Persoonia 45, 2020: 163–176
www.ingentaconnect.com/content/nhn/pimj

RESEARCH ARTICLE

ISSN (Online) 1878-9080
https://doi.org/10.3767/persoonia.2020.45.06

Mating genes in *Calonectria* and evidence for a heterothallic ancestral state

J.Q. Li[1,2,3], B.D. Wingfield[1], M.J. Wingfield[1], I. Barnes[1], A. Fourie[1], P.W. Crous[1,4], S.F. Chen[2,3]

Key words

Cylindrocladium
fungal biology
fungal pathogens
MAT locus
mating type
phylogeny
sexual reproduction

Abstract The genus *Calonectria* includes many important plant pathogens with a wide global distribution. In order to better understand the reproductive biology of these fungi, we characterised the structure of the mating type locus and flanking genes using the genome sequences for seven *Calonectria* species. Primers to amplify the mating type genes in other species were also developed. PCR amplification of the mating type genes and multi-gene phylogenetic analyses were used to investigate the mating strategies and evolution of mating type in a collection of 70 *Calonectria* species residing in 10 *Calonectria* species complexes. Results showed that the organisation of the *MAT* locus and flanking genes is conserved. In heterothallic species, a novel *MAT* gene, *MAT1-2-12* was identified in the *MAT1-2* idiomorph; the *MAT1-1* idiomorph, in most cases, contained the *MAT1-1-3* gene. Neither *MAT1-1-3* nor *MAT1-2-12* was found in homothallic *Calonectria* (*Ca.*) *hongkongensis*, *Ca. lateralis*, *Ca. pseudoturangicola* and *Ca. turangicola*. Four different homothallic *MAT* locus gene arrangements were observed. Ancestral state reconstruction analysis provided evidence that the homothallic state was basal in *Calonectria* and this evolved from a heterothallic ancestor.

Article info Received: 28 October 2019; Accepted: 14 February 2020; Published: 17 March 2020.

INTRODUCTION

Calonectria is an *Ascomycete* genus that accommodates many important plant pathogens having a broad global distribution (Crous 2002, Lombard et al. 2010c). Approximately 335 plant species residing in 100 plant families are hosts to these fungi (Crous 2002, Lombard et al. 2010c). *Calonectria* species reside in two main phylogenetic groups. These are known as the Prolate Group and the Sphaero-Naviculate Group, and they are differentiated based on the shape of the vesicles in their conidiogenous apparatuses (Lombard et al. 2010b, Pham et al. 2019).

Ten species complexes are defined in *Calonectria*. Eight of these are in the Prolate Group, which includes the *Ca. brassicae*, *Ca. candelabrum*, *Ca. colhounii*, *Ca. cylindrospora*, *Ca. mexicana*, *Ca. pteridis*, *Ca. reteaudii* and *Ca. spathiphylli* species complexes. The remaining two species complexes reside in the Sphaero-Naviculate Group and they include the *Ca. kyotensis* and the *Ca. naviculata* species complexes (Lombard et al. 2010b, 2016). To date, 172 *Calonectria* species have been identified based on comparisons of DNA sequence data. Of these, approximately 99 were isolated from diseased tissues and about 73 from soil samples (Lombard et al. 2010b, 2016, Marin-Felix et al. 2017, Crous et al. 2019, Pham et al. 2019).

Both homothallic and heterothallic mating systems have been reported in *Calonectria* spp., but their sexual morphs are rarely seen in nature or in laboratory culture (Crous 2002, Lombard et al. 2010a). This is not unusual given that sexual reproduction is a complex process that is commonly species-specific, and strongly influenced by the environment and the compatibility of isolates (Goodenough & Heitman 2014). Consequently, the absence of sexual structures in *Calonectria* does not preclude the fact that species may be capable of sexual outcrossing (Billiard et al. 2012). This is an important consideration given that sexual reproduction is the dominant mechanism generating genetic diversity, eliminating deleterious mutations, ensuring survival of species and their overall population health (Crow 1994, Gordo & Campos 2008, Lumley et al. 2015).

Ascomycetes have a bipolar mating system that is controlled by mating type (*MAT*) genes at a single *MAT* locus (*MAT1*) with two non-allelic forms referred to as the *MAT1-1* and *MAT1-2* idiomorphs (Turgeon & Yoder 2000). The *MAT1-1* idiomorph is characterised by a *MAT1-1-1* gene, which encodes an alpha box motif protein homologous to MATa1 of *Saccharomyces cerevisiae* (Turgeon & Yoder 2000). The *MAT1-2* idiomorph contains a *MAT1-2-1* gene that encodes a protein with a high mobility group (HMG) domain (Wilson et al. 2015a). Eight additional genes (*MAT1-1-2* to *MAT1-1-9*) have been identified in the *MAT1-1* idiomorph and 10 genes (*MAT1-2-2* to *MAT1-2-11*) in the *MAT1-2* idiomorph (Wilken et al. 2017). These have been named sequentially in the order of their discovery (Wilken et al. 2017). The expression of these genes is most often related to the sexual life cycle of the fungi in which they occur (Ferreira et al. 1998, Kim et al. 2012, Zheng et al. 2013).

In heterothallic *Ascomycetes*, the two opposite mating type idiomorphs exist in different isolates. These individuals are self-sterile and require a compatible partner to mate and produce sexual spores. In contrast, homothallic species are self-fertile, where a single individual possesses both mating type idiomorphs, and can therefore complete the sexual cycle on its own (Ni et al. 2011, Wilson et al. 2015b). Transitions between homothallism and heterothallism are well-known in genera of

[1] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0028, South Africa.
[2] State Key Laboratory of Tree Genetics and Breeding (SKLTGB), Chinese Academy of Forestry (CAF), Haidian District 100091, Beijing, China; corresponding author e-mail: shuaifei.chen@gmail.com.
[3] China Eucalypt Research Centre (CERC), Chinese Academy of Forestry (CAF), ZhanJiang 524022, GuangDong Province, China.
[4] Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands.

the *Ascomycetes* (Labarere & Noel 1992, Lin & Heitman 2007, Ni et al. 2011).

Mating strategy and the ratio of mating type genes are commonly used in population genetics and epidemiology studies of plant pathogens (McDonald & Linde 2002, Alby et al. 2009, Adamson et al. 2018). The *MAT* gene sequences have also been used to track the evolutionary direction of mating systems based on thallism and molecular phylogenies (James et al. 2006, Fraser et al. 2007, Nagel et al. 2018). These genes can be used as molecular markers to establish species boundaries and to delimitate cryptic species (O'Donnell et al. 2004, Lopes et al. 2017). Mating strategies have consequently served as important criteria in the taxonomy of *Calonectria* (Schoch et al. 1999, Lombard et al. 2010a). Similarly, using genome sequences and PCR amplification of *MAT* genes, populations of *Calonectria* species have been defined based on their mating type (Malapi-Wight et al. 2014, 2019). For example, Malapi-Wight et al. (2019) showed in a collection from four continents, that all isolates of *Ca. henricotiae* were *MAT1-1* whereas all isolates of *Ca. pseudonaviculata* were *MAT1-2*.

Some studies have considered the mating types of *Calonectria* spp., however, sexual reproduction is still not well understood in this genus. For example, it is not known which *MAT* genes occur at the *MAT* loci of homothallic *Calonectria* species, how they are arranged, or whether there is significant conservation of *MAT* genes or gene sequences at these loci. Universal mating type markers for *MAT1-1* idiomorph are not available to enable easy detection of the thallism in *Calonectria* species, although *MAT1-2-1* gene markers were designed for *Calonectria* by Schoch et al. (2000). In addition, nothing is known regarding the evolution of the mating systems in *Calonectria* and the probable ancestral state (homothallism or heterothallism) has not been determined.

An important basis to control the spread and prevalence of plant pathogens is to understand their life cycles and modes of reproduction. In order to further understand the possible role of sexual reproduction in *Calonectria*, we identified and characterised the *MAT* loci and flanking genes of seven species of *Calonectria* using whole genome sequences. Mating type primers were then designed to consider the mating strategies of 65 *Calonectria* species from 10 *Calonectria* species complexes. The data were also used to consider the evolutionary history of mating in the genus.

MATERIALS AND METHODS

Isolates, DNA extraction and identification

A total of 123 isolates, representing 65 *Calonectria* species residing in 10 *Calonectria* species complexes (Lombard et al. 2010b, 2016) were utilised in this study (Table 1). Two isolates were acquired from the culture collection of the China Eucalypt Research Centre (CERC), Chinese Academy of Forestry (CAF); 32 from the culture collection (CBS) of the Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands and 89 from the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa. Cultures were incubated and maintained on 2 % malt extract agar (MEA) at room temperature.

All cultures were purified using single hyphal tip transfers to ensure that they represented a single genotype. After three to five days of growth on MEA, the mycelium was harvested and genomic DNA was extracted using Prepman™ Ultra Sample Preparation Reagent (Thermo Fisher Scientific, Waltham, MA, USA) following a protocol described by Duong et al. (2012). DNA concentrations were determined using a NanoDrop ND-2000 spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) and diluted to 25–50 ng/µL using sterile distilled water.

The translation elongation factor 1-alpha (*tef1*) gene region was amplified for all 123 *Calonectria* isolates using the primers and protocols described by Lombard et al. (2016). Amplification reactions were conducted in 25 µL reaction volumes consisting of 12.5 µL 2 × TopTaq™ Master Mix (Qiagen Inc., Hilden, Germany), 1 µL of each of the two primers (10 mM), 2 µL genomic DNA and 8.5 µL sterile distilled water. The PCR products were visualized under UV light after 2 % agarose gel electrophoresis with 3 % SYBR Safe DNA gel stain (Thermo Fisher Scientific Inc., USA). Amplicons were sequenced in both directions using the same primers used for PCR amplification by the Beijing Genomics Institution, Guangzhou, China. The sequences were edited and assembled using Geneious v. 7.0 (Kearse et al. 2012). The *tef1* sequences were used to confirm the identification of isolates based on a pairwise similarity comparison with sequences published on NCBI (https://guides.lib.berkeley.edu/ncbi/blast).

Analysis of the MAT loci in seven Calonectria species and primer design

Genome sequences

The genome sequences of seven *Calonectria* species (eight isolates) were used to analyse the *MAT* locus. Three of the genomes were sequenced in this study. This included one isolate of *Ca. hongkongensis* (CMW 47271) that is self-fertile and resides in the Sphaero-Naviculate Group of *Calonectria* (Crous et al. 2004, Lombard et al. 2010b, Li et al. 2017) and two isolates of *Ca. pauciramosa* (CMW 5683 and CMW 7592) known to be self-sterile, of opposite mating type, and which reside in the Prolate Group of *Calonectria* (Lombard et al. 2010a, b). Genomic DNA was extracted using the phenol/chloroform method described by Goodwin et al. (1992). Pair-end libraries (350 bp average insert size) and mate pair libraries (5 000 bp average insert size) for CMW 47271 and CMW 5683, as well as pair-end libraries (350 bp average insert size) for CMW 7592, were prepared and sequenced using the Illumina HiSeq 2500 platform. Quality control procedures on the raw sequencing reads, and the removal of adapters, were done using Trimmomatic v. 0.36 (Bolger et al. 2014). Genome assembly, assembly of contigs into scaffolds and gap filling were conducted as described by Duong et al. (in Wingfield et al. 2016) for the genome assembly of CMW 2644 (*Grosmannia penicillata*). The completeness of assembly was evaluated with BUSCO v. 3 (https://busco.ezlab.org/) using the *Sordariomycetes* odb9 dataset (Simão et al. 2015). All three genomic sequences were deposited in GenBank.

Sequences for the other five species, including *Ca. henricotiae* (CBS 138102), *Ca. leucothoes* (CBS 109166), *Ca. naviculata* (CBS 101121), *Ca. pseudonaviculata* (CBS 139394) and *Ca. pseudoreteaudii* (YA51), were obtained from public genomic databases at NCBI with accession numbers PGWR00000000, NAJI00000000, NAGG00000000, JYJY00000000 and MOCD00000000, respectively (Malapi-Wight et al. 2016a, b, Ye et al. 2017). All additional available genome sequences for *Calonectria* spp. published to date (Malapi-Wight et al. 2016a, b, 2019, Ye et al. 2017, LeBlanc et al. 2019) were also screened for inclusion in this study of the mating type locus. These included three genome sequences of *Ca. henricotiae* (CB077, NL009 and NL017) with NCBI accession numbers PGSE00000000, PGSF00000000 and PHMY00000000, respectively, and seven genome sequences of *Ca. pseudonaviculata* (CB002, CBS 114417, CBS 139395, CT13, ICMP 14368, NC-BB1 and ODA1) with NCBI accession numbers RQSK00000000, PHMX00000000, PGGA00000000, PGWW00000000, PHNA00000000, PHMZ00000000 and PHNB00000000, respectively. All three genome sequences of *Ca. henricotiae* harboured the same *MAT1-1* idiomorph as the

Table 1 Species of *Calonectria* used in this study.

