectors in *Aspergillus niger*. *Applied Microbiology and Biotechnology*, **80**, 917-924.

Ostertag, E.M. and Jr. Kazazian, H.H. (2001). Biology of mammalian L1 retrotransposons. *Annual Review of Genetics*, **35**, 501-538.

Padmanabhan, M. and Dinesh-Kumar, S.P. (2009). Virus-induced gene silencing as a tool for delivery of dsRNA into plant. *Cold Spring Harbor Protocol*, **4(2)**, doi:10.1101/ pdb.prot 5139.

Pai, S.I., Lin, Y.Y., Macaes, B., Meneshian, A., Hung, C.F. and Wu, T.C. (2006). Prospects of RNA interference therapy for cancer. *Gene Therapy*, **13**, 464-477.

Pall, G.S., Codony-Servat, C., Byrne, J., Ritchie, L. and Hamilton, A. (2007). Carbodiimide-mediated cross-linking of RNA to nylon membrane improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Research*, **35(8)**, e60, doi: 10.1093/nar/gkm112.

Panwar, V., McCallum, B. and Bakkeren, G. (2013). Host-induced gene silencing of wheat leaf rust fungus *Puccinia triticina* pathogenicity genes mediated by the Barley stripe mosaic virus. *Plant Molecular Biology*, **81(6)**, 595-608.

Paquin, C.E. and Williams, V.M. (1988). In: Eukaryotic transposable elements as mutagenic agents, Banbury Report 30, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Paule, M.R. and White, R.J. (2000). Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Research*, **28**, 1283-1298.

Perrin, R.M., Fedorova, N.D., Bok, J.W., Cramer Jr R.A., Wortman, J.R., Kim, H.S., Nierman, W.C. and Keller, N.P. (2007). Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. *PLoS Pathogens*, **3(4)**, e50. doi:10.1371/journal.ppat.0030050.

Petersen, TN., Brunak, S., von Heijne, G. and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions, *Nature Methods*, **8**, 785-786.

Peyret, H. and Lomonossoff, G.P. (2013). The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. *Plant Molecular Biology*, **83(1-2)**, 51-58.

Presig, O., Moleleki, N., Smit, W.A. Wingfield, B.D. and Wingfield, M.J. (2000). A novel RNA mycovirus in a hypovirulent isolate of the plant pathogen *Diaporthe ambigua*. *Journal of General Virology*, **81**, 3107-3114.

Rappleye, C.A., Engle, J.T. and Goldman, W.E. (2004). RNA interference in *Histoplasma capsulatum* demonstrates a role for alpha-(1,3)-glucan in virulence. *Molecular Microbiology*, **53**, 153-165.

Rasmussen, M., Rossen, L. and Giese, H. (1993). *SINE*-like properties of a highly repetitive element in the genome of the obligateparasitic fungus *Erysiphe graminis* f. sp. *hordei* . *Molecular and General Genetics*, **239**, 298-303.

Ravindran, S. (2012). Barbara McClintock and the discovery of jumping genes. *Proceedings of the National Academy of Sciences*, *USA*, **109**, 20198-20199.

Rawlings, N.D., Barrett, A.J. and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*, **40** (database issue), D343-D350.

Refos, J.M., Vonk, A.G., Eadie, K., Lo-Ten-Foe, J.R., Verbrugh, H.A., van Diepeningen, A.D., Wendy, W.J. and van de Sande, W.J. (2013). Double-stranded RNA mycovirus infection of *Aspergillus fumigatus* is not dependent on the genetic make-up of the host. *PLoS ONE*, **8(10)**, e77381. doi:10.1371/journal.pone.0077381.

Robin, C. and Heiniger, U. (2001). Chestnut blight in Europe: diversity of *Cryphonectria parasitica*, hypovirulence and biocontol. *Forest, Snow and Landscape Research*, **76**, 361-367.

Robinson, H.L. and Deacon, J.W. (2002). Double-stranded RNA element in *Rhizoctonia solani* AG3. *Mycological Research*, **106**, 12-22.

Romaine, C. P. and Schlagnhaufer, B. (1995). PCR analysis of the viral complex associated with La France disease of *Agaricus bisporus*. *Applied and Environmental Microbiology*, **61**, 2322-2325.

Roossinck, M.J., Sabanadzovic, S., Okada, R. and Valverde, R.A. (2011). The remarkable evolutionary history of Endornaviruses. *Journal of General Virology*, **92**, 2674-2678.

Sainsbury, F., Thuenemann, E.C. and Lomonossoff, G.P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal*, **7**, 682-693.

Sambrook, J., Fritsch, E.M. and Maniatis, T. (1989). Molecular cloning: A laboratory manual, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Habor Labor, New York.

Sandy, P., Ventura, A. and Jacks, T. (2005). Mammalian RNA1: a practical guide. *BioTechniques*, **39**, 215-224.

Scherm, B., Orru, M., Balmas, V., Spanu, F., Azara, E., Delogu, G., Hammond, T. M., Keller, N. P. and Migheli, Q. (2011). Altered trichothecene biosynthesis in TRI6-silenced transformants of *Fusarium culmorum* influences the severity of crown and foot rot on durum wheat seedlings, *Molecular Plant Pathology*, **12**, 759-771.

Schmid, C.W. (1996). *Alu*: structure, origin, evolution, significance and function of one-tenth of human DNA. *Progress in Nucleic Acid Research and Molecular Biology*, **53**, 282-319.

Schmitt, M.J. and Neuhausen, F. (1994). Killer toxin-secreting double-stranded RNA mycoviruses in the yeasts *Hanseniaspora uvarum* and *Zygosaccharomyces bailii*. *Journal of Virology,* **68(3)**, 1765-1772.

Schmitt, M.J., Poravou, O., Trenz, K. and Rehfeldt, K. (1997). Unique double-stranded RNA responsible for the anti-*Candida* activity of the yeast *Hanseniaspora uvarum*. *Journal of Virology*, **71**, 8852-8855.

Schmitt, M.J. and Breinig, F. (2002). The viral killer system in yeast: From molecular biology to application. *FEMS Microbiology Reviews*, **26**, 257-276.

Schnell, M. J. and Conzelmann, K. K. (1995). Polymerase activity of *in vitro* mutated *rabies virus* L protein. *Virology*, **214**, 522-530.

Schrunder, J. And Meinhardt, F. (1995). Extrachromosomal inheritance: yeast linear killer plasmids as a tool in genetic engineering. *Progress in botany*, **56**, 332-353.

Segers, G.C., van Wezel, R., Zhang, X., Hong, Y. and Nuss, D.L. (2006). Hypovirus papain-like protease p29 suppresses RNA silencing in the natural fungal host and in a heterologous plant system. *Eukaryotic Cells*, **5**, 896-904.

Segers, G.C., Zhang, X., Deng, F., Sun, Q. and Nuss, D.L. (2007). Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proceedings of the National Academy of Sciences, USA*, **104**, 12902-12906.

Senthil-Kumar, M. and Mysore, K.S. (2011). New dimensions for VIGS in plant functional genomics. *Trends in Plant Sciences*, **16**, 656-665.

Serdobova, I.M. and Kramerov, D.A. (1997). Short retroposons of the B2 superfamily: evolution and application for the study of rodent phylogeny. *Journal of Molecular Evolution*, **46**, 202-214.

Shaaban, M. Palmer, J.M., El-Naggar, W.A., El-Sokkary, M.A., Habib, el-S.E. and Keller, N.P. (2010). Involvement of transposon-like elements in penicillin gene cluster regulation. *Fungal Genetics and Biology*, **47(5)**, 423–432.

Shen, H.B. and Chou, K.C. (2010). Virus-mPLoc: a fusion classifier for viral protein subcellular location prediction by incorporating multiple sites. *Journal of Biomolecular Structure & Dynamics*, **28**, 175-186.

Silva, K. and Schnierle, B.S. (2010). Selective gene silencing by viral delivery of short hairpin RNA. *Virology Journal*, **7**, 248.

Smit, A.F. and Riggs, A.D. (1995). *MIRs* are classic, tRNA-derived *SINEs* that amplified before the mammalian radiation. *Nucleic Acids Research*, **23**, 98-102.

Smit, A.F.A. (1996). The origin of interspersed repeats in the human genome. *Current Opinion in Genetics and Development*, **6**, 743-748.

Smit, A.F.A., Hubley, R. and Green, P. *RepeatMasker Open-4.0*. 2013-2015. <http://www.repeatmasker.org>.

Soifer, H.S, Zaragoza, A., Peyvan, M., Behlke, M.A. and Rossi, J.J. (2005). A potential role for RNA interference in controlling the activity of the human *LINE-1* retrotransposon. *Nucleic Acids Research*, **33**, 846-856.

Solorzano, A., Rodriguez-Cousini, N., Esteban, R. and Fujimura, T. (2000). Persistent yeast single-stranded RNA viruses exist in vivo as genomic RNA-RNA polymerase complexes in 1:1 stoichiometry. *The Journal of Biological Chemistry*, **275(34)**, 26428-26435.

Sone, T., Suto, M. and Tomita, F. (1993). Host species-specific repetitive DNA sequence in the genome of *Magnaporthe grisea,* the rice blast fungus. *Bioscience, Biotechnology, **and** Biochemistry*, **57**, 1228-1230**.**

Spagnolo, J.F. and Hogue, B.G. (2001). Requirement of the Poly (A) Tail in Coronavirus Genome Replication. *Advances in Experimental Medicine and Biology*, **494**, 467-474.

Spanu, P.D., Abbott, J.C., Amselem, J., Burgess, T.A., Soanes, D.M., Stuber, K**.,** Ver Loren van Themaat, E., Brown, J.K. (2010). Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science*, **330,** 1543-1546.

Spear, A., Sisterson, M.S., Yokomi, R. and Stenger, D.C. (2010). Plant-feeding insects harbor double-stranded RNA viruses encoding a novel proline-alanine rich protein and a polymerase distantly related to that of fungal viruses. *Virology*, **404**, 304-311.

Strauss, E.E., Lakshman, D.K. and Tavantzis, S.M. (2000). Molecular characterization of the genome of a partitivirus from the basidiomycete *Rhizoctonia solani. Journal of General Virology*, **81**, 549-555.

Sun, Q., Choi, G.H. and Nuss, D.L. (2009). A single Argonaute gene is required for induction of RNA silencing antiviral defense and promotes viral RNA recombination. *Proceedings of the National Academy of Sciences, USA*, **106**, 17927-17932.

Suzuki, N., Geletka, L.M. and Nuss, D.L. (2000). Essential and dispensable virus-encoded replication elements revealed by efforts to develop hypoviruses as gene expression vectors. *Journal of Virology*, **74**, 7568-7577.

Suzuki, N., Supyani, S., Maruyama, K. and Hillman, B.I. (2004). Complete genome sequence of Mycoreovirus-1/Cp9B21, a member of a novel genus within the family *Reoviridae*, isolated from the chestnut blight fungus *Cryphonectria parasitica*. *Journal of General Virology*, **85**, 3437-3448.

Szewczyk, E., Nayak, T., Oakley, C.E., Edgerton, H., Xiong, Y., Taheri-Talesh, N., Osmani, S.A. and Oakley, B.R. (2006). Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nature Protocols*, **1(6)**, 3111-3120.

Tao, J., Ginsberg, I., Banerjee, N., Held, W., Koltin, Y. and Bruenn J.A. (1990). *Ustilago maydis* KP6 killer toxin: structure, expression in *Saccharomyces cerevisiae*, and relationship to other cellular toxins. *Molecular and Cellular Biology*, **10**, 1373-1381.

Tavantzis, S.M. and Bandy, B.P. (1998). Properties of a mycovirus from *Rhizoctonia solani* and its virion-associated RNA polymerase. *Journal of General Virology*, **69**, 1465-1477.

Tavantzis, S. (2008). Partitiviruses of fungi. In: Mahy, B.W.J., Van Regenmortel, M.H.V. (eds) Encyclopedia of virology, 3rd edn. Elsevier, Oxford, 63-68.

Tooley, P.W. amd Hewings, A.D. (1989). Detection of double-stranded RNA in P*hytophthora infectans*. *Phytopathology*, **79**, 470-474.

Ullu, E. and Tschudi, C. (1984). *Alu* sequences and processed 7SL RNA genes. *Nature*, **312**, 171-172.

Valent, B. and Chumley, F.G. (1991). Molecular genetic analysis of the rice blast fungus, *Magnaporthe orysea*. *Annual Review of Phytopathology*, **29**, 443-467.

van Diepeningen, A.D., Debets, A.J.M., Slakhorst, S.M., Fekete, C., Hornok, L., Hoekstra, R.F. (2000). Interspecies virus transfer *via* protoplast fusions between *Fusarium poae* and black *Aspergillus* strains. *Fungal Genetics Newsletters*, **47**, 99-100.

van Diepeningen, A.D., Debets, A.J.M. and Hoekstra, R.F. (2006). Dynamics of dsRNA mycoviruses in black *Aspergillus* populations. *Fungal Genetics and Biology*, **43**, 446-452.

van de Sande, W.W.J., Lo-Ten-Foe, J.R., van Belkum, A., Netea, M.G., Kullberg, B.J. and Vonk, A.G. (2010). Mycoviruses: future therapeutic agents of invasive fungal infections in human? *European Journal of clinical Microbiology and Infectious Diseases*, **29**, 755-763.

Vetukuri, R.R., Tian, Z., Avrova, A.O., Savenkov, E.I., Dixelius, C. and Whisson, S.C. (2011a). Silencing of the *PiAvr3a* effector-encoding gene from *Phytophthora infestans* by transcriptional fusion to a short interspersed element. *Fungal Biology*, **115**, 1225-1233.

Vetukuri, R.R., Avrova, A.O., Grenville-Briggs, L.J., van West, P., Soderbom, F., Sevenkov, E.I., Whisson, S.C. and Dixelius, C. (2011b). Evidence for involvement of Dicer-like, Argonaute and histone deacetylase proteins in gene silencing in *Phytophthora infestans*. *Molecular Plant Pathology*, 12, 772-785.

Vilfan, I.D., Kamping, W., van den Hout, M. and Candelli, A. (2007). An RNA toolbox for single - molecule force spectroscopy studies. *Nucleic Acid Research*, **35**, 6625-6639.

Walker, J.A., Hughes, D.A., Andres, B.A., Shewale, J., Sinha, S.K. and Batzer, M.A. (2003). Quantitative intra-short interspersed element PCR for species-specific DNA identification. *Analytical Biochemistry*, **316**, 259-269.

Wang, Z., Rao, D.D., Senzer, N. and Nemunaitis, J. (2011). RNA interference and cancer therapy. *Pharmaceutical Research*, **28(11)**, 2983-2995.

Wei, Y.D., Collinge, D.B., Smedegaard-Petersen, V. and Thordal-Christensen, H. (1996). Characterization of the transcript of a new class of retroposon-type repetitive element cloned from the powdery mildew fungus *Erysiphe graminis*. *Molecular and General Genetics*, **250**, 477-482.

Wei, C.Z., Osaki, H., Iwanami, T., Matsumoto, N. and Ohtsu, Y. (2004). Complete nucleotide sequences of genome segments 1 and 3 of *Rosellina* anti-rot virus in the family *Reoviridae*. *Achieves of Virology*, **149**, 773-777.

Weiner, A.M. (2002). *SINEs* and *LINEs*: the art of biting the hand that feeds you. *Current Opinion in Cell Biology*, **14**, 343-350.

Wessler, S.R, (1996). Turned on by stress. Plant retrotransposons. *Current Biology*, **6**, 959-961.

Whisson, S.C., Avrova, A. O., Lavrova, O. and Pritchard, L. (2005a). Families of short interspersed elements in the genome of the oomycete plant pathogen, *Phytophthora infestans*. *Fungal Genetics and Biology*, **42**, 351-365.

Whisson, S.C., Avrova, A.O., West, P.V. and Jones, J.T. (2005b). A method for double-stranded RNA-mediated transient gene silencing in *Phytophthora infestans. Molecular Plant Pathology*, **6**, 153-163.

Wichman, H.A., Van den Bussche, R.A., Hamilton, M.J. and Baker, R.J. (1992). Transposable elements and evolution of the genome organization in mammals. *Genetica*, **86**, 287-293.

Wickner, R.B. (1996). Prions and RNA viruses of *Saccharomyces cerevisiae*. *Annual Review of Genetics*, **30**, 109-139.

Wood, H.A., Bozarth, R.F., Adler, J. andMackenzie, D.W. (1974). Proteinaceous virus-like particles from an isolate of *Aspergillus flavus*. *Journal of Virology*, **13**, 532-534.

Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J. and Baker, S. (2002). The genome sequence of *Schizosaccharomyces pombe*. *Nature*, **415**, 871-880.

Wu, M., Jin, F., Zhang, J., Yang, L., Jiang, D. and Li, G. (2012). Characterizationof a novel bipartite double-stranded RNA mycovirus confering hypovirulence in the phytopathogenic fungus B*otrytis porri*. *Journal of Virology*, **86**, 6605-6619.

Xie, J., Wei, D., Jiang, D., Fu, Y., Li, G., Ghabrial, S. and Peng, Y. (2006). Characterization of debilitation-associated mycovirus infecting the plant-pathogenic fungus *Sclerotinia sclerotiorum*. *Journal of General Virology*, **87**, 241-249.

Yaegashi, H., Yoshikawa, N., Ito, T. and Kanematsu, S. (2013). A mycoreovirus suppresses RNA silencing in the white rot fungus *Rosellinia necatrix. Virology*, **444**, 409-416.

Yu, L. and Markoff, L. (2005). The topology of bulges in the long stem of the flavivirus 3' stem-loop is a major determinant of RNA replication competence. *Journal of Virology*, **79**, 2309-2324.

Yu, J., Kwon, S.J., Lee, M.K., Son, M. and Kim, K.H. (2009). Complete nucleotide sequence of double-stranded RNA viruses from *Fusarium graminearum* strain DK3. *Archives of Virology*. **154**, 1855-1858.

Yu, X., Li, B., Fu, Y., Jiang, D., Ghabrial, S.A., Li, G., Peng, Y., Xie, J., Cheng, J., Huang, J., Yi, X. (2010). A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proceedings of the National Academy of Sciences, USA*, **107**, 8387-8392.