Species	Isolate number[1]	Host	Origin	Thallism[2]	Mating type	GenBank accession No.[3]							
						MAT1-1-1	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*tub2*	*cmdA*	*his3*	*tef1*
Ca. acaciicola	**CBS 143557**[4,5]; CMW 47173	Soil in *Acacia auriculiformis* plantation	Nghe An, Vietnam	P_HE	MAT1-1	MN959486	No[6]	No	No	MH119285	MH119252	MH119186	MH119219
	CBS 143558; CMW 47174	Soil in *A. auriculiformis* plantation	Nghe An, Vietnam	P_HE	MAT1-1	MN959487	No	No	No	MH119286	MH119253	MH119187	MH119220
Ca. aciculata	**CBS 142883**[5]; CMW 47645; CERC 5342	*Eucalyptus urophylla* × *E. grandis* leaf	YunNan, China	HO	homothallic	MN959488	MN959560	MN959612	MN959697	MF442989	MF442874	MF442759	MF442644
Ca. aeknauliensis	**CBS 143559**[5]; CMW 48253	Soil in *Eucalyptus* plantation	North Sumatra, Indonesia	P_HE	MAT1-2	No	No	MN959613	No	–[7]	MH119259	MH119193	MH119226
	CBS 143560; CMW 48254	Soil in *Eucalyptus* plantation	North Sumatra, Indonesia	P_HE	MAT1-2	No	No	MN959614	No	–	MH119260	MH119194	MH119227
Ca. amazonica	CBS 115486; CMW 51223; CPC 3894	*E. tereticornis*	Brazil	HE	MAT1-2	No	No	MN959615	No	KX784611	KX784554	–	KX784681
	CBS 116250[5]; CMW 51234; CPC 3534	*E. tereticornis*	Brazil	HE	MAT1-1	MN959489	MN959561	No	No	KX784612	KX784555	–	KX784682
Ca. arbusta	**CBS 136079**[5]; CMW 31370; CERC 1705	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959490	MN959562	MN959616	No	KJ462904	KJ463018	KJ463135	KJ462787
	CBS 136098; CMW 37981; CERC 1944; CPC 23519	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959491	MN959563	MN959617	No	–	KJ463019	KJ463136	KJ462788
Ca. auriculiformis	**CBS 143561**[5]; CMW 47178	Soil in *A. auriculiformis* plantation	Thanh Hoa, Vietnam	P_HE	MAT1-2	No	No	MN959618	MN959698	MH119287	MH119254	MH119188	MH119221
	CBS 143562; CMW 47179	Soil in *A. auriculiformis* plantation	Thanh Hoa, Vietnam	P_HE	MAT1-2	No	No	MN959619	MN959699	MH119288	MH119255	MH119189	MH119222
Ca. baviensis	**CBS 143563**[5]; CMW 47410	*E. urophylla* leaf	Hanoi, Vietnam	P_HE	MAT1-1	MN959492	No	No	No	MH119289	MH119256	MH119190	MH119223
	CBS 143564; CMW 47433	*E. pellita* leaf	Hanoi, Vietnam	P_HE	MAT1-1	MN959493	No	No	No	MH119290	MH119257	MH119191	MH119224
Ca. blephiliae	**CBS 136425**[5]; CMW 51321; CPC 21859	*Blephilia ciliata* stem	North Carolina, USA	P_HE	MAT1-1	MN959494	No	No	No	KF777246	–	–	KF777243
Ca. brachiatica	**CBS 123700**[5]; CMW 25298	*Pinus maximinoi*	Buga, Colombia	P_HE	MAT1-2	No	No	MN959620	MN959700	FJ696388	GQ267366	FJ696396	GQ267296
	CMW 25302	*P. tecunumanii*	Buga, Colombia	P_HE	MAT1-2	No	No	MN959621	MN959701	FJ716708	GQ267365	FJ716712	GQ267295
	CMW 25307	*P. tecunumanii*	Buga, Colombia	P_HE	MAT1-2	No	No	MN959622	MN959702	FJ716709	GQ267366	FJ716713	GQ267296
Ca. brasiliana	**CBS 111484**[5]; CMW 51187; CPC 1924	Soil	Brazil	P_HE	MAT1-2	No	No	MN959623	MN959703	KX784616	KX784559	–	KX784686
	CBS 111485; CMW 51188; CPC 1929	Soil	Brazil	P_HE	MAT1-2	No	No	MN959624	MN959704	KX784617	KX784560	–	KX784687
Ca. brasiliensis	**CBS 230.51**[5]; CMW 23670; CPC 2390; CMW 51160	*Eucalyptus* sp.	Brazil	P_HE	MAT1-1	MN959495	MN959564	No	No	GQ267241	GQ267421	GQ267259	GQ267328
Ca. brevistipitata	CBS 110837; CMW 51163; CPC 913	Soil	Mexico	HE	MAT1-2	No	No	MN959625	MN959705	KX784621	KX784563	–	KX784691
	CBS 110928; CMW 51170; CPC 951	Soil	Mexico	HE	MAT1-1	MN959496	MN959565	No	No	KX784622	KX784564	–	KX784692
	CBS 115671[5]; CMW 51226; CPC 949	Soil	Mexico	HE	MAT1-1	MN959497	MN959566	No	No	KX784623	KX784565	–	KX784693
Ca. bumicola	**CBS 143575**[5]; CMW 48257	Soil in *Eucalyptus* plantation	North Sumatra, Indonesia	HO	homothallic	MN959498	MN959567	MN959626	No	–	MH119271	MH119205	MH119238
Ca. candelabra	CMW 31000[5]; CPC 1675	*Eucalyptus* sp.	Brazil	HE	MAT1-1	MN959499	MN959568	No	No	FJ972426	GQ267367	FJ972476	FJ972525
	CMW 31001; CPC 1679	*Eucalyptus* sp.	Brazil	HE	MAT1-2	No	No	MN959627	MN959706	GQ421779	GQ267368	GQ267246	GQ267298
Ca. clavata	**CBS 114557**[5]; CMW 23690; CPC 2536	*Callistemon viminalis*	USA	HE	MAT1-1	MN959500	MN959569	No	No	AF333396	GQ267377	DQ190623	GQ267305
	CBS 114666; CMW 30994; CPC 2537	Root debris in peat	USA	HE	MAT1-2	No	No	MN959628	MN959707	DQ190549	GQ267378	DQ190624	GQ267306
Ca. colombiana	CBS 115638[5]; CMW 30766; CPC 1161	Soil	Colombia	P_HE	MAT1-1	MN959501	MN959570	No	No	FJ972422	GQ267456	FJ972441	FJ972491
Ca. colombiensis	CBS 112221[5]; CMW 30985; CPC 724	*E. grandis*	Colombia	HO	homothallic	MN959502	MN959571	MN959629	No	AY725620	AY725749	AY725663	AY725712
Ca. crousiana	CBS 127199[5]; CMW 27253	*E. grandis*	FuJian, China	HO	homothallic	MN959503	MN959572	MN959630	MN959708	HQ285795	MF527085	HQ285809	HQ285823
Ca. curvispora	**CBS 116159**[5]; CMW 23693; CPC 765	Soil	Tamatave, Madagascar	P_HE	MAT1-1	MN959504	MN959573	No	No	AF333395	GQ267374	AY725664	GQ267302
Ca. densa	**CBS 125261**[5]; CMW 31182	Soil	Pichincha, Ecuador	P_HE	MAT1-1	MN959505	MN959574	No	No	GQ267232	GQ267444	GQ267281	GQ267352
Ca. ericae	CBS 114456; CMW 51209; CPC 1984	*Erica capensis*	California, USA	P_HE	MAT1-2	No	No	MN959631	MN959709	KX784627	KX784569	–	KX784697
	CBS 114457; CMW 51210; CPC 1985	*Erica capensis*	California, USA	P_HE	MAT1-2	No	No	MN959632	MN959710	KX784628	KX784570	–	KX784698
	CBS 114458[5]; CMW 51211; CPC 2019	*Erica capensis*	California, USA	P_HE	MAT1-2	No	No	MN959633	MN959711	KX784629	KX784571	–	KX784699
Ca. eucalypti	**CBS 125275**[5]; CMW18444	*E. grandis* leaf	Sumatra Utara, Indonesia	HO	homothallic	MN959506	MN959575	MN959634	MN959712	GQ267218	GQ267430	GQ267267	GQ267338
	CBS 125276; CMW 18445	*E. grandis* leaf	Sumatra Utara, Indonesia	HO	homothallic	MN959507	MN959576	MN959635	MN959713	GQ267219	GQ267431	GQ267268	GQ267339

Table 1 (cont.)

Species	Isolate number[1]	Host	Origin	Thallism[2]	Mating type	GenBank accession No.[3]							
						MAT1-1-1	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*tub2*	*cmdA*	*his3*	*tef1*
Ca. expansa	**CBS 136247**[5]; CMW 31392; CERC 1727	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959508	MN959577	MN959636	No	KJ462914	KJ463029	KJ463146	KJ462798
Ca. foliicola	**CBS 136641**[5]; CMW 31393; CERC 1728	*E. urophylla* × *E. grandis* leaf	Guangxi, China	P_HE	MAT1-2	No	No	MN959637	MN959714	KJ462916	KJ463031	KJ463148	KJ462800
Ca. fujianensis	CBS 127200; CMW 27254	*E. grandis* leaf in plantation	FuJian, China	HO	homothallic	MN959509	MN959578	MN959638	MN959715	HQ285791	MF527088	HQ285805	HQ285819
	CBS 127201[5]; CMW 27257	*E. grandis* leaf in plantation	FuJian, China	HO	homothallic	MN959510	MN959579	MN959639	MN959716	HQ285792	MF527089	HQ285806	HQ285820
Ca. gracilis	CBS 111284; CMW 51175	Soil	Brazil	HO	homothallic	MN959511	No	MN959640	MN959717	DQ190567	GQ267408	DQ190647	GQ267324
	CBS 111807[5]; CMW 51189	*Manilkara zapota*	Brazil	HO	homothallic	MN959512	No	MN959641	MN959718	AF232858	GQ267407	DQ190646	GQ267323
Ca. guangxiensis	**CBS 136092**[5]; CMW 35409; CERC 1900; CPC 23506	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959513	MN959580	MN959642	No	KJ462919	KJ463034	KJ463151	KJ462803
	CBS 136094; CMW 35411; CERC 1902; CPC 23507	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959514	MN959581	MN959643	No	KJ462920	KJ463035	–	KJ462804
*Ca. henricotiae**-1	**CBS 138102**[5,8]	*Buxus sempervirens*	Lokeren, East Flanders, Belgium	HE	MAT1-1					JX535308	KF815157	KF815185	–
Ca. heveicola	CBS 143571[5]; CMW 49928	Soil	Binh Phuoc, Vietnam	P_HE	MAT1-2	No	No	MN959644	No	MH119296	MH119267	MH119201	MH119234
	CBS 143572; CMW 49935	Soil	Binh Phuoc, Vietnam	P_HE	MAT1-2	No	No	MN959645	No	MH119297	MH119268	MH119202	MH119235
Ca. honghensis	CBS 142884; CMW 47668; CERC 5571	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959515	MN959582	MN959646	MN959719	MF442996	MF442894	MF442779	MF442664
	CBS 142885[5]; CMW 47669; CERC 5572	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959516	MN959583	MN959647	MN959720	MF442997	MF442895	MF442780	MF442665
Ca. hongkongensis	**CBS 114828**[5]; CMW 51217; CPC 4670	Soil	Hong Kong	HO	homothallic	MN959517	No	MN959648	No	AY725622	AY725755	AY725667	AY725717
*Ca. hongkongensis**-2	CMW 47271; CERC 3570	Soil in *Eucalyptus* plantation	GuangXi, China	HO	homothallic	MN959518	No	MN959649	No	MF443001	MF442899	MF442784	MF442669
	CMW 47499; CERC 7132	Soil	FuJian, China	HO	homothallic	MN959519	No	MN959650	No	MF443004	MF442902	MF442787	MF442672
Ca. indonesiae	**CBS 112823**[5]; CMW 23683; CPC 4508	Soil	Warambunga, Indonesia	P_HE	MAT1-2	No	No	MN959651	No	AY725623	AY725756	AY725668	AY725718
Ca. lantauensis	CBS 142887; CMW 47251; CERC 3301	Soil	Hong Kong, China	P_HE	MAT1-2	No	No	MN959652	No	–	MF442906	MF442791	MF442676
	CBS 142888[5]; CMW 47252; CERC 3302	Soil	Hong Kong, China	P_HE	MAT1-2	No	No	MN959653	No	–	MF442907	MF442792	MF442677
Ca. lateralis	**CBS 136629**[5]; CMW 31412; CERC 1747	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959520	No	MN959654	No	KJ462955	KJ463070	KJ463186	KJ462840
Ca. lauri	**CBS 749.70**[5]; CMW 23682	*Llex aquifolium*	Netherlands	P_HE	MAT1-1	MN959521	No	No	No	GQ267210	GQ267388	GQ267250	GQ267312
*Ca. leucothoes**-3	**CBS 109166**[5,8]; CMW 30977	*Leucothoe axillaris* leaf	Florida, USA	HE	MAT1-2					FJ918508	GQ267392	FJ918523	FJ918553
Ca. lichi	**CERC 8866**[5]; CGMCC3.18733	Soil	HeNan, China	HO	homothallic	MN959522	MN959584	MN959655	MN959721	MF527097	MF527071	MF527055	MF527039
	CERC 8890; CGMCC3.18734	Soil	HeNan, China	HO	homothallic	MN959523	MN959585	MN959656	MN959722	MF527099	MF527073	MF527057	MF527041
Ca. malesiana	CBS 112710; CMW 51199; CPC 3899	Leaf litter	Thailand	P_HE	MAT1-1	MN959524	MN959586	No	No	AY725626	AY725759	AY725671	AY725721
	CBS 112752[5]; CMW 23687; CPC 4223	Soil	Indonesia	P_HE	MAT1-1	MN959525	MN959587	No	No	AY725627	AY725760	AY725672	AY725722
Ca. mossambicensis	**CBS 137243**[5]; CMW 36327	*E. grandis* × *E. camaldulensis* cutting	Manica, Mozambmbique	P_HE	MAT1-2	No	No	MN959657	MN959723	–	JX570722	JX570726	JX570718
	CMW 36329	*E. grandis* and *E. urophylla* cutting	Zambézia, Mozambmbique	P_HE	MAT1-2	No	No	MN959658	MN959724	–	JX570721	JX570725	JX570717
*Ca. naviculata**-4	**CBS 101121**[5,8]; CMW 30974	Leaf litter	Joao Pessoa, Brazil	HE	MAT1-1					GQ267211	GQ267399	GQ267252	GQ267317
Ca. orientalis	CBS 125259; CMW 20273	Soil	Teso East, Indonesia	P_HE	MAT1-1	MN959526	MN959588	No	No	GQ267237	GQ267449	GQ267286	GQ267357
	CBS 125260[5]; CMW 20291	Soil	Lagan, Indonesia	P_HE	MAT1-1	MN959527	MN959589	No	No	GQ267236	GQ267448	GQ267285	GQ267356
Ca. ovata	CBS 111299[5]; CMW 16724	*E. tereticornis*	Tucuruí, Para, Brazil	HE	MAT1-2	No	No	MN959659	No	GQ267212	GQ267400	GQ267253	GQ267318
	CBS 111307; CMW 30979	*E. tereticornis*	Tucuruí, Para, Brazil	HE	MAT1-1	MN959528	No	No	No	AF210868	GQ267401	GQ267254	GQ267319
Ca. papillata	CBS 136096; CMW 37972; CERC 1935; CPC 23515	Soil in *Eucalyptus* plantation	Guangdong, China	P_HE	MAT1-1	MN959529	No	No	No	KJ462963	KJ463078	KJ463194	KJ462848
	CBS 136097[5]; CMW 37976; CERC 1939; CPC 23517	Soil in *Eucalyptus* plantation	Guangdong, China	P_HE	MAT1-1	MN959530	No	No	No	KJ462964	KJ463079	KJ463195	KJ462849
Ca. parakyotensis	**CBS 136085**[5]; CMW 35169; CERC 1845	Soil in *Eucalyptus* plantation	Guangdong, China	HO	homothallic	MN959531	MN959590	MN959660	No	–	KJ463081	KJ463197	KJ462851
*Ca. pauciramosa**-5	**CBS 138824**[5]; CMW 5683; CPC 971	*E. grandis*	South Africa	HE	MAT1-2	No	No	MN959661	MN959725	FJ918514	GQ267405	FJ918531	FJ918565
	CMW 2151	*E. nitens*	South Africa	HE	MAT1-2	No	No	MN959662	MN959726	FJ972400	–	FJ972468	FJ972517
*Ca. pauciramosa**-6	CMW 7592	*E. grandis*	Uruguay	HE	MAT1-1	MN959532	MN959591	No	No	FJ972380	–	FJ972447	FJ972497
	CMW 9151	*A. mearnsii*	South Africa	HE	MAT1-2	No	No	MN959663	MN959727	FJ972384	–	FJ972451	FJ972501
	CMW 30823; CPC 416	*E. grandis*	South Africa	HE	MAT1-1	MN959533	MN959592	No	No	FJ918515	GQ267404	FJ918532	FJ918566
	CMW 30875; CPC 415	*Eucalyptus* sp.	South Africa	HE	MAT1-1	MN959534	MN959593	No	No	FJ972390	–	FJ972457	FJ972507

Table 1 (cont.)

Species	Isolate number[1]	Host	Origin	Thallism[2]	Mating type	GenBank accession No.[3]							
						MAT1-1-1	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*tub2*	*cmdA*	*his3*	*tef1*
Ca. pentaseptata	**CBS 133349**[5]; CMW 51318	*Eucalyptus* hybrid	Bavi, Hanoi, Vietnam	P_HE	MAT1-1	MN959535	MN959594	No	No	JX855942	–	JX855946	JX855958
	CBS 133351; CMW 51319	*Macadamia* sp.	Bavi, Hanoi, Vietnam	P_HE	MAT1-1	MN959536	MN959595	No	No	JX855944	–	JX855948	JX855960
Ca. plurilateralis	**CBS 111401**[5]; CMW 51178; CPC 1637	Soil	Ecuador	P_HE	MAT1-2	No	No	MN959664	MN959728	KX784648	KX784586	–	KX784719
Ca. polizzii	**CBS 123402**[5]; CMW 51312	*Arbutus unedo*	Sicily, Italy	HE	MAT1-1	MN959537	MN959596	No	No	FJ972419	–	FJ972438	FJ972488
	CBS 125270; CMW 7804; CPC 2681	*Callistemon citrinus*	Sicily, Italy	HE	MAT1-1	MN959538	MN959597	No	No	FJ972417	GQ267461	FJ972436	FJ972486
	CBS 125271; CMW 10151; CPC 2771	*Arbutus unedo*	Sicily, Italy	HE	MAT1-2	No	No	MN959665	MN959729	FJ972418	GQ267462	FJ972437	FJ972487
Ca. pseudocolhounii	**CBS 127195**[5]; CMW 27209	*E. dunnii* leaf in plantation	FuJian, China	HO	homothallic	MN959539	MN959598	MN959666	MN959730	HQ285788	MF527091	HQ285802	HQ285816
	CBS 127196; CMW 27213	*E. dunnii* leaf in plantation	FuJian, China	HO	homothallic	MN959540	MN959599	MN959667	MN959731	HQ285789	MF527092	HQ285803	HQ285817
Ca. pseudoecuadoriae	CBS 111412[5]; CMW 51180; CPC 1648	Soil	Ecuador	P_HE	MAT1-2	No	No	MN959668	MN959732	DQ190601	KX784590	–	KX784724
Ca. pseudomexicana	**CBS 130354**[5]; CMW 51313	*Callistemon* sp. (rouge)	Carthage, Tunis, Tunisia	P_HE	MAT1-2	No	No	MN959669	MN959733	JN607281	–	JN607266	JN607296
	CBS 130355; CMW 51314	*Callistemon* sp. (rouge)	Carthage, Tunis, Tunisia	P_HE	MAT1-2	No	No	MN959670	MN959734	JN607282	–	JN607267	JN607297
Ca. pseudonaviculata[*-7]	CBS 139394[5,8]	*Sarcococca hookeriana*	Maryland, USA	HE	MAT1-2					KR011242	–	–	–
Ca. pseudopteridis	**CBS 163.28**[5]; CMW 51159	*Washingtonia robusta*	USA	P_HE	MAT1-1	MN959541	MN959600	No	No	–	KM396076	–	KM395902
Ca. pseudoreteaudii[*-8]	YA51[5,8]	*Eucalyptus* sp.	Fujian, China	HE	MAT1-2					–	–	–	–
Ca. pseudoscoparia	CBS 125255; CMW 15215	*E. grandis*	Pichincha, Ecuador	P_HE	MAT1-2	No	No	MN959671	MN959735	GQ267227	GQ267439	GQ267276	GQ267347
	CBS 125257[5]; CMW 15218	*E. grandis*	Pichincha, Ecuador	P_HE	MAT1-2	No	No	MN959672	MN959736	GQ267229	GQ267441	GQ267278	GQ267349
Ca. pseudoturangicola	**CBS 142890**[5]; CMW 47496; CERC 7126	Soil	FuJian, China	HO	homothallic	MN959542	No	MN959673	No	MF443080	MF442980	MF442865	MF442750
	CBS 142891; CMW 47497; CERC 7127	Soil	FuJian, China	HO	homothallic	MN959543	No	MN959674	No	MF443081	MF442981	MF442866	MF442751
Ca. pseudouxmalensis	CBS 110923; CMW 51165; CPC 941	Soil	Mexico	P_HE	MAT1-2	No	No	MN959675	MN959737	KX784653	–	–	KX784725
	CBS 110924[5]; CMW 51166; CPC 942	Soil	Mexico	P_HE	MAT1-2	No	No	MN959676	MN959738	KX784654	–	–	KX784726
	CBS 115677; CMW 51228; CPC 943	Soil	Mexico	P_HE	MAT1-2	No	No	MN959677	MN959739	KX784655	–	–	KX784727
Ca. pseudoyunnanensis	**CBS 142892**[5]; CMW 47655; CERC 5376	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959544	MN959601	MN959678	No	MF443083	MF442983	MF442868	MF442753
	CBS 142893; CMW 47656; CERC 5377	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959545	MN959602	MN959679	No	MF443084	MF442984	MF442869	MF442754
	CBS 142894; CMW 47657; CERC 5378	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959546	MN959603	MN959680	No	MF443085	MF442985	MF442870	MF442755
Ca. putriramosa	**CBS 111449**[5]; CMW 51181; CPC 1951	*Eucalyptus* cutting	Brazil	P_HE	MAT1-2	No	No	MN959681	MN959740	KX784656	KX784591	–	KX784728
	CBS 111470; CMW 51182; CPC 1940	Soil	Brazil	P_HE	MAT1-2	No	No	MN959682	MN959741	KX784657	KX784592	–	KX784729
	CBS 111477; CMW; 51183; CPC 1928	Soil	Brazil	P_HE	MAT1-2	No	No	MN959683	MN959742	KX784658	KX784593	–	KX784730
	CBS 116076; CMW 51230; CPC 604	*Eucalyptus* cutting	Brazil	P_HE	MAT1-2	No	No	MN959684	MN959743	–	–	–	KX784731
Ca. seminaria	**CBS 136632**[5]; CMW 31450; CERC 1785; CPC 23488	*E. urophylla* × *E. grandis* seedling leaf	Guangdong, China	P_HE	MAT1-2	No	No	MN959685	MN959744	KJ462998	KJ463115	KJ463231	KJ462885
	CBS 136639; CMW 31489; CERC 1824	*E. urophylla* × *E. grandis* seedling leaf	Guangdong, China	P_HE	MAT1-2	No	No	MN959686	MN959745	KJ462999	KJ463116	KJ463232	KJ462886
Ca. sphaeropedunculata	**CBS 136081**[5]; CMW 31390; CERC 1725	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959547	MN959604	MN959687	No	KJ463003	KJ463120	KJ463236	KJ462890
Ca. sulawesiensis	CBS 125253; CMW 14879	*Eucalyptus* sp.	Sulawesi, Indonesia	P_HE	MAT1-1	MN959548	No	No	No	GQ267222	GQ267434	GQ267271	GQ267342
	CBS 125277[5]; CMW 14878	*Eucalyptus* sp.	Sulawesi, Indonesia	P_HE	MAT1-1	MN959549	No	No	No	GQ267220	GQ267432	GQ267269	GQ267340
Ca. sumatrensis	**CBS 112829**[5]; CMW 23698; CPC4518	Soil	Indonesia	P_HE	MAT1-1	MN959550	MN959605	No	No	AY725649	AY725771	AY725696	AY725733
	CBS 112934; CMW 30987; CPC 4516	Soil	Indonesia	P_HE	MAT1-1	MN959551	MN959606	No	No	AY725651	AY725773	AY725698	AY725735
Ca. terrestris	**CBS 136642**[5]; CMW 35180; CERC 1856	Soil in *Eucalyptus* plantation	Guangdong, China	P_HE	MAT1-2	No	No	MN959688	MN959746	KJ463004	KJ463121	KJ463237	KJ462891

Table 1 (cont.)