Zabalgogeazcoa, I., Petrunak, D., Christ, B.J. and Gildow, F.E. (1997). Unencapsidated double-stranded RNA associated with membrane vesicles in isolates of *Alternaria solani*. *Mycology Research*, **101**, 604-608.

Zamore, P.D. (2001). RNA interference: listening to the sound of silence. *Nature Structural and Molecular Biology*, **8**, 746-750.

Zhang, X., Seger, G.C., Sun, Q., Deng, F. and Nuss, D.L. (2008). Characterization of hypovirus-derived small RNAs generated in the chestnut blight fungus by an inducible DCL-2-dependent pathway. *Journal of virology*, **82**, 2613-2619.

Zhang, T., Jiang, Y. and Dong, W. (2014). A novel monopartite dsRNA virus isolated from the phytopathogenic fungus *Ustilaginoidea virens* and ancestrally related to a mitochondria-associated dsRNA in the green alga *Bryopsis*. *Virology*, **462-463**, 227-235.

Zuker, M (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research,* **31**, 3406-3415.

Appendices

Appendix I: Media and reagents

1) *Aspergillus* Complete Medium (ACM)

Adenine	0.075 g
Glucose	10 g
Yeast extract	1 g
Bacteriological peptone	2 g
Casamino acids	1 g
Vitamin solution*	10 ml
Salt solution*	20 ml
Ammonium tartrate (500 mM)	10 ml
Sterile distilled water	to 1000 ml
pH 6.5 (adjusted with 10 M NaOH)	

To make AMC solid medium, add 1.2 % Oxoid agar No. 3. Solution was sterilised at 121 °C for 15 min

2) *Aspergillus* Salt Solution

KCl	26 g
$MgSO_4.7H_2O$	26 g
KH_2PO_4	76 g
Trace element solution*	50 ml
Sterile distilled water	to 1000 ml

Do not autoclave. Solution should be kept in the refrigerator after 1.5 ml chloroform is added as a preservative.

3) Biotin

Stock 100 µg/ml final concentration was prepared by dissolving 10 mg in 100 ml H_2O. The solution was sterilised at 121 °C for 15 min and then kept in the fridge.

4) Vitamin Solution

4-aminobenzoic acid (PABA)	400 mg
Aneurin (thiamine)	50 mg

Biotin	1 mg
Meso-inositol	24 g
Nicotinic acid	100 mg
DL-pantothenic acid (hemicalcium salt)	200 mg
Pyridoxine monohydrochloride	250 mg
Riboflavin	100 mg
Choline chloride	1.4 g
Sterile distilled water	to 1000 ml

The solution was sterilized at 121 °C for 15 min and covered with aluminium foil and kept at 4 °C.

5) Trace element solution

$Na_2B_4O_7.10H_2O$	40 mg
$CuSo_4.5H_2O$	400 mg
$FePO_4.2H_2O$	800 mg
$MnSO_4.2H_2O$	800 mg
$Na_2MoO_4.2H_2O$	800 mg
$ZnSO_4.7H_2O$	8 g
Sterile distilled water	to 1000 ml

The solution was sterilised at 121 °C for 15 min 121 °C and kept at 4 °C. Mix before use.

6) Minimal medium (MM)

Glucose (1% w/v)	10 g
Aspergillus salt solution	20 ml
Ammonium tartrate (500 mM)	10 ml
Sterile distilled water	to 1000 ml
pH 6.5 (adjusted with 10 M NaOH)	

To make MM solid medium, add 1.2 % Oxoid agar No. 3. The solution was sterilised at 121 °C for 15 min

7) LB medium (Luria-Bertani medium)

Tryptone (Oxoid)	10 g
Yeast extract (Oxoid)	5 g

NaCl (BDH)	10 g
Sterile distilled water	to 1000 ml

The solution was sterilized at 121 °C for 15 min

8) IPTG stock solution (200 mg/ml)

isopropyl-β-D-thiogalactopyranoside (IPTG)	2.0 g

Add water to 10 ml final volume. Filter-sterilize using 0.22 µm disposable filter. Disperse solution into 1 ml aliquots and store at -20°C.

9) X-Gal (20 mg/ml)

5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal)	100 mg
N,N´-dimethylformamide	5 ml
Cover with aluminium foil and store at –20°C.	

10) LB plates with ampicillin

15 g of Technical agar (Oxoid) was added to 1 litre of LB medium and then autoclaved. The medium was allowed to cool to 50°C before adding 0.1 g/ml ampicillin to a final concentration of 0.1 mg/ml. Then 30 ml of the medium was poured into Petri plates, the medium was allowed to solidify and the plates were then stored at 4°C for up to 1 month.

11) LB plates with ampicillin/IPTG/X-Gal

The LB plates with ampicillin were made as described above. IPTG/X-Gal solution (10 µl IPTG, 20 µl X-Gal and 90 µl sterile dH_2O) was spread over the surface of an LB ampicillin plate and allowed to absorb for 30 min at 37°C prior to use.

12) Commonly used electrophoresis buffers

Buffer	Working solution	Concentrated stock solution (per litre)
Tris-acetate (TAE)	1X: 0.04 M Tris-acetate 0.001 M EDTA	50X: 242 g Tris-base (dissolve in H_2O) 57.1 ml glacial acetic acid 100 ml 0.5 M EDTA (pH 8)
Tris-borate (TBE)	0.5: 0.045 M Tris-borate 0.001 M EDTA	5X: 54 g Tris base 27.5 g boric acid 20 ml 0.5 M EDTA (pH 8)

All buffer solutions were then sterilized at 121 °C for 15 min and then kept in the fridge.

13) Transformation storage solution (TSS)

2X LB	5 ml
PEG 8000	1 g
DMSO	0.5 ml
$MgCl_2$ (1M)	0.5 ml
Sterile distilled water	3.0 ml

14) DEPC-treated water (0.1% v/v)

Diethyl pyrocarbonate (Sigma)	1.0 ml
Distilled water	to 1000 ml

The solution was shaken in a water bath at 37 °C overnight and then sterilized at 121 °C for 45 min to inactivate traces of DEPC.

15) TE buffer

1 M Tris-HCl, pH 7.5	The solution was filter sterilized using 0.22 µm disposable filter
0.5 mM EDTA, pH 8	The solution was autoclaved at 121 °C for 15 min

TE buffer was prepared by mixing 50 ml 1 M Tris-HCl with 2 ml 0.5 M EDTA of the stock solution to make a final concentration of 0.5 mM Tris-HCl and 1.0 mM EDTA. Sterile distilled water was added to make up the volume of 1 L TE buffer.

16) Solutions for DIG northern blot hybridization

Solution for gel denaturation:	0.25 N HCl;	dH_2O	78.4 ml
	conc. HCl (SIGMA)	26.1 ml	
	100 mM NaOH;	NaOH (SIGMA)	4 g
		dH_2O	1 L
	100 mM Tris-HCl (pH 8.0)	Tris-HCl (Trizma SIGMA)	157.6 g
		dH_2O	1 L

DIG Easy-Hyb: Add 64 ml sterile DEPC H_2O to DIG Easy-Hyb Granules. Dissolve by stirring at 37 ˚C for 5 min

Washing buffer, pH 7.5:	0.1 M maleic acid	11.607 g
	0.15 M NaCl	8.766 g
	0.3% (v/v) Tween 20	3 ml
	DEPC-treated H_2O	to 1000 ml
Maleic acid buffer, pH 7.5:	0.1 M maleic acid	11.607 g
	0.15 M NaCl	8.766 g
	DEPC-treated H_2O	to 1000 ml
Detection buffer, pH 9.5:	0.1 M Tris-HCl	12.14 g
	0.1 M NaCl	5.855 g
	DEPC-treated H_2O	to 1000 ml
Blocking solution:	prepare 1x blocking solution by diluting 10x blocking solution in maleic acid buffer (1:10)	
Antibody solution:	centrifuge anti-digoxigenin-AP for 5 min at 10,000 rpm. Dilute anti-dig-AP: 1x blocking solution = 1:10,000. Keep in ice.	
2x SSC + 0.1% SDS:	20x SSC	20 ml
	10% SDS	2 ml
	DEPC-treated H_2O	to 200 ml
0.1x SSC + 0.1% SDS:	20x SSC	1 ml
	10% SDS	2 ml
	DEPC-treated H_2O	to 200 ml

17) Protoplasting Media and Solutions:

Osmotic Medium (OM), for 500 ml:

1.2 M $MgSO_4$	(147.88 g of $MgSO_4.7H_2O$)
10 mM Sodium phosphate buffer	(2.5 ml of 2M Sodium phosphate buffer)

Adjust pH to 5.8 with 1M Na_2HPO_4. Filter Sterilize and store at 4°C

2 M Sodium phosphate buffer stock, for 100 ml: pH to 6.5

Na_2HPO_4	(9.09 g NaH_2PO)
$NaH_2PO_4 .H_2O$	(16.34 g OR 18.79g)

Protoplast Trapping Buffer, for 1 L:

0.6 M Sorbitol	(109.3 g of Sorbitol)
0.1 M Tris-HCl, pH 7.0	(100 ml of 1 M Tris-HCl, pH 7.0 stock)

Aliquot in 100 ml, autoclave, store at 4°C

1 M Tris-HCl, pH 7.0, for 500 ml:

60.7 g Tris (in approx 400 ml water), pH to 7.0 with HCl, water to 500 ml

STC Buffer, for 1L:

1.2 M Sorbitol	(218.6 g)
10 mM $CaCl_2$	(1.47 g of $CaCl_2$.2H2O)
10 mMTris-HCl, pH 7.5	(10 ml of 1 M Tris-HCl, pH 7.5 stock)

Aliquot in 50 ml, autoclave, store at 4°C

Polyethylene Glycol (PEG) Solution, for 100 ml:

60% PEG	(60 g of PEG 4000)
50 mM $CaCl_2$	(0.735 g of $CaCl_2.2H_2O$)
50 mMTris-HCl, pH 7.5	(5 ml of 1 M Tris-HCl, pH 7.5 stock)

Final volume needs to be 100 ml, PEG takes up a lot of volume, slowly dissolved PEG into the Tris solution plus about 30 ml of water, slowly add more water as needed, but be careful not to exceed 100 ml of final volume. Autoclave, store at room temp.

20X Salt solution, for 1L:

$NaNo_3$	120 g
KCl	10.4 g
$MgSO_4.7H_2O$	10.4 g
KH_2PO4	30.4 g
dH_2O	to 1000 ml

Stored at room temperature.

SMM (stabilized minimal medium, 1.5%) selection plates, for 1 L:

**used for plating transformed protoplasts

20X Salt solution*	20 ml
Trace element	1 ml

1% D-Glucose (dextrose)	10 g
1.2 M Sorbitol	218.6 g
Yeast extract	1 g
Sterile distilled water	to 1000 ml
pH 6.5 (adjusted with 10 M NaOH)	

To make solid medium, add 1.5 % Oxoid agar No. 3 (15 g). The solution was sterilised at 121 °C for 15 min. For 0.7% Top agar, use SMM with 7 g agar/L.

Appendix II: Supplementary data

Sequences for plasmid construction

>pAfuTmV-1 dsRNA2 (4932 bp)

T7: 5´- **TAATACGACTCACTATAGGG**- 3´

Oligo-RNA2FS 5'- CCCAAGCTTTCTAGA**TAATACGACTCACTATAGGG**TGAATAAAAGATTTGTTCTG -3'
GC%=25 Tm=43.5

Oligo-RNA2RS 5'- CCCGAATTCGGTACC**TAATACGACTCACTATAGGG**CCCCCAGGCCCGGCGGGGGG -3'
GC%=95 Tm=76.3

Rev CCCCCCGCCGGGCCTGGGGG**CCCTATAGTGAGTCGTATTA**GGTACCGAATTCGGG

BAMF 5'- CTCAACAAGGATCCCATGCTCG -3' 1130 bp GC%=54.5 Tm=58

BAMR 5'- CGAGCATGGGATCCTTGTTGAG -3' 1195 bp GC%=54.5 Tm=58

HindIII = AAGCTT, XbaI = TCTAGA, EcoRI = GAATTC, KpnI = GGTACC

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT
CAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACC
GCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTAC
GCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTT
TCTAGATAATACGACTCACTATAGGGTGAATAAAAGATTTGTTCTGGGCAAGCACGCTGCAACGTCGAGCAACTCCCATATAGGGGACGATGAGACATGGC
GGACCTAACCCGTCTTAGGAATATTGTTCTGTCGGGCGATATGAAAGTCGCCCTCACTGCAGCAATTCTGTACCACCTCTCCCCCAACCCCTCCGCGATCA
CCGATTACGGGGTGGATGAGGTCAAGAGCGTGGTGTCATGGCTCGTCTCCTCGCCCGGGATGTTCTACGGCGACGTTGTTGCCGCCACTGGAATCACCCCT
GACATCCTCTCTACTTTCCTTCCCGCGCTCGATGATCAAGCCTGCGCCGACGCCGCCTCTCTTGCCAACGCCATGGCGACCCCCGATGCCGAGTGGGACCG
CGAGCACCGCAACCGTTTCACGTCCGCACTCGCGCACCACTCCGCCGACGCTGCTAGGCTCGCCGCCCAGTCGTCGGCTGAGTTCGCTGCGGTCGACGCCA
TGGAAGAGGCTGGCTACGAGAATCAGGACCACGCTCGCACCGCGGTCAACGAGCTCAACCGCGTCTATGGCAATACCTACCGGGTCTTCACTCGTCAACAC
GCCCGCTCCCGCGTGGCCATCGCCGCCTCTTTCAAGATCGATGACCACGGCACCCTCTTCCACGTCCTCACCCCCTGGTGCCTCTCCCGAACATCGGCCAT
GATCTACGTCGCTGTCATCGCCAGGCATGTGAGGCGTCGCCACATCGAGGAGCGCGCCTCCAGATTGGCACGCACTGCCCTGACCTACGCCGCTCGCGCCG
GCTTAGAGGCTGTACGCGCCATCATAGAGCCCGCCATTGCCATAGTGACAAGTTTCCGTTACGACCCTGCCAAGATCGCCTTCGTGGACGATTTTGACCGG
GTCAGGGAGTCCATCACTACGAGGAGTCCTCCAACGCTTGTCGCCTTCGCTTGCGTCCATATGTGCGGGAACCAGGACGTCGTTTCCTTGGCGGCGCGACT
CCGCATGGAGTCGCTACCACAACTCCATGAGACCCCGCCAGCCGTCGCCGAGTACACCGACCACGTCGACGACGAACAGTCCGCCCTCCTCATGTTCTCTC
GCGTCCTCAAGAGACTGAGACATGAGCGCGACAGGGGGGTGCACGGCCCCCGCCTCAACAAGGATCCCATGCTCGGCTACATAGCCCGGGGAGTACGCCGT
TCCGAGAGGAACACCATGGCGTGCATCAACAGGGTTGAGGAGGCCGTGTCCTTTCTGCGCTCGAAGGGGGTCGACACCCTCAATACACTGTTCGTCATCGA
ATGGGGCGGTGAATTTGACTATCCTACCGTCATGGCTGCCCTCGCCGTTGCCAGGTTGGATGTCTGTGTCGACATTGGTCCCTCGGGCATTGACCTACCCG
GCGCCGACGCCAAGGCCGAGGATGGCACCGAAGAGTACAACTACTCTTTGCTGCTCGCCTCAGCCGAGGAGCGGCGCATGCCTCGTATGCCCAGGGTGCCC
TATCCGTCCGACGAGCCCCTTGAGGCCAAACTGCGGCGCGTCGCGCAGTTCTACTCTCTCTCGGGCTCTGCCAATATAGCATACATCAGCGGGGGGGTCAC
CCAAAACGAGGGGGTTCCTGTCAACATCTGTATCGACACCGCCAACAGGCTTAATGCCATACGCAATGTCACCCCGGACGTTAACGTCGTTTATGCCTGCG
CCGAAGCTCTCCTCCCCCCCCTCTGCGGGCACGGACTCGAGCTCGGGCCCGACGACTACCTTCAGTGCGCGTCCCTGGTCGATGAGGATTGCCCCGCCTGC
AATGCTCACTACCGCGCCACCGCCGTCCTCTCTGCTGCCTGCGCTGACAGACAGGTGCGCATCGTTAAGCCAAGATCCGCCTACGCTCATAACGGGCACCT
GTCCTTGGAGCTCGTCCCGGGCCACCCTACATGCGTGGACGACACTCTCCTCACCGTCGACTCCATGGTCGCGGCCAATGTCCTCCGCAACCACGATTGGG
GCGATGAGCCCGTTACTCCACCAGCCGGCGGCAACCCGACATCGCCTACCCTCGCCCGCGTCACAGCGGAGAACATGCGCACTGTGTACTCCGCTCTCAAT
AACGGCGCCGTGGGGCCCTTCGACTATGACGACATGGCGTCAATTGCCGCTTCAATAGCGCCGTACACTGCGCTGCATCGTGCTTGACCAGCGTCCGACGG
GGAGGGGCTTAATTGCCCCCCCGCCGGGCCTGGGGGCCCTATAGTGAGTCGTATTAGGTACCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTTCCTG
TGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT
GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG
CTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG
CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA
CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA
CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCT
CAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT
CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTA
TCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGT
AAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG
TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
AGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC
GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC
ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCA
ATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCC
GAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC

>pAfuTmV-1-ALB1 (4199 bp)