Species	Isolate number[1]	Host	Origin	Thallism[2]	Mating type	GenBank accession No.[3]							
						MAT1-1-1	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*tub2*	*cmdA*	*his3*	*tef1*
Ca. terrestris (cont.)	CBS 136645; CMW 35178; CERC 1854	Soil in *Eucalyptus* plantation	Guangdong, China	P_HE	MAT1-2	No	No	MN959689	MN959747	KJ463007	KJ463124	KJ463240	KJ462894
Ca. tetraramosa	**CBS 136635**[5]; CMW 31474; CERC 1809; CPC 23489	*E. urophylla* × *E. grandis* seedling leaf	Guangdong, China	P_HE	MAT1-2	No	No	MN959690	MN959748	KJ463011	KJ463128	KJ463244	KJ462898
	CBS 136637; CMW 31476; CERC 1811	*E. urophylla* × *E. grandis* seedling leaf	Guangdong, China	P_HE	MAT1-2	No	No	MN959691	MN959749	KJ463012	KJ463129	KJ463245	KJ462899
Ca. tonkinensis	**CBS 143576**[5]; CWM 47430	Soil in *Eucalyptus* plantation	Hanoi, Vietnam	P_HE	MAT1-1	MN959552	No	No	No	MH119291	MH119258	MH119192	MH119225
Ca. turangicola	**CBS 136077**[5]; CMW 31411; CERC 1746; CPC 23479	Soil in *Eucalyptus* plantation	Guangxi,China	HO	homothallic	MN959553	No	MN959692	No	KJ463013	–	KJ463246	KJ462900
	CBS 136093; CMW 35410; CERC 1901	Soil in *Eucalyptus* plantation	Guangxi, China	HO	homothallic	MN959554	No	MN959693	No	KJ463014	KJ463130	KJ463247	KJ462901
Ca. vegrandis	**CBS 143565**[5]; CMW 48245	Soil in *Eucalyptus* plantation	North Sumatra, Indonesia	P_HE	MAT1-1	MN959555	MN959607	No	No	–	MH119261	MH119195	MH119228
	CBS 143566; CMW 48246	Soil in *Eucalyptus* plantation	North Sumatra, Indonesia	P_HE	MAT1-1	MN959556	MN959608	No	No	–	MH119262	MH119196	MH119229
Ca. yunnanensis	CBS 142895; CMW 47642; CERC 5337	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959557	MN959609	MN959694	No	MF443086	MF442986	MF442871	MF442756
	CBS 142897[5];CMW 47644; CERC 5339	Soil in *Eucalyptus* plantation	YunNan, China	HO	homothallic	MN959558	MN959610	MN959695	No	MF443088	MF442988	MF442873	MF442758
Ca. zuluensis	**CBS 125268**[5]; CMW 9188	*E. grandis*	Kwa-Zulu Natal, South Africa	HE	MAT1-2	No	No	MN959696	MN959750	FJ972414	GQ267459	FJ972433	FJ972483
	CBS 125272; CMW 9896	*E. grandis* × *E. urophylla* cutting	Pietermarizburg, South Africa	HE	MAT1-1	MN959559	MN959611	No	No	FJ972415	GQ267460	FJ972434	FJ972484

[1] CBS: Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands; CERC: China Eucalypt Research Centre, Chinese Academy of Forestry, Zhanjiang, GuangDong Province, China; CMW: culture collection of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa; CPC: Pedro Crous working collection housed at CBS; CGMCC: Microbiological Culture Collection Center, Beijing, China; YA: Quanzhu Chen working culture collection number (Ye et al. 2017).

[2] HE = Heterothallic; HO = Homothallic; P_HE = Putative heterothallic.

[3] *tub2* = β-tubulin; *cmdA* = calmodulin; *his3* = histone H3; *tef1* = translation elongation factor 1-alpha.

[4] Isolates representing ex-type cultures are indicated in bold.

[5] Isolate sequences were used in phylogenetic analyses.

[6] 'No' represents the relative *MAT* locus was not amplified successfully by the primers designed in the current study.

[7] '–' represents sequences that are not available.

[8] Genome sequences of the isolate were from public genomic databases and for which no cultures were available in this study.

[9] The genome sequences were generated in this study.

Genome *Ca. henricotiae**-1 = PGWR00000000[8]; *Ca. hongkongensis**-2 = JAACJA000000000[9]; *Ca. leucothoes**-3 = NAJI00000000[8]; *Ca. naviculata**-4 = NAGG00000000[8]; *Ca. pauciramosa**-5 = JAACIZ000000000[9]; *Ca. pauciramosa**-6 = JAACIY000000000[9]; *Ca. pseudonaviculata**-7 = JYJY00000000[8]; *Ca. pseudoreteaudii**-8 = MOCD00000000[8].

ex-type isolate of this species (CBS 138102) and all seven genome sequences of *Ca. pseudonaviculata* contained the same *MAT1-2* idiomorph as CBS 139394. The genome sequences of CBS 114417, which is the ex-type culture for *Ca. pseudonaviculata*, harboured only partial *MAT* gene sequences while CBS 139394 contained the full *MAT* gene sequences. Consequently, isolates CBS 138102 (*Ca. henricotiae*) and CBS 139394 (*Ca. pseudonaviculata*) were chosen to describe their *MAT* loci.

Determination of the MAT locus structures

The *MAT* genes in each of the available eight *Calonectria* genome sequences were characterised using a tBLASTx search on the CLC Main Workbench v. 7.9.1 using the *MAT* genes (*MAT1-2-1*, *MAT1-1-3*, *MAT1-1-2* and *MAT1-1-1*) reported in *Fusarium anguioides* NRRL 25385 (heterothallic, NCBI accession number MH742713; Jacobs-Venter et al. 2018) and *F. graminearum* 3639 (homothallic, NCBI accession number AF318048; Yun et al. 2000). These *Fusarium* spp., for which data are available regarding the *MAT* genes, are close relatives of *Calonectria* in the *Nectriaceae*. The contigs that produced hits with an E-value ≤ 10^{-2} were used to predict *MAT* genes and flanking regions using the online AUGUSTUS tool (http://bioinf.uni-greifswald.de/augustus/; Stanke et al. 2004). The *MAT* genes and their flanking regions were identified by BLASTp (NCBI), and further confirmed by comparison of homologs published on NCBI. The functional domains of the *MAT* genes were determined using the Conserved Domain search on NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

Comparison of MAT loci

A comparison of the *MAT* loci mined from genome sequences of the eight *Calonectria* isolates was generated using BLASTn with a maximum E-value cut off of 0.0001, and visualized using Easyfig v. 2.2.2 (Sullivan et al. 2011). Easyfig is a Python application used to create linear comparative figures of multiple genomic loci with an easy-to-use graphical user interface. Pairwise similarity comparisons (BLASTn, tBLASTx) between multiple genomic regions were generated using the Easyfig interface (Sullivan et al. 2011).

Primer design for MAT genes

MAT1-1-1 and *MAT1-2-1* primers were designed to determine the mode of sexual reproduction in a collection of 65 *Calonectria* species residing in 10 *Calonectria* species complexes. In addition, the available genome sequences were used to design primers for *MAT1-1-3* or *MAT1-2-12*, which were present in the heterothallic *Calonectria* isolates but absent in the one homothallic species (*Ca. hongkongensis*, CMW 47271).

The sequences of the *MAT1-1-1* and *MAT1-1-3* genes extracted from the genomes of *Ca. henricotiae* (CBS 138102), *Ca. hongkongensis* (CMW 47271, only for *MAT1-1-1* due to absence of *MAT1-1-3*), *Ca. naviculata* (CBS 101121) and *Ca. pauciramosa* (CMW 7592) were aligned. This alignment was used to design primers using the primer design function in CLC Main Workbench v. 7.9.1. following the software instructions. The alpha box domain in the *MAT1-1-1* gene and the HMG box domain in the *MAT1-1-3* gene were specifically targeted for primer design because these regions had the greatest similarity across all species.

The *MAT1-2-1* primers designed previously by Schoch et al. (2000) were based on the partial HMG box domain and produced fragments of approximately 170 bp. The whole *MAT1-2-1* gene region was used to design *MAT1-2-1* primers again in this study and aimed to obtain a longer *MAT1-2-1* fragment. The target areas for primer design for the *MAT1-2-1* and *MAT1-2-12* genes were based on the aligned sequences of the *MAT1-2-1* or *MAT1-2-12* gene found in the genomes of *Ca. hongkongensis* (CMW 47271, only for *MAT1-2-1* due to absence of *MAT1-2-12*), *Ca. leucothoes* (CBS 109166), *Ca. pauciramosa* (CMW 5683), *Ca. pseudonaviculata* (CBS 139394) and *Ca. pseudoreteaudii* (YA51) using CLC Main Workbench v. 7.9.1. The *MAT1-2-1* primers were designed in HMG box domain and overlapped with those designed by Schoch et al. (2000); *MAT1-2-12* primers were designed in the conserved areas.

MAT gene amplification and mating type assignment

All 123 isolates representing 65 *Calonectria* species were screened for four *MAT* genes (*MAT1-1-1*, *MAT1-1-3*, *MAT1-2-1* and *MAT1-2-12*). PCR amplification reaction conditions for these *MAT* genes were as follows: initial denaturation at 95 °C for 3 min, followed by 30 cycles of 95 °C denaturation for 30 s, 53 °C (*MAT1-1-1*) or 58 °C (*MAT1-2-1*) or 48 °C (*MAT1-1-3* or *MAT1-2-12)* annealing for 30 s, and 72 °C extension for 1 min, followed by a final extension at 72 °C for 10 min. PCR amplification mixtures, verification of PCR products, amplicon sequencing and sequence editing, assembly tools for *MAT* gene amplification and analyses were the same as those used to obtain the *tef1* gene regions described above. The sequences were aligned using the online version of MAFFT v. 7 (http://mafft.cbrc.jp/alignment/server/; Katoh & Standley 2013). Alignments of four *MAT* gene sequences were deposited in TreeBASE (http://treebase.org).

The conserved domains for each *MAT* gene sequence in all 123 *Calonectria* isolates were determined by the Pfam domain search on CLC Main Workbench v. 7.9.1. All of these sequences were deposited in GenBank (Table 1). Species having both *MAT1-1-1* and *MAT1-2-1* genes in a single isolate were designated as homothallic. Heterothallic species were identified by the presence of either *MAT1-1-1* or *MAT1-2-1* in different isolates. Species were considered to be putatively heterothallic when only the *MAT1-1-1* or *MAT1-2-1* gene was detected in all the isolates of a particular species (Duong et al. 2016).

Phylogenetic analysis and ancestral state reconstruction

To investigate the evolutionary history of sexual reproduction in *Calonectria*, a multi-gene phylogenetic tree based on Maximum Likelihood (ML) analysis for the combined dataset of the *tef1*, histone H3 (*his3*), calmodulin (*cmdA*) and partial β-tubulin (*tub2*) gene regions was generated using PhyML v. 3.1 (Guindon & Gascuel 2003). A single isolate representing each of 70 *Calonectria* species (Table 1) was selected for the phylogenetic analyses. These included the five species for which the genome sequences are publicly available and for which cultures were not used in this study (Table 1). All sequences used to construct the phylogenetic tree were either downloaded directly from NCBI (http://www.ncbi.nlm.nih.gov) or extracted from the genome sequences. Confidence levels for the nodes were determined with 1 000 bootstrap replicates. *Curvicladiella cignea* (CBS 109167) was used as the outgroup taxon in the analyses (Lombard et al. 2016). Alignment of sequence combination of four gene regions was deposited in TreeBASE (http://treebase.org).

The homothallic or heterothallic mode of reproduction in each of the 70 *Calonectria* species was mapped onto the backbone of the multi-gene phylogenetic tree. Ancestral state reconstruction based on the ML approach was performed using an unordered parsimony model in Mesquite v. 3.5 (Maddison & Maddison 2018).

RESULTS

Isolates and identification

The DNA for all 123 isolates representing 65 *Calonectria* spp. was successfully extracted. Confirmation of these previously

identified and published isolates was achieved based on a comparison of *tef1* sequences generated in this study and published on NCBI (Table 1).

Genome sequencing

For CMW 47271 (*Ca. hongkongensis*), CMW 5683 (*Ca. pauciramosa*) and CMW 7592 (*Ca. pauciramosa*), the estimated genome sizes were 61.7 Mb, 62.4 Mb and 62.3 Mb, respectively. The average coverage of all three assembled genomes were higher than 736×. The assembled genome of CMW 47271 (*Ca. hongkongensis*) had 76 scaffolds larger than 500 bp, a N50 contig size of 1.7 Mb and a mean GC content of 49.0 %. The genomes for CMW 5683 and CMW 7592 (*Ca. pauciramosa*) contained 83 scaffolds (> 500 bp) with N50 of 3.1 Mb, and 104 scaffolds (> 500 bp) with N50 of 1.4 Mb, respectively. These two genomes had a similar GC content of 49.3 %. The BUSCO analysis indicated a high level of completeness for all three assemblies based on the *Sordariomycetes* dataset and less than 1.2 % BUSCO orthologs were missing. GenBank accession numbers of these three genome sequences were JAACJA000000000, JAACIZ000000000 and JAACIY000000000, respectively (Table 1).

MAT locus structure and MAT genes in the eight Calonectria genomes

The *MAT* idiomorphs in each of the eight selected *Calonectria* isolates for which genome sequences were available were detected in a single contig (scaffold) based on a tBLASTx search on the CLC Main Workbench. Contigs from *Ca. leucothoes* (CBS 109166), *Ca. pauciramosa* (CMW 5683), *Ca. pseudonaviculata* (CBS 139394) and *Ca. pseudoreteaudii* (YA51) contained sequences very similar to those of the *MAT1-2-1* gene sequences in *F. graminearum* 3639 (E-value: 2.31E-8 to 4.14E-5). None of the contigs had similarity to the gene sequences of the *MAT1-1* idiomorph. These isolates were considered to contain only a *MAT1-2* idiomorph. *Calonectria henricotiae* (CBS 138102), *Ca. naviculata* (CBS 101121) and *Ca. pauciramosa* (CMW 7592) were designated as containing the *MAT1-1* idiomorph based on the presence of a *MAT1-1-1* gene and the absence of a *MAT1-2-1* gene in the *MAT* locus of each isolate. In addition, *Ca. hongkongensis* (CMW 47271) was found to have both *MAT1-1-1* and *MAT1-2-1* in a single scaffold and was confirmed as homothallic.

The length of the *MAT* idiomorph of *Ca. hongkongensis* (CMW 47271) was 4.66 kb. The *MAT1-1* idiomorph of *Ca. henricotiae* (CBS 138102), *Ca. naviculata* (CBS 101121) and *Ca. pauciramosa* (CMW 7592) were approximately 4.3 kb long, and the length of the *MAT1-2* idiomorph in *Ca. leucothoes* (CBS 109166), *Ca. pauciramosa* (CMW 5683), *Ca. pseudonaviculata* (CBS 139394) and *Ca. pseudoreteaudii* (YA51) was approximately 3.3 kb. The structural arrangement of the *MAT* locus and flanking genes was conserved in all isolates (Fig. 1). The *MAT* locus was flanked by the genes *APN2* (DNA lyase) and *SLA2* (cytoskeleton assembly control protein) gene.

The *MAT1-1* and *MAT1-2* idiomorphs in the genomes of the six heterothallic *Calonectria* species were identical in order and orientation (Fig. 1). The *MAT1-1* idiomorph in *Ca. henricotiae* (CBS 138102), *Ca. naviculata* (CBS 101121) and *Ca. pauciramosa* (CMW 7592) possessed the *MAT1-1-1*, *MAT1-1-2* and *MAT1-1-3* genes. A *MAT1-2-1* gene as well as an open reading frame (ORF) of unknown function were observed in the *MAT1-2* idiomorph of *Ca. leucothoes* (CBS 109166), *Ca. pauciramosa* (CMW 5683), *Ca. pseudonaviculata* (CBS 139394) and *Ca. pseudoreteaudii* (YA51). The *MAT1-1-3* gene and the ORF of un-

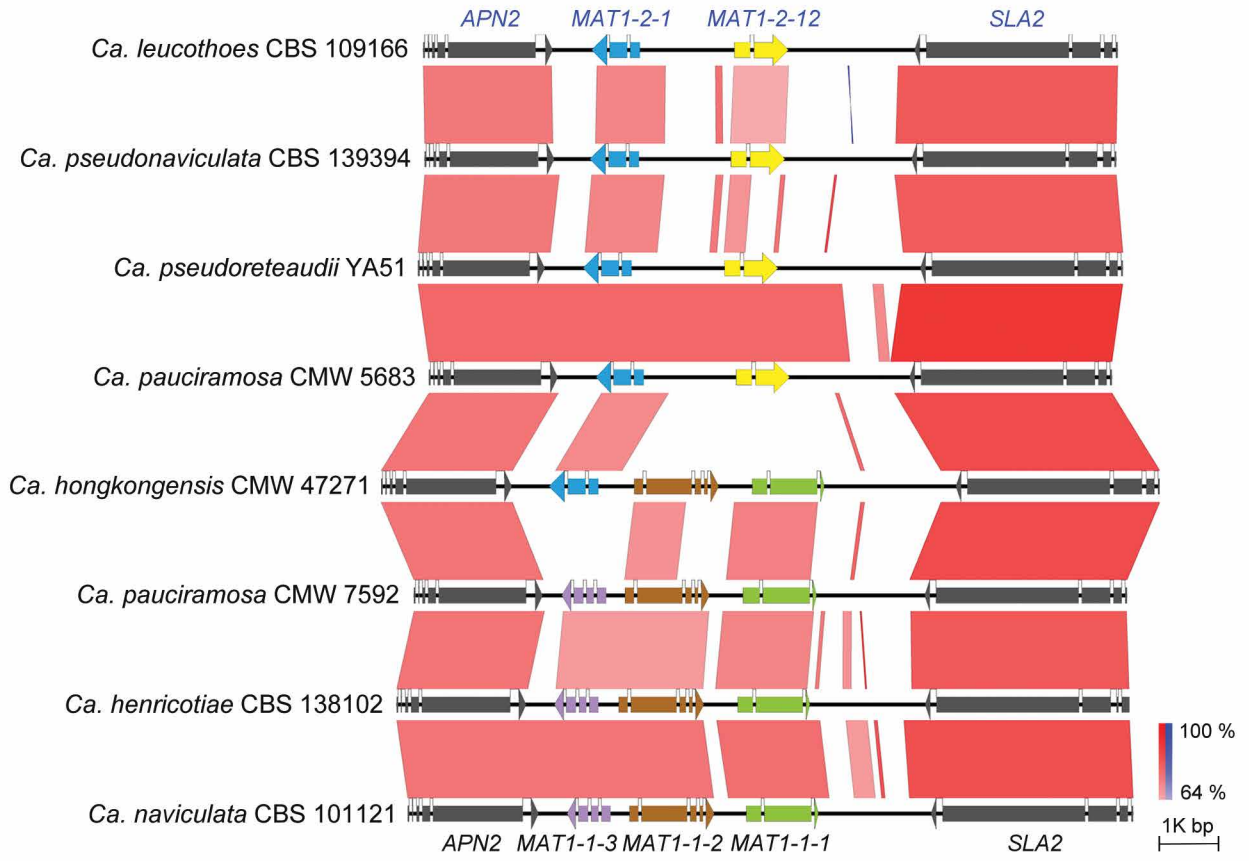

Fig. 1 Pairwise *MAT* loci comparison among eight *Calonectria* isolates representing seven species. Black horizontal lines represent genomic sequences. Colour coded arrows represent annotated genes. Red or blue boxes between genomic sequences indicates pairwise similarity based on BLASTn; red suggest that both regions are in the same orientation and blue are in opposite directions. *Calonectria hongkongensis* CMW 47271 represents the only homothallic individual containing both *MAT1-1* and *MAT1-2* idiomorph.

Table 2 Nucleotide and amino acid conservation of mating type and flanking genes in the genomes of eight *Calonectria* isolates.

Isolates	Nucleotide conservation (%)						
	SLA2	*MAT1-1-1*	*MAT1-1-2*	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*APN2*
Ca. henricotiae CBS 138102	66.37 (2 463/3 711)[1]	60.82 (742/1 220)	45.63 (657/1 440)	66.93 (500/747)			54.20 (1 188/2 192)
Ca. naviculata CBS 101121	71.95 (2 463/3 423)	60.77 (742/1 221)	45.72 (657/1 437)	67.84 (500/737)			53.71 (1 188/2 212)
Ca. pauciramosa CMW 7592	71.89 (2 463/3 426)	59.50 (742/1 247)	45.94 (657/1 430)	66.58 (500/751)			54.57 (1 188/2 177)
Ca. hongkongensis CMW 47271	71.31 (2 463/3 454)	60.92 (742/1 218)	45.98 (657/1 429)		56.99 (477/837)		53.71 (1 188/2 212)
Ca. leucothoes CBS 109166	71.62 (2 463/3 439)				58.24 (477/819)	49.34 (452/916)	54.22 (1 188/2 191)
Ca. pauciramosa CMW 5683	71.87 (2 463/3 427)				58.96 (477/809)	49.83 (452/907)	54.57 (1 188/2 177)
Ca. pseudonaviculata CBS 139394	71.08 (2 463/3 465)				57.26 (477/833)	49.24 (452/918)	54.20 (1 188/2 192)
Ca. pseudoreteaudii YA51	71.81 (2 463/3 430)				58.10 (477/821)	49.83 (452/907)	55.38 (1 188/2 145)
Isolates	Amino acid conservation (%)						
	SLA2	*MAT1-1-1*	*MAT1-1-2*	*MAT1-1-3*	*MAT1-2-1*	*MAT1-2-12*	*APN2*
Ca. henricotiae CBS 138102	83.48 (945/1 132)[2]	68.10 (254/373)	45.61 (187/410)	75.00 (150/200)			67.75 (416/614)
Ca. naviculata CBS 101121	89.83 (945/1 052)	68.10 (254/373)	45.61 (187/410)	76.53 (150/196)			66.99 (416/621)
Ca. pauciramosa CMW 7592	89.83 (945/1 052)	66.32 (254/383)	45.95 (187/407)	75.00 (150/200)			68.53 (416/607)
Ca. hongkongensis CMW 47271	89.83 (945/1 052)	68.28 (254/372)	45.95 (187/407)		62.30 (152/244)		66.99 (416/621)
Ca. leucothoes CBS 109166	89.83 (945/1 052)				62.81 (152/242)	39.65 (113/285)	68.42 (416/608)
Ca. pauciramosa CMW 5683	89.83 (945/1 052)				63.87 (152/238)	40.07 (113/282)	68.53 (416/607)
Ca. pseudonaviculata CBS 139394	89.83 (945/1 052)				62.04 (152/245)	39.51 (113/286)	67.75 (416/614)
Ca. pseudoreteaudii YA51	89.83 (945/1 052)				62.81 (152/242)	40.07 (113/282)	68.99 (416/603)

[1] The percentage of conserved nucleotides including exon and intron (length of conserved nucleotides/full-length of nucleotides).

[2] The percentage of conserved amino acid (length of conserved amino acid/full-length of amino acid).

known function, found respectively in the *MAT1-1* and *MAT1-2* locus of the heterothallic species, were absent in the *MAT* locus of homothallic *Ca. hongkongensis* (CMW 47271), which contained the *MAT1-1-1*, *MAT1-1-2* and *MAT1-2-1* genes. The ORF found in the *MAT1-2* locus of heterothallic *Calonectria* species was different to all other genes previously observed at a *MAT* locus. This was consequently recognised as a new mating type gene and is designated here as *MAT1-2-12*. This gene was previously designated as *MAT1-2-2* by Malapi-Wight et al. (2019).