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGT
CAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACC
GCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTAC
GCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTT
TCTAGATAATACGACTCACTATAGGGTGAATAAAAGATTTGTTCTGGGCAAGCACGCTGCAACGTCGAGCAACTCCCATATAGGGGACGATGAGACATGGC
GGACCTAACCCGTCTTAGGAATATTGTTCTGTCGGGCGATATGAAAGTCGCCCTCACTGCAGCAATTCTGTACCACCTCTCCCCCAACCCCTCCGCGATCA
CCGATTACGGGGTGGATGAGGTCAAGAGCGTGGTGTCATGGCTCGTCTCCTCGCCCGGGATGTTCTACGGCGACGTTGTTGCCGCCACTGGAATCACCCCT
GACATCCTCTCTACTTTCCTTCCCGCGCTCGATGATCAAGCCTGCGCCGACGCCGCCTCTCTTGCCAACGCCATGGCGACCCCCGATGCCGAGTGGGACCG
CGAGCACCGCAACCGTTTCACGTCCGCACTCGCGCACCACTCCGCCGACGCTGCTAGGCTCGCCGCCCAGTCGTCGGCTGAGTTCGCTGCGGTCGACGCCA
TGGAAGAGGCTGGCTACGAGAATCAGGACCACGCTCGCACCGCGGTCAACGAGCTCAACCGCGTCTATGGCAATACCTACCGGGTCTTCACTCGTCAACAC
GCCCGCTCCCGCGTGGCCATCGCCGCCTCTTTCAAGATCGATGACCACGGCACCCTCTTCCACGTCCTCACCCCCTGGTGCCTCTCCCGAACATCGGCCAT
GATCTACGTCGCTGTCATCGCCAGGCATGTGAGGCGTCGCCACATCGAGGAGCGCGCCTCCAGATTGGCACGCACTGCCCTGACCTACGCCGCTCGCGCCG
GCGCGGTTGGATCGAAAGGTCTCACCATCAGTGCACCACCCGAAATTCTCGACAACTTTATCGAAGAAGGTCTTCCGAAGGAGTACAAACACTTCAAGGCT
CCTGGAGTCAGTGGTCCGTACCACGCGCCCCATCTGTACAATGACCGAGAAATTCGCAATATCCTCAGCTTCTGCTCCGAGGACGTGATTCTGCGCCACAC
ACCACGGGTTCCACTGGTCTCGAGCAACACAGGGAAGCTGGTCCAGGTAAAGAGCATGCGTGATCTGCTAAAGGTGGCTCTGGAGGAAATCCTCTTGCGCA
AGATCTGCTGGGACAAAGTCACCGAGTCATGCCTTTCCATCGTTCAGGCTACCAACGACAAGCCCTGGAGGATTCTCCCTATCGCCAGCAACGCCACGCAA
GGCTTGGTTACTGCACTCCAGCGTATGGGAAACTGCCAGATCGAGGTAGACACCGGGGTCGGCGCTCCTCAAATGGACCCGGCCGCTCCCAATGCAACGCC
GGCGGCAACCCGACATCGCCTACCCTCGCCCGCGTCACAGCGGAGAACATGCGCACTGTGTACTCCGCTCTCAATAACGGCGCCGTGGGGCCCTTCGACTA
TGACGACATGGCGTCAATTGCCGCTTCAATAGCGCCGTAGACTGCGCTGCATCGTGCTTGACCAGCGTCCGACGGGGAGGGGCTTAATTGCCCCCCCGCC
GGGCCTGGGGGCCCTATAGTGAGTCGTATTAGGTACCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACT
CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG
CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTA
CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTAT
TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTT
TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACG
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAG
GGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT
GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTT
GCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGG
ATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAC
AGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTT
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAA
ACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC

>Construct of pAF1-ALB1: *AfuSINE2-1a-ALB1/PKSP* - sense orientation.

1-SINE1aPmeI-F	CCCGTTTAAACCAAGTGTACATAGAGC	*Pme*I	Tm=48
2-SINE1b-R	GAGACCTTTCGATCCAACCGCCCTCTATGTACACTTAGTAG	ALB1/PKSP underlined	Tm=45
3-SINE1c-F	CTACTAAGTGTACATAGAGGGCGGTTGGATCGAAAGGTCTC	ALB1/PKSP underlined	Tm=69
4-ALB1aPmeI-R	GGGGTTTAAACGTTGCATTGGGAGCGGCC	*Pme*I	Tm=71

****GTTT'AAAC** *Pme*I site (blunt end)

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATA
AATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACC
GTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA
CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAAC
GACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGATCTAGTTGGCAGTAGGAGGCTGTTC
AGGCGGAGGCGATGAAGCTTTGGGGGCCTCTGACGGTTGTGGCTTGTTGGATTCAACAGGAGGCCGCTTCCCGACATCTGCGGGCTCCTTTGGTGATATTC
GCGGCAACCCTACCTTGTCGGTCTCAATTGTAGCGATTTCTTCGTCCCGCTCAACGTAATCGCCCACCTCTGCCGAATAATGAGTTAAAGATTAGGAGACG
GGTCAAGAGACCTACGCTTCGAGAATTGCTTCAATGTACCCTCGGTGATCGATTCCGCCATCTGAGGGACCTTCACTATGGTATCAGCTGCGGCGAGATCA
GCGAACAACGTCATTGAATCTAAAGCAAATTGCGATACCGTAATTTCTGATCTGATAGCCGCCGAGGCGGGCGAGACTGCCCCGGGGTAAATTGGCTCTAT
GGAATGTCGACGGTCAGCCAGAGCAGCATAATGTGCGCGAACCTCCTTACCCCAAGAACGGAGCATGCAGCAAAACCGCGCCAGACTTATGCCTTCTAAGC
GCCAAAGAGGTGATACTCCTGCGAGCACCAATGCAAGGCAAGCGGGAGCATTGAGCTGTTCGAGGAATAGACCTCAGGTGTAGCTCTGACAGACGGTGAAA
TGGTAACCTAGACGCCATATTTTATCCGATTTGATTGGGTAGGTAATCAGAAGTTAAGGAGCAGGAGGATTTGGCACGATTGTCTGTCCTCAGCAACACAC
GCACATCACCTGCGTTCCATGCGACCGGTGGATGACGGAGTATTATTGTTACGCTTGTTCGGCTTTCAGTAAAGACCAACACGGCAGATTCGGCGCACCCT
ATCAAATTCACTTGGTGCACATCGCAAGAGACACTGCCTTGTAGTACGTGGACAGGATCGTGAAAAATCACTAGAAATTTACAGATTTATTCTTCAGCTAG
CGTCTCTCTCTGCATTCGTTGTGCCAAGTCTCAGGGTGTTCTCCCCTTACGACATTGACAGAGAGTAATCATGTCTGCTAACACTGTTGCGTAAGTTGCTC
TCAGATGGAAGAATAGGCTGCTTTGCTAATCTCAACTTCGCCTTTCAGGTCACTCTACCGACGGAGTCTCAAGCTTGCCCTTGATTGGGCAGTTCACAGAC