The predicted *MAT1-1-1* (1.2 kb) gene in the eight *Calonectria* genomes contain two introns, and encode a 372 to 383 amino acid (aa) protein with a conserved MATalpha_HMGbox domain (GenBank: pfam04769) that spans a 49 bp intron. Both the *MAT1-1-3* (737 bp to 751 bp) and *MAT1-2-1* gene (809 bp to 837 bp) encode an HMG box domain (GenBank: cd01389), which is interrupted by an intron (about 50 bp). The predicted *MAT1-1-3* gene has a CDS approximately 600 bp in size and contains three introns. The putative *MAT1-2-1* gene has a CDS of approximately 720 bp and contains two introns. A conserved putative protein 1-1-2 domain (GenBank: pfam17043) was found in all *MAT1-1-2* (1.4 kb) genes. Although four introns were present in the *MAT1-1-2* gene, the conserved putative protein 1-1-2 domain was not interrupted by any of them. The novel mating type gene defined in this study as *MAT1-2-12* was approximately 910 bp long, has a predicted 60 bp intron and encodes for a putative protein around 285 aa with unknown domains.

A comparison of nucleotide and amino acid sequences of mating type genes among the eight isolates for which whole genome sequences were available, showed that non-coding intronic regions were more variable than the coding regions. This was with the exception of *MAT1-1-2* and *MAT1-2-12* (Table 2). The full nucleotide sequence (around 49 %) of the *MAT1-2-12* gene was more conserved than amino acid sequences (about 40 %), and both sequences had very similar variation in *MAT1-1-2* genes. The sequences of *APN2* were more variable than *MAT1-1-1* and *MAT1-1-3* in the eight *Calonectria* isolates (Table 2) used in this study and for which whole genome sequences were available.

MAT loci amplification and mating type assignment

Mating type markers designed in this study (Table 3) were used in PCRs to amplify portions of the *MAT1-1-1* (primers Cal_MAT111_F and Cal_MAT111_R), *MAT1-1-3* (primers Cal_MAT113_F and Cal_MAT113_R), *MAT1-2-1* (primers Cal_MAT121_F and Cal_MAT121_R) and *MAT1-2-12* (primers Cal_MAT1212_F and Cal_MAT1212_R) genes in the 123 *Calonectria* isolates representing 10 *Calonectria* species complexes. These resulted in PCR products of approximately 330 bp, 430 bp, 240 bp and 670 bp, respectively. The *MAT1-1-1* DNA sequences produced by PCR amplification all encoded a putative 110 amino acid sequence that included an alpha box domain. The *MAT1-1-3* encoded a sequence of 104 amino acids and *MAT1-2-1* encoded a sequence of 61 amino acids; the former having two predicted introns of about 50 bp and the latter an intron of 55 bp. Both sequences had an HMG domain that was interrupted by a single intron (Table 3). The alignments of each of the datasets of four *MAT* genes were deposited in TreeBASE (TreeBASE no 25663; http://treebase.org). An alignment analysis of the *MAT1-1-1*, *MAT1-1-3*, *MAT1-2-1* and *MAT1-2-12* sequences revealed little or no sequence variation in the genes within species but a high level of variation in the genes between species.

Based on the *MAT* gene amplification profile, 21 species (36 isolates) were identified as homothallic and 22 isolates representing eight species were heterothallic (Table 1). The remain-

Table 3 Primers for amplification of mating type gene fragments.

Target gene	Primer name	Primer sequence (5' to 3')	Tm (°C)	Fragment size (bp)	Target area
MAT1-1-1	Cal_MAT111_F Cal_MAT111_R	ATGCTTCCTCAGTCTTTGCT CTTGAAYRGGGTTGGTGG	53	330	Cal_MAT111_F → *MAT1-1-1* ← Cal_MAT111_R
MAT1-1-3	Cal_MAT113_F Cal_MAT113_R	CCTCCAGAAGTACCGACT GCTGTCGTTCTTCTTCCT	48	430	← Cal_MAT113_F *MAT1-1-3* Cal_MAT113_R →
MAT1-2-1	Cal_MAT121_F Cal_MAT121_R	GCAAGGAYCGCCACCRAAT GACACCTCKGCGTTTCTTCTCAG	58	240	← Cal_MAT121_F *MAT1-2-1* Cal_MAT121_R →
MAT1-2-12	Cal_MAT1212_F Cal_MAT1212_R	TCATCAGTTTCGCCCATT CGTCGTACTTCTTCTTCCG	48	670	Cal_MAT1212_F → *MAT1-2-12* ← Cal_MAT1212_R

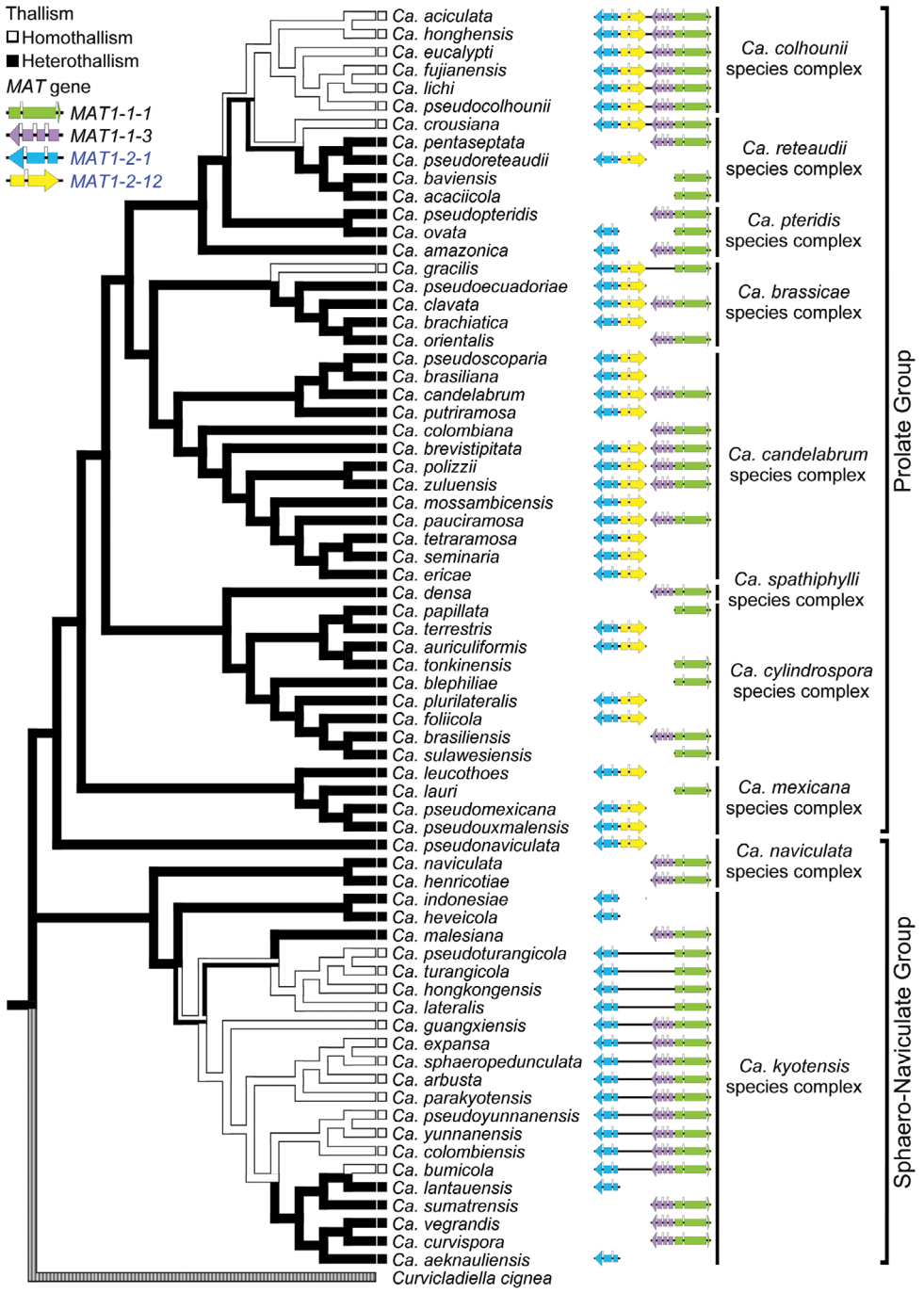

Fig. 2 Ancestral state reconstruction of sexual thallism of 70 *Calonectria* species. Homothallic species are marked with an open line, heterothallic species are marked with a solid line. Green, purple, blue and yellow coded arrows represent the *MAT1-1-1*, *MAT1-1-3*, *MAT1-2-1* and *MAT1-2-12* gene, respectively.

ing 36 species (65 isolates) were tentatively designated as heterothallic because only a *MAT1-1-1* or a *MAT1-2-1* gene was detected in isolates of these species. For the 21 homothallic species, 17 were first described from China, two (*Ca. eucalypti* CBS 125275 and *Ca. bumicola* CBS 143575) from Indonesia, *Ca. colombiensis* CBS 112221 from Colombia and *Ca. gracilis* CBS 111807 was from Brazil (Table 1).

The PCR amplification results revealed four different homothallic *MAT* loci in *Calonectria* (Fig. 2). In the Prolate Group, the *MAT* locus of most homothallic species contained the *MAT1-1-1*, *MAT1-1-3*, *MAT1-2-1* and *MAT1-2-12* genes. This was with the exception of *Ca. gracilis* in which the *MAT1-1-3* gene was not detected. In the Sphaero-Naviculate Group, the *MAT1-2-12* gene was absent in all homothallic species. In the clade represented by *Ca. lateralis*, the *MAT1-1-3* gene was absent in all of these species.

Ancestral state reconstruction of sexual thallism

The alignment of sequence combination of *tef1*, *his3*, *cmdA* and *tub2* genes was deposited in TreeBASE (TreeBASE no 25663; http://treebase.org). The ancestral state reconstruction analysis suggested that heterothallism is the ancestral state in *Calonectria*. This emerged from tracing the history of mating type characters onto the multi-gene phylogenetic species tree (Fig. 2). Three independent transitions from heterothallism to homothallism appear to have occurred across the phylogeny. One transition from homothallism to heterothallism was observed in the *Ca. kyotensis* species complex. Either a homothallic or a heterothallic lifestyle has occurred across *Calonectria* species in both the Prolate and Sphaero-Naviculate Groups. In most of the cases, the species with the same thallism grouped together in the phylogeny. Heterothallism was the most common state across the genus but homothallism was dominant for species in the Sphaero-Naviculate Group.

DISCUSSION

Analyses of genome sequences enabled the characterisation of the *MAT* loci in eight isolates representing seven species of *Calonectria*. In addition, the mating strategies of 65 *Calonectria* species were revealed using primers developed for four *MAT* genes. The *MAT* locus and flanking region was shown to have a conserved *APN2-MAT1-SLA2* structure, with differences observed in the genes of the *MAT* locus. From these results, and using ancestral state reconstruction, heterothallism was found to represent the ancestral reproductive state in *Calonectria*.

MAT loci and mating type genes

Species residing in the *Hypocreales* have commonly been found to harbour the *MAT1-1-1*, *MAT1-1-2* and *MAT1-1-3* genes in the *MAT1-1* idiomorph (Bushley et al. 2013). This is consistent with the results of the present study for heterothallic *Calonectria* species. In the *MAT1-2* idiomorph, in addition to the *MAT1-2-1* gene that was always present, the *MAT1-2-12* gene was described in this study. The discovery of this *MAT* gene in *Calonectria* represents a third gene to be discovered in this idiomorph in the *Hypocreales*. The other two genes include the *MAT1-2-8* in *Ustilaginoidea* (Yu et al. 2015, Wilken et al. 2017) and *MAT1-2-9* in *Fusarium* (Martin et al. 2011, Wilken et al. 2017). These three genes have not been detected in any fungi outside the *Hypocreales*, suggesting that they are probably restricted to this order. Gene deletions showed the *MAT1-2-9* (previously named *MAT1-2-3*, Wilken et al. 2017) have a similar expression pattern to the *MAT1-1-1* and *MAT1-2-1* in *F. graminearum* and *F. asiaticum* (Kim et al. 2012). The function of *MAT1-2-8* and *MAT1-2-12* in sexual reproduction has yet to be determined (Wilken et al. 2017, Malapi-Wight et al. 2019).

Neither the *MAT1-1-3* nor *MAT1-2-12* genes were observed in the *MAT* locus of the homothallic *Ca. hongkongensis*, *Ca. lateralis*, *Ca. pseudoturangicola* and *Ca. turangicola*. The *MAT1-1-3* gene has been reported as absent in the *MAT1-1* idiomorph of other *Hypocreales* fungi (Yokoyama et al. 2006, Bushley et al. 2013). Interestingly the *MAT1-1-3* gene was present in the various closely related species including *Ca. arbusta*, *Ca. bumicola*, *Ca. colombiensis*, *Ca. expansa*, *Ca. guangxiensis*, *Ca. parakyotensis*, *Ca. pseudoyunnanensis*, *Ca. sphaeropedunculata* and *Ca. yunnanensis*. This could reflect two different branches of evolution for the *MAT* locus in *Calonectria* spp. Mutation analyses of *MAT1-1-2* and *MAT1-1-3* have shown that these two genes have similar expression profiles and may possess overlapping functions in sexual development (Ferreira et al. 1998, Zheng et al. 2013). In addition, species maintaining the *MAT1-1-3* gene in the *Hypocreales* are also located at a more ancestral position in the mating type tree than species lacking the *MAT1-1-3* gene (Yokoyama et al. 2006). We consequently hypothesize that the *MAT* locus lacking the *MAT1-1-3* gene in *Calonectria* may have evolved from an ancestral locus containing all three genes (*MAT1-1-1*, *MAT1-1-2* and *MAT1-1-3*).