ACGTATGGAGAGGACAAGCTGTGTATATTCGGTCCCTGTTTGAGGCAAATAAGGACGTCCGTGATCCTCGTCAACAAAGGGTAAACCCCCTCTGACCCCAG
CAACCGTCTAGCCGTTGCACAACGCCGAGTCATTAGCTTATACTAATGAATTGAGCAGGTGCTACTCCGCGAAACTGAGAAGTCTGATAAGCTAGAAAAAG
GCTGCGTCATGCGCAATTGGCTGATAATGTCTTCCAGGCAGTAAATACGAGAGAAACCTTCCTGCTCGGCAGTTACCTTGTATGGTCACCTTTGATCTGTC
CACTCTGCCTCGATTAACATACGTCTTAGATGCTTCTCAACATGCAGATGCCCACTAAGCTTGGACGGCGTACTAGATCCATGTATTTGCGTCAATTGGAT
GAGACCACAATGAGTTCCCTACTCGAAGATAGACATGGCGTTTAAACCAAGTGTACATAGAGCCGGTAGCGGTAGCGTAGTGGTAAGCGCTCCGAGGCAGC
CTCTAGAGAAGTAGGATTGGTGACCCTGCTTTTTTGGAGGTTATGGTTCGATTCCCGTCGCTGGCACAACATTTACCACCACAATGGAAGATCACTTCCC
ACAATGGTATCAAGGCCACTCCCTTATCGCAAGGTGGTGGGGAAGTTGGAACAATCACAGGCCTGTAAGGCGAGGCTCTAAATTCGCCCTCATGTAATGGA
ACAAAATTGTAACTAGACACACAAGGATTAGCTATAGTCGATACCTGCATATCGCCCAAGGCGAGGGGTCAGCGTATGAGTACTACTACTACTACTACTACT
AAGTGTACATAGAGGGCGGTTGGATCGAAAGGTCTCACCATCAGTGCACCACCCGAAATTCTCGACAACTTTATCGAAGAAGGTCTTCCGAAGGAGTACAA
ACACTTCAAGGCTCCTGGAGTCAGTGGTCCGTACCACGCGCCCCATCTGTACAATGACCGAGAAATTCGCAATATCCTCAGCTTCTGCTCCGAGGACGTGA
TTCTGCGCCACACACCACGGGTTCCACTGGTCTCGAGCAACACAGGGAAGCTGGTCCAGGTAAAGAGCATGCGTGATCTGCTAAAGGTGGCTCTGGAGGAA
ATCCTCTTGCGCAAGATCTGCTGGGACAAAGTCACCGAGTCATGCCTTTCCATCGTTCAGGCTACCAACGACAAGCCCTGGAGGATTCTCCCTATCGCCAG
CAACGCCACGCAAGGCTTGGTTACTGCACTCCAGCGTATGGGAAACTGCCAGATCGAGGTAGACACCGGGGTCGGCGCTCCTCAAATGGACCCGGCCGCTC
CCAATGCAACAAACAAACGGTGATGTCTGCTCAAGCGGGGTAGCTGTTAGTCAAGCTGCGATGAAGTGGGAAAGCTCGAACTGAAAGGTTCAAAGGAATAA
GGGATGGGAAGGATGGAGTATGGATGTAGCAAAGTACTTACTTAGGGGAAATAAAGGTTCTTGGATGGGAAGATGAATATACTGAAGATGGGAAAAGAAAG
AGAAAAGAAAAGAGCAGCTGGTGGGGAGAGCAGGAAAATATGGCAACCAGCTCCTGAGCGGCCTTTCCGGTTTCATACACCGGGCAAAGATCGATCCCCCG
GGCTGCAGGAATTCGATGGCCACTCAGGCCAATTGATTACGGGATCCCATTGGTAACGAAATGTAAAAGCTAGGAGATCGTCCGCCGATGTCAGGATGATT
TCACTTGTTTCTTGTCCGGCTCACCGGTCAAAGCTAAAGAGGAGCAAAAGGAACGGATAGAATCGGGTGCCGCTGATCTATACGGTATAGTGCCCTTATCA
CGTTGACTCAACCCATGCTATTTAACTCAACCCCTCCTTCTGAACCCCACCATCTTCTTCCTTTTCCTCTCATCCCACACAATTCTCTATCTCAGATTTGA
ATTCCAAAAGTCCTCGGACGAAACTGAACAAGTCTTCCTCCCTTCGATAAACCTTTGGTGATTGGAATAACTGACCATCTTCTATAGTTCCCAAACCAACC
GACAATGTAAATACACTCCTCGATTAGCCCTCTAGAGGGCATACGATGGAAGTCATGGAATACTTTTGGCTGGACTCTCACAATGATCAAGGTATCTTAGG
TAACGTCTTTGGCGTGGGCCGGTGTTCGTTCCCAGTCATCGATGCATTCACATGCCCTCCCTAAGCTGGGCCCTAGACTCTAGGATCCTAGTCTAGAAGGA
CATGGCATCGATGGACTGGGTTCGTTCTGAGATTATACGGGTCCCCGTTGAGGGAGCTTCTCAGACCAAGCTGTTGGACCATTTCGGTGGCAAGTGGGACG
AGTTCAAGTTCGCCCCTATCCGCGAAAGCCAGGTCTCTCGTGCCATGACCAGACGTTACTTTGAGGACCTGGACAAGTACGCTGAAAGTGACGTTGTCATT
GTTGGTGCTGGTTCCTGCGGTCTGAGCACTGCGTACGTCTTGGCCAAGGCTCGTCCGGACCTGAAGATTGCTATCGTCGAGGCCAGCGTCTCTCCTGGTCA
GTAGTCCATGATGGATTGCCTTGCACTCAGCTTTCCGGAACTAACGTGCAATAGGTGGCGGTGCCTGGTTGGGTGGCCAACTCTTTTCTGCTATGGTCATG
CGCCGTCCCGCGGAAGTCTTCCTGAACGAGCTGGGTGTTCCTTACGAAGAGGACGCAAACCCCAACTACGTTGTCGTCAAGCACGCCTCCCTGTTTACCTC
GACACTCATGTCGAAGGTTCTCTCCTTCCCCAATGTCAAGCTCTTCAATGCTACCGCTGTTGAGGACTTGATCACCCGTCCGACCGAGAACGGCAACCCCC
AGATTGCTGGTGTTGTCGTCAACTGGACGCTGGTCACCCTTCACCACGATGATCACTCCTGCATGGACCCCAACACTATCAACGCTCCTGTCATCATCAGT
ACCACTGGTCACGATGGGCCATTCGGCGCCTTCGGTGCCATGGTTCTCAGTGGTGTCAAGGCTGCCGAGGAGGCATTGAAGGTGTTCGACGAGCGTCAGCG
CGAGTGTGCTGAGTAAATGACTCACTACCCGAATGGGTTCAGTGCATGAACCGGATTTGTCTTACGGTCTTTGACGATAGGGGAATGATGATTATGTGATA
GTTCTGAGATTTGAATGAACTCGTTAGCTCGTAATCCACATGCATATGTAAATGGCTGTGTCCCGTATGTAACGGTGGGGCATTCTAGAATAATTATGTGT
AACAAGAAAGACAGTATAATACAAACAAAGATGCAAGAGGCCATCTAGGCCATCAAGCTTATCGATACCGTCGACCTCGAGCTCGAGGGGGGGCCCGGTAC
CCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC
AACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG
AAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC
GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG
TCAGAGGTGGCGAAACCCGACAGGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA
AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAA
CGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA
GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

>Construct of pAF2-ALB1: *ALB1/PKSP-AfuSINE2-1a* - sense orientation.

5-ALB1bPmeI-F	CCCGTTTAAACGCGGTTGGATCGAAAGGT	*Pme*I	Tm=65
6-SINE1d-R	ACCGGCTCTATGTACACTTGGTTGCATTGGGAGCGGCCG	afuSINE1 underlined	Tm=75
7-SINE1e-F	CGGCCGCTCCCAATGCAACCAAGTGTACATAGAGCCGGT	afuSINE1 underlined	Tm=59
8-SINE1fPmeI-R	GGGGTTTAAACCCTCTATGTACACTTA	*Pme*I	Tm=38

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATA
AATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACC
GTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA
CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAAC
GACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGATCTAGTTGGCAGTAGGAGGCTGTTC
AGGCGGAGGCGATGAAGCTTTGGGGGCCTCTGACGGTTGTGGCTTGTTGGATTCAACAGGAGGCCGCTTCCCGACATCTGCGGGCTCCTTTGGTGATATTC
GCGGCAACCCTACCTTGTCGGTCTCAATTGTAGCGATTTCTTCGTCCCGCTCAACGTAATCGCCCACCTCTGCCGAATAATGAGTTAAAGATTAGGAGACG
GGTCAAGAGACCTACGCTTCGAGAATTGCTTCAATGTACCCTCGGTGATCGATTCCGCCATCTGAGGGACCTTCACTATGGTATCAGCTGCGGCGAGATCA
GCGAACAACGTCATTGAATCTAAAGCAAATTGCGATACCGTAATTTCTGATCTGATAGCCGCCGAGGCGGGCGAGACTGCCCCGGGGTAAATTGGCTCTAT
GGAATGTCGACGGTCAGCCAGAGCAGCATAATGTGCGCGAACCTCCTTACCCCAAGAACGGAGCATGCAGCAAAACCGCGCCAGACTTATGCCTTCTAAGC