Distribution of mating types

Previous studies have shown that most species in *Calonectria* are heterothallic with a biallelic mating system (Crous et al. 1998, Crous 2002, Lombard et al. 2010a–c). This was supported in the results of the present study, where 44 of 65 *Calonectria* species were found to be heterothallic. These results also suggest that heterothallism is the ancestral state in *Calonectria*. The 21 homothallic species reside primarily in the *Ca. colhounii* and *Ca. kyotensis* species complexes. But in both these complexes, heterothallism is basal. This suggests that these species had a common homothallic ancestor, which has evolved from a heterothallic state.

The *MAT* genes observed in *Ca. bumicola*, *Ca. crousiana* and *Ca. gracilis* suggest that these species are homothallic while their closest neighbours in the same clade/group are all heterothallic. This is unusual and in contrast to views in a previous study (Duong et al. 2016) where species residing in the same complex consistently shared the same mode of sexual reproduction. The fact that only the *MAT1-1-1* or *MAT1-2-1* genes amplified in a number of isolates of *Calonectria*, provides a level of confidence in our results. It is, however, possible that the primers designed for the *MAT1-1-3* and *MAT1-2-12* failed to allow the detection of these genes and whole genome sequences would be needed to confirm this result.

Evolution of mating type

The results of this study indicated that heterothallism represents the ancestral reproductive state in *Calonectria*. Furthermore, that one independent transition from homothallism back to heterothallism has occurred in the *Ca. kyotensis* species complex. Evolution of homothallism from heterothallism has apparently occurred due to unequal crossing over and translocation of the *MAT* idiomorphs in various *Ascomycete* fungi, including *Bipolaris* = *Cochliobolus* (Yun et al. 1999), *Stemphylium* = *Pleospora* (Inderbitzin et al. 2005), *Crivellia* = *Alternaria* (Inderbitzin et al. 2006), *Neurospora* (Nygren et al. 2011, Gioti et al. 2012) and *Eutiarosporella* (Thynne et al. 2017). In contrast, fewer studies have shown heterothallic fungi have been derived from homothallic ancestors via gene loss. In this way, partial gene sequences of the genes residing in the *MAT1-2* idiomorph have been incorporated into the *MAT1-1* idiomorph or vice versa, such as *Aspergillus fumigatus* (Paoletti et al. 2005), *Botrytis cinerea* (Amselem et al. 2011) and *Cordyceps takaomontana* (Yokoyama et al. 2003). Although it is possible that the transition between homothallism and heterothallism in

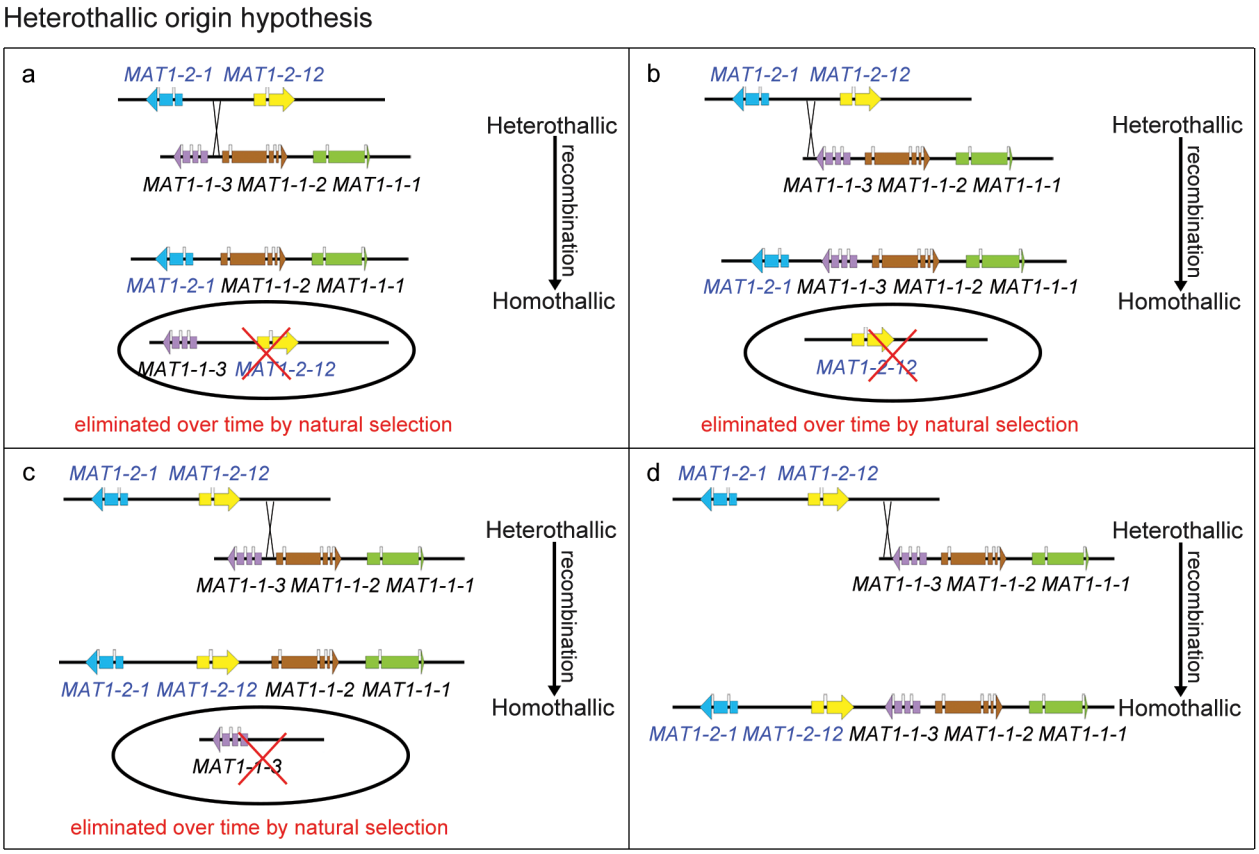

Fig. 3 Evolution models of mating type in *Calonectria* spp.: Heterothallic origin hypothesis. a–d. Four scenarios under which the mating type loci of heterothallic ancestors undergo an independent recombination event (unequal crossing over), resulting in the present homothallic mating type locus.

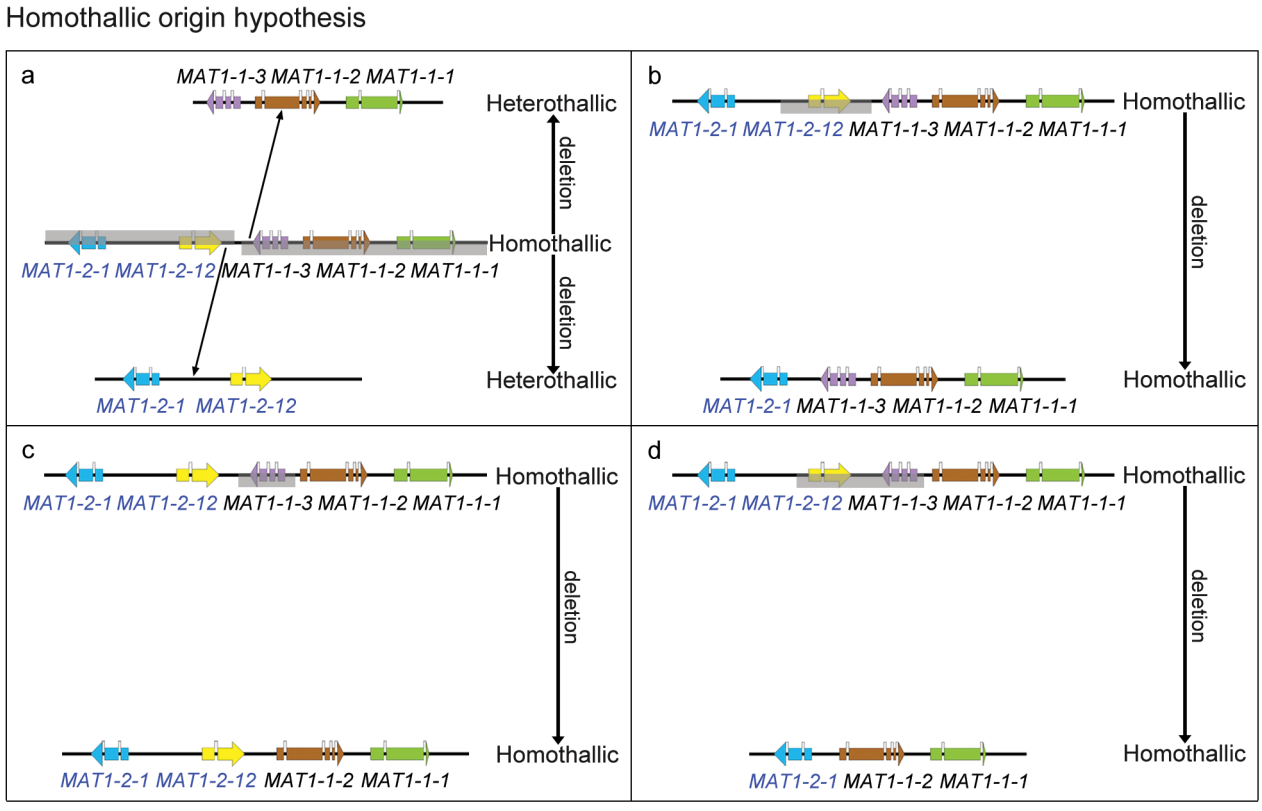

Fig. 4 Evolution models of mating type in *Calonectria* spp.: Homothallic origin hypothesis. a. Primary homothallic ancestor mating type locus undergoes two deletions events (gene loss) and this results in the mating type locus of two heterothallic offspring; b–d. primary homothallic ancestor mating type locus undergoes an independent deletion event which results in the present homothallic mating type locus.

Ascomycetes could occur in either direction, a switch from one state should logically reflect an evolutionary advantage. In this regard, heterothallism would offer the advantage of enhanced genetic diversity and adaption to the environment (Lumley et al. 2015). In contrast, homothallism offers the benefits of sexual recombination without needing isolates of the opposite mating type (Wilson et al. 2015b).

A proposed evolution model for mating type

The structure of mating type loci in *Calonectria* species revealed in this study makes it possible to explain the evolution of the mating types following two possible hypotheses (Fig. 3, 4). In one case, which we consider as the recombination hypothesis, there has been an ancestral shift from heterothallism to homothallism in four independent unequal recombination events (Fig. 3a–d). These would have resulted in the mating type idiomorphs observed in the present study.

An alternative hypothesis would involve a shift from a homothallic ancestor containing all the *MAT* genes (*MAT1-1-1*, *MAT1-1-2*, *MAT1-1-3*, *MAT1-2-12* and *MAT1-2-1*) to a heterothallic state via at least two deletion events (Fig. 4a–d). In this case, the homothallic ancestor would have also undergone three independent deletion events to arrive at the currently identified homothallic species. This hypothesis is less parsimonious than the recombination hypothesis. Based on parsimony (Rasmussen & Ghahramani 2001), a heterothallic origin hypothesis is more probable than the homothallic origin hypothesis. However, it is not possible to rule out the possibility that the original ancestor of the heterothallic species was in fact not homothallic and that species in this genus have evolved from homothallism to heterothallism and then some have switched back to homothallism.

Reproductive modes and pathogenicity

Results of this study have made it possible to easily characterise the mating type of important *Calonectria* spp. This will enhance the value of population genetic studies on these fungi where the presence or absence of sexual reproduction can be considered. The results will also support quarantine regulations that should seek to prevent the introduction of opposite mating type strains in heterothallic *Calonectria* spp., where only one of these is known to be present in a country. This can preclude the generation of new genotypes of such pathogens and a breakdown of resistance developed in the host (McDonald & Linde 2002, Lombard et al. 2010a, Malapi-Wight et al. 2014).

Acknowledgements This study was supported financially by the special fund for basic scientific research of State Key Laboratory of Tree Genetics and Breeding (SKLTGB) of China (project no. TGB2017001), the National Natural Science Foundation of China (NSFC) (project no. 31622019), the National Key R&D Program of China (project no. 2017YFD0600103), the National Ten-thousand Talents Program (project No. W03070115), the GuangDong Top Young Talents Program (project No. 20171172), the Genomics Research Institute and members of the Tree Protection and Cooperation Programme (TPCP), South Africa. We thank Dr. Seonju Marincowitz for assistance in sourcing cultures and Mr. Jan Nagel, Dr. Tuan Duong, Dr. Markus Wilken and Ms. Katrin Fitza for advice.

REFERENCES

Adamson K, Mullett MS, Solheim H, et al. 2018. Looking for relationships between the populations of Dothistroma septosporum in northern Europe and Asia. Fungal Genetics and Biology 110: 15–25.

Alby K, Schaefer D, Bennett RJ. 2009. Homothallic and heterothallic mating in the opportunistic pathogen Candida albicans. Nature 460: 890–893.

Amselem J, Cuomo CA, Van Kan JA, et al. 2011. Genomic analysis of the necrotrophic fungal pathogens Sclerotinia sclerotiorum and Botrytis cinerea. PLoS Genetics 7: e1002230.

Billiard S, López-Villavicencio M, Hood ME, et al. 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25: 1020–1038.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Bushley KE, Li Y, Wang WJ, et al. 2013. Isolation of the MAT1-1 mating type idiomorph and evidence for selfing in the Chinese medicinal fungus Ophiocordyceps sinensis. Fungal Biology 117: 599–610.

Crous PW. 2002. Taxonomy and pathology of Cylindrocladium (Calonectria) and allied genera. APS Press, USA.

Crous PW, Alfenas AC, Junghans TG. 1998. Variability within Calonectria ovata and its anamorph Cylindrocladium ovatum from Brazil. Sydowia 50: 1–3.

Crous PW, Carnegie AJ, Wingfield MJ, et al. 2019. Fungal Planet description sheets: 868–950. Persoonia 42: 291–473.

Crous PW, Groenewald JZ, Risède JM, et al. 2004. Calonectria species and their Cylindrocladium anamorphs: species with sphaeropedunculate vesicles. Studies in Mycology 50: 415–430.

Crow JF. 1994. Advantages of sexual reproduction. Developmental Genetics 15: 205–213.

Duong TA, De Beer ZW, Wingfield BD, et al. 2012. Phylogeny and taxonomy of species in the Grosmannia serpens complex. Mycologia 104: 715–732.

Duong TA, De Beer ZW, Wingfield BD, et al. 2016. Mating type markers reveal high levels of heterothallism in Leptographium sensu lato. Fungal Biology 120: 538–546.

Ferreira AV, An Z, Metzenberg RL, et al. 1998. Characterization of mat A-2, mat A-3 and ΔmatA mating-type mutants of Neurospora crassa. Genetics 148: 1069–1079.

Fraser JA, Stajich JE, Tarcha EJ, et al. 2007. Evolution of the mating type locus: insights gained from the dimorphic primary fungal pathogens Histoplasma capsulatum, Coccidioides immitis, and Coccidioides posadasii. Eukaryotic Cell 6: 622–629.

Gioti A, Mushegian AA, Strandberg R, et al. 2012. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. Molecular Biology and Evolution 29: 3215–3226.

Goodenough U, Heitman J. 2014. Origins of eukaryotic sexual reproduction. Cold Spring Harbor Perspectives in Biology 6: a016154

Goodwin SB, Drenth A, Fry WE. 1992. Cloning and genetic analyses of two highly polymorphic, moderately repetitive nuclear DNAs from Phytophthora infestans. Current Genetics 22: 107–115.

Gordo I, Campos PR. 2008. Sex and deleterious mutations. Genetics 179: 621–626.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52: 696–704.

Inderbitzin P, Harkness J, Turgeon BG, et al. 2005. Lateral transfer of mating system in Stemphylium. Proceedings of the National Academy of Sciences 102: 11390–11395.

Inderbitzin P, Shoemaker RA, O'Neill NR, et al. 2006. Systematics and mating systems of two fungal pathogens of opium poppy: the heterothallic Crivellia papaveracea with a Brachycladium penicillatum asexual state and a homothallic species with a Brachycladium papaveris asexual state. Botany 84: 1304–1326.

Jacobs-Venter A, Laraba I, Geiser DM, et al. 2018. Molecular systematics of two sister clades, the Fusarium concolor and F. babinda species complexes, and the discovery of a novel microcycle macroconidium-producing species from South Africa. Mycologia 110: 1189–1204.

James TY, Kauff F, Schoch CL, et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443: 818–822.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30: 772–780.

Kearse M, Moir R, Wilson A, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

Kim HK, Cho EJ, Lee S, et al. 2012. Functional analyses of individual mating-type transcripts at MAT loci in Fusarium graminearum and Fusarium asiaticum. FEMS Microbiology Letters 337: 89–96.

Labarere J, Noel T. 1992. Mating type switching in the tetrapolar basidiomycete Agrocybe aegerita. Genetics 131: 307–319.

LeBlanc N, Gehesquière B, Salgado-Salazar C, et al. 2019. Limited genetic diversity across pathogen populations responsible for the global emergence of boxwood blight identified using SSRs. Plant Pathology 68: 861–868.

Li J, Wingfield MJ, Liu Q, et al. 2017. Calonectria species isolated from Eucalyptus plantations and nurseries in South China. IMA Fungus 8: 259–294.

Lin X, Heitman J. 2007. Mechanisms of homothallism in fungi and transitions between heterothallism and homothallism. In: Heitman J, Kronstad JW, Taylor JW, et al. (eds), Sex in fungi: molecular determination and evolutionary implications: 35–57. ASM Press, Washington DC.

Lombard L, Crous PW, Wingfield BD, et al. 2010a. Multigene phylogeny and mating tests reveal three cryptic species related to Calonectria pauciramosa. Studies in Mycology 66: 15–30.

Lombard L, Crous PW, Wingfield BD, et al. 2010b. Phylogeny and systematics of the genus Calonectria. Studies in Mycology 66: 31–69.
Lombard L, Crous PW, Wingfield BD, et al. 2010c. Species concepts in Calonectria (Cylindrocladium). Studies in Mycology 66: 1–14.
Lombard L, Wingfield MJ, Alfenas AC, et al. 2016. The forgotten Calonectria collection: pouring old wine into new bags. Studies in Mycology 85: 159–198.
Lopes A, Phillips AJ, Alves A. 2017. Mating type genes in the genus Neofusicoccum: mating strategies and usefulness in species delimitation. Fungal Biology 121: 394–404.
Lumley AJ, Michalczyk Ł, Kitson JJ, et al. 2015. Sexual selection protects against extinction. Nature 522: 470–473.
Maddison WP, Maddison DR. 2018. Mesquite: a modular system for evolutionary analysis. Version 3.5. http://www.mesquiteproject.org [accessed 18 June 2018].
Malapi-Wight M, Demers JE, Veltri D, et al. 2016a. LAMP detection assays for boxwood blight pathogens: a comparative genomics approach. Scientific Reports 6: 26140.
Malapi-Wight M, Hébert JB, Rivera Y, et al. 2014. Comparative genomics in the boxwood blight system: Insights into the global diversity of the mating-type locus. (Abstr.) Phytopathology 104: S74.
Malapi-Wight M, Salgado-Salazar C, Demers JE, et al. 2016b. Sarcococca blight: use of whole-genome sequencing for fungal plant disease diagnosis. Plant Disease 100: 1093–1100.
Malapi-Wight M, Veltri D, Gehesquière B, et al. 2019. Global distribution of mating types shows limited opportunities for mating across populations of fungi causing boxwood blight disease. Fungal Genetics and Biology 131: 103246.
Marin-Felix Y, Groenewald JZ, Cai L, et al. 2017. Genera of phytopathogenic fungi: GOPHY 1. Studies in Mycology 86: 99–216.
Martin SH, Wingfield BD, Wingfield MJ, et al. 2011. Structure and evolution of the Fusarium mating type locus: new insights from the Gibberella fujikuroi complex. Fungal Genetics and Biology 48: 731–740.
McDonald BA, Linde C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annual Review of Phytopathology 40: 349–379.
Nagel JH, Wingfield MJ, Slippers B. 2018. Evolution of the mating types and mating strategies in prominent genera in the Botryosphaeriaceae. Fungal Genetics and Biology 114: 24–33.
Ni M, Feretzaki M, Sun S, et al. 2011. Sex in fungi. Annual Review of Genetics 45: 405–430.
Nygren K, Strandberg R, Wallberg A, et al. 2011. A comprehensive phylogeny of Neurospora reveals a link between reproductive mode and molecular evolution in fungi. Molecular Phylogenetics and Evolution 59: 649–663.
O'Donnell K, Ward TJ, Geiser DM, et al. 2004. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the Fusarium graminearum clade. Fungal Genetics and Biology 41: 600–623.
Paoletti M, Rydholm C, Schwier EU, et al. 2005. Evidence for sexuality in the opportunistic fungal pathogen Aspergillus fumigatus. Current Biology 15: 1242–1248.
Pham NQ, Barnes I, Chen SF, et al. 2019. Ten new species of Calonectria from Indonesia and Vietnam. Mycologia 111: 78–102.
Rasmussen CE, Ghahramani Z. 2001. Occam's razor. In: Leen TK, Dietterich, TG, Tresp V (eds), Advances in Neural Information Processing Systems 13: 294–300. MIT Press, Cambridge MA.
Schoch CL, Crous PW, Wingfield BD, et al. 1999. The Cylindrocladium candelabrum species complex includes four distinct mating populations. Mycologia 91: 286–298.
Schoch CL, Crous PW, Witthuhn RC, et al. 2000. Recombination in Calonectria morganii and phylogeny with other heterothallic small-spored Calonectria species. Mycologia 92: 665–673.
Simão FA, Waterhouse RM, Ioannidis P, et al. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210–3212.
Stanke M, Steinkamp R, Waack S, et al. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Research 32: W309–W312.
Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27: 1009–1010.
Thynne E, McDonald MC, Solomon PS. 2017. Transition from heterothallism to homothallism is hypothesised to have facilitated speciation among emerging Botryosphaeriaceae wheat-pathogens. Fungal Genetics and Biology 109: 36–45.
Turgeon BG, Yoder OC. 2000. Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genetics and Biology 31: 1–5.
Wilken PM, Steenkamp ET, Wingfield MJ, et al. 2017. Which MAT gene? Pezizomycotina (Ascomycota) mating-type gene nomenclature reconsidered. Fungal Biology Reviews 31: 199–211.
Wilson AM, Godlonton T, Van der Nest MA, et al. 2015a. Unisexual reproduction in Huntiella moniliformis. Fungal Genetics and Biology 8: 1–9.
Wilson AM, Wilken PM, Van der Nest MA, et al. 2015b. Homothallism: an umbrella term for describing diverse sexual behaviours. IMA Fungus 6: 207–214.
Wingfield BD, Duong TA, Hammerbacher A, et al. 2016. Draft genome sequences for Ceratocystis fagacearum, C. harringtonii, Grosmannia penicillata, and Huntiella bhutanensis. IMA Fungus 7: 317–323.
Ye X, Liu H, Jin Y, et al. 2017. Transcriptomic analysis of Calonectria pseudoreteaudii during various stages of Eucalyptus infection. PLoS One 12: e0169598.
Yokoyama E, Arakawa M, Yamagishi K, et al. 2006. Phylogenetic and structural analyses of the mating-type loci in Clavicipitaceae. FEMS Microbiology Letters 264: 182–191.
Yokoyama E, Yamagishi K, Hara A. 2003. Structures of the mating-type loci of Cordyceps takaomontana. Applied and Environmental Microbiology 69: 5019–5022.
Yu JJ, Sun WX, Yu MN, et al. 2015. Characterization of mating-type loci in rice false smut fungus Villosiclava virens. FEMS Microbiology Letters 362: fnv014.
Yun SH, Arie T, Kaneko I, et al. 2000. Molecular organization of mating type loci in heterothallic, homothallic, and asexual Gibberella/Fusarium species. Fungal Genetics and Biology 31: 7–20.
Yun SH, Berbee ML, Yoder OC, et al. 1999. Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. Proceedings of the National Academy of Sciences 96: 5592–5597.
Zheng Q, Hou R, Ma J, et al. 2013. The MAT locus genes play different roles in sexual reproduction and pathogenesis in Fusarium graminearum. PLoS One 8: e66980.