GCCAAAGAGGTGATACTCCTGCGAGCACCAATGCAAGGCAAGCGGGAGCATTGAGCTGTTCGAGGAATAGACCTCAGGTGTAGCTCTGACAGACGGTGAAA
TGGTAACCTAGACGCCATATTTTATCCGATTTGATTGGGTAGGTAATCAGAAGTTAAGGAGCAGGAGGATTTGGCACGATTGTCTGTCCTCAGCAACACAC
GCACATCACCTGCGTTCCATGCGACCGGTGGATGACGGAGTATTATTGTTACGCTTGTTCGGCTTTCAGTAAAGACCAACACGGCAGATTCGGCGCACCCT
ATCAAATTCACTTGGTGCACATCGCAAGAGACACTGCCTTGTAGTACGTGGACAGGATCGTGAAAAATCACTAGAAATTTACAGATTTATTCTTCAGCTAG
CGTCTCTCTCTGCATTCGTTGTGCCAAGTCTCAGGGTGTTCTCCCCTTACGACATTGACAGAGAGTAATCATGTCTGCTAACACTGTTGCGTAAGTTGCTC
TCAGATGGAAGAATAGGCTGCTTTGCTAATCTCAACTTCGCCTTTCAGGTCACTCTACCGACGGAGTCTCAAGCTTGCCCTTGATTGGGCAGTTCACAGAC
ACGTATGGAGAGGACAAGCTGTGTATATTCGGTCCCTGTTTGAGGCAAATAAGGACGTCCGTGATCCTCGTCAACAAAGGGTAAACCCCCTCTGACCCCAG
CAACCGTCTAGCCGTTGCACAACGCCGAGTCATTAGCTTATACTAATGAATTGAGCAGGTGCTACTCCGCGAAACTGAGAAGTCTGATAAGCTAGAAAAAG
GCTGCGTCATGCGCAATTGGCTGATAATGTCTTCCAGGCAGTAAATACGAGAGAAACCTTCCTGCTCGGCAGTTACCTTGTATGGTCACCTTTGATCTGTC
CACTCTGCCTCGATTAACATACGTCTTAGATGCTTCTCAACATGCAGATGCCCACTAAGCTTGGACGGCGTACTAGATCCATGTATTTGCGTCAATTGGAT
GAGACCACAATGAGTTCCCTACTCGAAGATAGACATGGCGTTTAAACGCGGTTGGATCGAAAGGTCTCACCATCAGTGCACCACCCGAAATTCTCGACAAC
TTTATCGAAGAAGGTCTTCCGAAGGAGTACAAACACTTCAAGGCTCCTGGAGTCAGTGGTCCGTACCACGCGCCCCATCTGTACAATGACCGAGAAATTCG
CAATATCCTCAGCTTCTGCTCCGAGGACGTGATTCTGCGCCACACACCACGGGTTCCACTGGTCTCGAGCAACACAGGGAAGCTGGTCCAGGTAAAGAGCA
TGCGTGATCTGCTAAAGGTGGCTCTGGAGGAAATCCTCTTGCGCAAGATCTGCTGGGACAAAGTCACCGAGTCATGCCTTTCCATCGTTCAGGCTACCAAC
GACAAGCCCTGGAGGATTCTCCCTATCGCCAGCAACGCCACGCAAGGCTTGGTTACTGCACTCCAGCGTATGGGAAACTGCCAGATCGAGGTAGACACCGG
GGTCGGCGCTCCTCAAATGGACCCGGCCGCTCCCAATGCAACCAAGTGTACATAGAGCCGGTAGCGGTAGCGTAGTGGTAAGCGCTCCGAGGCAGCCTCTA
GAGAAGTAGGATTGGTGACCCTGCTTTTTTGGAGGTTATGGTTCGATTCCCGTCGCTGGCACAACATTTACCACCACAATGGAAGATCACTTCCCACAAT
GGTATCAAGGCCACTCCCTTATCGCAAGGTGGTGGGGAAGTTGGAACAATCACAGGCCTGTAAGGCGAGGCTCTAAATTCGCCCTCATGTAATGGAACAAA
ATGTAACTAGACACACAAGGATTAGCTATAGTCGATACCTGCATATCGCCCAAGGCGAGGGGTCAGCGTATGA*GTACTACTACTACTACTACTACTAA*GTG
TACATAGAGGAAACAAACGGTGATGTCTGCTCAAGCGGGGTAGCTGTTAGTCAAGCTGCGATGAAGTGGGAAAGCTCGAACTGAAAGGTTCAAAGGAATAA
GGGATGGGAAGGATGGAGTATGGATGTAGCAAAGTACTTACTTAGGGGAAATAAAGGTTCTTGGATGGGAAGATGAATATACTGAAGATGGGAAAAGAAAG
AGAAAAGAAAAGAGCAGCTGGTGGGGAGAGCAGGAAAATATGGCAACCAGCTCCTGAGCGGCCTTTCCGGTTTCATACACCGGGCAAAGATCGATCCCCCG
GGCTGCAGGAATTCGATGGCCACTCAGGCCAATTGATTACGGGATCCCATTGGTAACGAAATGTAAAAGCTAGGAGATCGTCCGCCGATGTCAGGATGATT
TCACTTGTTTCTTGTCCGGCTCACCGGTCAAAGCTAAAGAGGAGCAAAAGGAACGGATAGAATCGGGTGCCGCTGATCTATACGGTATAGTGCCCTTATCA
CGTTGACTCAACCCATGCTATTTAACTCAACCCCTCCTTCTGAACCCCACCATCTTCTTCCTTTTCCTCTCATCCCACACAATTCTCTATCTCAGATTTGA
ATTCCAAAAGTCCTCGGACGAAACTGAACAAGTCTTCCTCCCTTCGATAAACCTTTGGTGATTGGAATAACTGACCATCTTCTATAGTTCCCAAACCAACC
GACAATGTAAATACACTCCTCGATTAGCCCTCTAGAGGGCATACGATGGAAGTCATGGAATACTTTTGGCTGGACTCTCACAATGATCAAGGTATCTTAGG
TAACGTCTTTGGCGTGGGCCGGTGTTCGTTCCCAGTCATCGATGCATTCACATGCCCTCCCTAAGCTGGGCCCTAGACTCTAGGATCCTAGTCTAGAAGGA
CATGGCATCGATGGACTGGGTTCGTTCTGAGATTATACGGGTCCCCGTTGAGGGAGCTTCTCAGACCAAGCTGTTGGACCATTTCGGTGGCAAGTGGGACG
AGTTCAAGTTCGCCCCTATCCGCGAAAGCCAGGTCTCTCGTGCCATGACCAGACGTTACTTTGAGGACCTGGACAAGTACGCTGAAAGTGACGTTGTCATT
GTTGGTGCTGGTTCCTGCGGTCTGAGCACTGCGTACGTCTTGGCCAAGGCTCGTCCGGACCTGAAGATTGCTATCGTCGAGGCCAGCGTCTCTCCTGGTCA
GTAGTCCATGATGGATTGCCTTGCACTCAGCTTTCCGGAACTAACGTGCAATAGGTGGCGGTGCCTGGTTGGGTGGCCAACTCTTTTCTGCTATGGTCATG
CGCCGTCCCGCGGAAGTCTTCCTGAACGAGCTGGGTGTTCCTTACGAAGAGGACGCAAACCCCAACTACGTTGTCGTCAAGCACGCCTCCCTGTTTACCTC
GACACTCATGTCGAAGGTTCTCTCCTTCCCCAATGTCAAGCTCTTCAATGCTACCGCTGTTGAGGACTTGATCACCCGTCCGACCGAGAACGGCAACCCCC
AGATTGCTGGTGTTGTCGTCAACTGGACGCTGGTCACCCTTCACCACGATGATCACTCCTGCATGGACCCCAACACTATCAACGCTCCTGTCATCATCAGT
ACCACTGGTCACGATGGGCCATTCGGCGCCTTCGGTGCCATGGTTCTCAGTGGTGTCAAGGCTGCCGAGGAGGCATTGAAGGTGTTCGACGAGCGTCAGCG
CGAGTGTGCTGAGTAAATGACTCACTACCCGAATGGGTTCAGTGCATGAACCGGATTTGTCTTACGGTCTTTGACGATAGGGGAATGATGATTATGTGATA
GTTCTGAGATTTGAATGAACTCGTTAGCTCGTAATCCACATGCATATGTAAATGGCTGTGTCCCGTATGTAACGGTGGGGCATTCTAGAATAATTATGTGT
AACAAGAAAGACAGTATAATACAAACAAAGATGCAAGAGGCCATCTAGGCCATCAAGCTTATCGATACCGTCGACCTCGAGCTCGAGGGGGGGCCCGGTAC
CCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC
AACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGG
AAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC
GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG
TCAGAGGTGGCGAAACCCGACAGGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCA
AGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATAC
GGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCC
GAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAA
CGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC
CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA
GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

>Construct of pAF3-ALB1: *ALB1/PKSP – GFP – PSKP/1BLA*

5-ALB1bPmeI-F	CCCGTTTAAACGCGGTTGGATCGAAAGGT	*Pme*I
9-ALB1d-R	CAGCTCCTCGCCCTTGCTCACCATGTTGCATTGGGAGCGGCC	GFP underlined
10-GFP1a-F	GGCCGCTCCCAATGCAACATGGTGAGCAAGGGCGAGGAGCTG	GFP underlined
11-GFP1b-R	GGGATCGATAAGTCGTGCTGCTTCATG	*Cla*I
12-ALB1eClaI-F	CCCATCGATCCCGTTGCATTGGGAGCGGCC	*Cla*I
13-ALB1fPmeI-R	GGGGTTTAAACGCGGTTGGATCGAAAGGT	*Pme*I
***AT'CGAT	***Cla*I (5' overhang)**	

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATA
AATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACC
GTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCC
CCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA
CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTG
CGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAAC
GACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGATCTAGTTGGCAGTAGGAGGCTGTTC
AGGCGGAGGCGATGAAGCTTTGGGGGCCTCTGACGGTTGTGGCTTGTTGGATTCAACAGGAGGCCGCTTCCCGACATCTGCGGGCTCCTTTGGTGATATTC
GCGGCAACCCTACCTTGTCGGTCTCAATTGTAGCGATTTCTTCGTCCCGCTCAACGTAATCGCCCACCTCTGCCGAATAATGAGTTAAAGATTAGGAGACG
GGTCAAGAGACCTACGCTTCGAGAATTGCTTCAATGTACCCTCGGTGATCGATTCCGCCATCTGAGGGACCTTCACTATGGTATCAGCTGCGGCGAGATCA
GCGAACAACGTCATTGAATCTAAAGCAAATTGCGATACCGTAATTTCTGATCTGATAGCCGCCGAGGCGGGCGAGACTGCCCCGGGGTAAATTGGCTCTAT
GGAATGTCGACGGTCAGCCAGAGCAGCATAATGTGCGCGAACCTCCTTACCCCAAGAACGGAGCATGCAGCAAAACCGCGCCAGACTTATGCCTTCTAAGC
GCCAAAGAGGTGATACTCCTGCGAGCACCAATGCAAGGCAAGCGGGAGCATTGAGCTGTTCGAGGAATAGACCTCAGGTGTAGCTCTGACAGACGGTGAAA
TGGTAACCTAGACGCCATATTTTATCCGATTTGATTGGGTAGGTAATCAGAAGTTAAGGAGCAGGAGGATTTGGCACGATTGTCTGTCCTCAGCAACACAC
GCACATCACCTGCGTTCCATGCGACCGGTGGATGACGGAGTATTATTGTTACGCTTGTTCGGCTTTCAGTAAAGACCAACACGGCAGATTCGGCGCACCCT
ATCAAATTCACTTGGTGCACATCGCAAGAGACACTGCCTTGTAGTACGTGGACAGGATCGTGAAAAATCACTAGAAATTTACAGATTTATTCTTCAGCTAG
CGTCTCTCTCTGCATTCGTTGTGCCAAGTCTCAGGGTGTTCTCCCCTTACGACATTGACAGAGAGTAATCATGTCTGCTAACACTGTTGCGTAAGTTGCTC
TCAGATGGAAGAATAGGCTGCTTTGCTAATCTCAACTTCGCCTTTCAGGTCACTCTACCGACGGAGTCTCAAGCTTGCCCTTGATTGGGCAGTTCACAGAC
ACGTATGGAGAGGACAAGCTGTGTATATTCGGTCCCTGTTTGAGGCAAATAAGGACGTCCGTGATCCTCGTCAACAAAGGGTAAACCCCCTCTGACCCCAG
CAACCGTCTAGCCGTTGCACAACGCCGAGTCATTAGCTTATACTAATGAATTGAGCAGGTGCTACTCCGCGAAACTGAGAAGTCTGATAAGCTAGAAAAAG
GCTGCGTCATGCGCAATTGGCTGATAATGTCTTCCAGGCAGTAAATACGAGAGAAACCTTCCTGCTCGGCAGTTACCTTGTATGGTCACCTTTGATCTGTC
CACTCTGCCTCGATTAACATACGTCTTAGATGCTTCTCAACATGCAGATGCCCACTAAGCTTGGACGGCGTACTAGATCCATGTATTTGCGTCAATTGGAT
GAGACCACAATGAGTTCCCTACTCGAAGATAGACATGGCGTTTAAACGCGGTTGGATCGAAAGGTCTCACCATCAGTGCACCACCCGAAATTCTCGACAAC
TTTATCGAAGAAGGTCTTCCGAAGGAGTACAAACACTTCAAGGCTCCTGGAGTCAGTGGTCCGTACCACGCGCCCCATCTGTACAATGACCGAGAAATTCG
CAATATCCTCAGCTTCTGCTCCGAGGACGTGATTCTGCGCCACACACCACGGGTTCCACTGGTCTCGAGCAACACAGGGAAGCTGGTCCAGGTAAAGAGCA
TGCGTGATCTGCTAAAGGTGGCTCTGGAGGAAATCCTCTTGCGCAAGATCTGCTGGGACAAAGTCACCGAGTCATGCCTTTCCATCGTTCAGGCTACCAAC
GACAAGCCCTGGAGGATTCTCCCTATCGCCAGCAACGCCACGCAAGGCTTGGTTACTGCACTCCAGCGTATGGGAAACTGCCAGATCGAGGTAGACACCGG
GGTCGGCGCTCCTCAAATGGACCCGGCCGCTCCCAATGCAACATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC
TGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTATCGATGTTG
CATTGGGAGCGGCCGGGTCCATTTGAGGAGCGCCGACCCCGGTGTCTACCTCGATCTGGCAGTTTCCCATACGCTGGAGTGCAGTAACCAAGCCTTGCGTG
GCGTTGCTGGCGATAGGGAGAATCCTCCAGGGCTTGTCGTTGGTAGCCTGAACGATGGAAAGGCATGACTCGGTGACTTTGTCCCAGCAGATCTTGCGCAA
GAGGATTTCCTCCAGAGCCACCTTTAGCAGATCACGCATGCTCTTTACCTGGACCAGCTTCCCTGTGTTGCTCGAGACCAGTGGAACCCGTGGTGTGTGGC
GCAGAATCACGTCCTCGGAGCAGAAGCTGAGGATATTGCGAATTTCTCGGTCATTGTACAGATGGGGCGCGTGGTACGGACCACTGACTCCAGGAGCCTTG
AAGTGTTTGTACTCCTTCGGAAGACCTTCTTCGATAAAGTTGTCGAGAATTTCGGGTGGTGCACTGATGGTGAGACCTTTCGATCCAACCGCAAACAAACG
GTGATGTCTGCTCAAGCGGGGTAGCTGTTAGTCAAGCTGCGATGAAGTGGGAAAGCTCGAACTGAAAGGTTCAAAGGAATAAGGGATGGGAAGGATGGAGT
ATGGATGTAGCAAAGTACTTACTTAGGGGAAATAAAGGTTCTTGGATGGGAAGATGAATATACTGAAGATGGGAAAAGAAAGAGAAAAGAAAAGAGCAGCT
GGTGGGGAGAGCAGGAAAATATGGCAACCAGCTCCTGAGCGGCCTTTCCGGTTTCATACACCGGGCAAAGATCGATCCCCCGGGCTGCAGGAATTCGATGG
CCACTCAGGCCAATTGATTACGGGATCCCATTGGTAACGAAATGTAAAAGCTAGGAGATCGTCCGCCGATGTCAGGATGATTTCACTTGTTTCTTGTCCGG
CTCACCGGTCAAAGCTAAAGAGGAGCAAAAGGAACGGATAGAATCGGGTGCCGCTGATCTATACGGTATAGTGCCCTTATCACGTTGACTCAACCCATGCT
ATTTAACTCAACCCCTCCTTCTGAACCCCACCATCTTCTTCCTTTTCCTCTCATCCCACACAATTCTCTATCTCAGATTTGAATTCCAAAAGTCCTCGGAC
GAAACTGAACAAGTCTTCCTCCCTTCGATAAACCTTTGGTGATTGGAATAACTGACCATCTTCTATAGTTCCCAAACCAACCGACAATGTAAATACACTCC
TCGATTAGCCCTCTAGAGGGCATACGATGGAAGTCATGGAATACTTTTGGCTGGACTCTCACAATGATCAAGGTATCTTAGGTAACGTCTTTGGCGTGGGC
CGGTGTTCGTTCCCAGTCATCGATGCATTCACATGCCCTCCCTAAGCTGGGCCCTAGACTCTAGGATCCTAGTCTAGAAGGACATGGCATCGATGGACTGG
GTTCGTTCTGAGATTATACGGGTCCCCGTTGAGGGAGCTTCTCAGACCAAGCTGTTGGACCATTTCGGTGGCAAGTGGGACGAGTTCAAGTTCGCCCCTAT
CCGCGAAAGCCAGGTCTCTCGTGCCATGACCAGACGTTACTTTGAGGACCTGGACAAGTACGCTGAAAGTGACGTTGTCATTGTTGGTGCTGGTTCCTGCG
GTCTGAGCACTGCGTACGTCTTGGCCAAGGCTCGTCCGGACCTGAAGATTGCTATCGTCGAGGCCAGCGTCTCTCCTGGTCAGTAGTCCATGATGGATTGC
CTTGCACTCAGCTTTCCGGAACTAACGTGCAATAGGTGGCGGTGCCTGGTTGGGTGGCCAACTCTTTTCTGCTATGGTCATGCGCCGTCCCGCGGAAGTCT
TCCTGAACGAGCTGGGTGTTCCTTACGAAGAGGACGCAAACCCCAACTACGTTGTCGTCAAGCACGCCTCCCTGTTTACCTCGACACTCATGTCGAAGGTT
CTCTCCTTCCCCAATGTCAAGCTCTTCAATGCTACCGCTGTTGAGGACTTGATCACCCGTCCGACCGAGAACGGCAACCCCCAGATTGCTGGTGTTGTCGT
CAACTGGACGCTGGTCACCCTTCACCACGATGATCACTCCTGCATGGACCCCAACACTATCAACGCTCCTGTCATCATCAGTACCACTGGTCACGATGGGC
CATTCGGCGCCTTCGGTGCCATGGTTCTCAGTGGTGTCAAGGCTGCCGAGGAGGCATTGAAGGTGTTCGACGAGCGTCAGCGCGAGTGTGCTGAGTAAATG
ACTCACTACCCGAATGGGTTCAGTGCATGAACCGGATTTGTCTTACGGTCTTTGACGATAGGGGAATGATGATTATGTGATAGTTCTGAGATTTGAATGAA
CTCGTTAGCTCGTAATCCACATGCATATGTAAATGGCTGTGTCCCGTATGTAACGGTGGGGCATTCTAGAATAATTATGTGTAACAAGAAAGACAGTATAA
TACAAACAAAGATGCAAGAGGCCATCTAGGCCATCAAGCTTATCGATACCGTCGACCTCGAGCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTA
GTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCA
TAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTG
CATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCG
GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA
GGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA
GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG
AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTG
CAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACA
GGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGC
GGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTGATGT
AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA
AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

Publications

Kanhayuwa, L. and Coutts, R.H.A. Short interspersed nuclear elements (*SINEs*) in the genome of the human pathogenic fungus, *Aspergillus fumigatus*. *PLoSOne*. (Manuscript submmitted).

Kanhayuwa, L., Loizou, I.K., Özkan, S., Gunning, A.P. and Coutts, R.H.A. A novel mycovirus from *Aspergillus fumigatus* contains four unique double stranded RNAs as its genome and is infectious as dsRNA. *Proceedings of the National Academy of Sciences of the Unites State of America*, **112 (29)**, 9100-9105.

Kanhayuwa, L., Spanu, P. and Coutts R.H.A. (2014). A novel mycovirus in the human pathogenic fungus *Aspergillus fumigatus* and its application as a virus-induced gene silencing (VIGS) vector. *FEBS Journal*, **281**, 1-783.

Kozlakidis, Z., Herrero, N., Ozkan, S., **Kanhayuwa, L.**, Jamal, A., Bhatti, M.F. and Coutts, R.H. (2013). Sequence determination of a quadripartite dsRNA virus isolated from *Aspergillus foetidus*. *Achieves of Virology,* **158(1)**, 267-272.

Ali, I., **Kanhayuwa, L.**, Rachdawong, S. and Rakshit, S.K., 2013. Identification, phylogenetic analysis and characterization of obligate halophilic fungi isolated from a man-made solar saltern in Phetchaburi province, Thailand. *Annals of Microbiology*, **63(3)**, 887-895.