3 Biotech (2021) 11:402
https://doi.org/10.1007/s13205-021-02950-1

GENOME REPORTS

Whole genome sequences of two *Trichophyton indotineae* clinical isolates from India emerging as threats during therapeutic treatment of dermatophytosis

Pawan Kumar[1,2] · Shukla Das[3] · Richa Tigga[3,5] · Rajesh Pandey[1] · S. N. Bhattacharya[4] · Bhupesh Taneja[1,2]

Received: 27 February 2021 / Accepted: 28 July 2021 / Published online: 9 August 2021
© King Abdulaziz City for Science and Technology 2021

Abstract

In the current study, we report the genome sequence of two different clinical isolates from India, *Trichophyton indotineae* UCMS-IGIB-CI12 and *Trichophyton indotineae* UCMS-IGIB-CI14. The resulting genome assembly achieved a 143-fold coverage in 824 contigs for *T. indotineae* UCMS-IGIB-CI12 and a 136-fold coverage in 904 contigs for *T. indotineae* UCMS-IGIB-CI14. Both the clinical isolates contain a c.1342G>A mutation corresponding to Ala448Thr amino acid substitution in *erg1* and exhibit an intermittent drug response to terbinafine. Comparative genomics analysis with available genomes of *Trichophyton interdigitale/Trichophyton mentagrophytes* species complex revealed a similar genome architecture and identified large number of genes associated with virulence and pathogenicity, namely, lipases, proteases, LysM domain-containing factors, carbon metabolism enzymes and cytochrome P450 enzymes, in all the genomes. An analysis of single amino acid polymorphisms (SAPs) in the protein sequences of subtilisin and lipase enzyme families identified a higher frequency of SAPs in functionally important proteins, Sub3 and Sub6 and their possible use in multilocus phylogenetic analysis of *T. interdigitale/T. mentagrophytes* species complex. The whole genome sequences of *T. indotineae* clinical isolates provided in this report will, hence, serve as a key reference point for investigation of clinical strains and emerging drug resistance among dermatophytes originating from different parts of the world.

Keywords *Trichophyton* spp. · Dermatophytes, whole-genome sequencing · Phylogenetic analysis

Abbreviations

wgs	Whole genome sequence
cds	Coding DNA sequences
ITS	Internal transcribed spacer
TiCI12	*T. indotineae* UCMS-IGIB-CI12
TiCI14	*T. indotineae* UCMS-IGIB-CI14
MIC	Minimum inhibitory concentration
AFST	Antifungal susceptibility test
SNP	Single nucleotide polymorphism
SAP	Single amino-acid polymorphism

✉ Bhupesh Taneja
btaneja@igib.res.in

1 CSIR-Institute of Genomics and Integrative Biology, New Delhi 110025, India

2 Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India

3 Department of Microbiology, UCMS and GTB Hospital, Dilshad Garden, Delhi 110095, India

4 Department of Dermatology, UCMS and GTB Hospital, Dilshad Garden, Delhi 110095, India

5 Present Address: Department of Microbiology, All India Institute of Medical Sciences, Raipur, India

Dermatophytes are a group of keratinophilic filamentous fungi belonging to the genera *Trichophyton*, *Epidermophyton* and *Microsporum* that initiate infection with the attachment of arthroconidia to the stratum corneum and consume keratin, a main component of outer layer of skin, as its primary nutrient source (White et al. 2008; Baldo et al. 2012). Further spread of the developing hyphae through the keratin layer is aided by a repertoire of secretory keratinases, proteases and lipases (Martinez et al. 2012; Latka et al. 2015), that act as important virulence factors (Monod 2008; Peres et al. 2010). During recent years, there are many reports of drug resistance in several *Trichophyton* spp. among dermatophytes emerging from India and rest of the world. While majority of the recalcitrant infections indicate resistance to terbinafine, prevalence of resistance across azole family of drugs has also been reported, with

most reports for resistance to either drug from *Trichophyton mentagrophytes/Trichophyton interdigitale* isolates (Yamada et al. 2017; Rudramurthy et al. 2018; Singh et al. 2018; Saunte et al. 2019; Ebert et al. 2020). Failure of allylamine and azole drugs to make interference in ergosterol synthesis pathway, the primary target of antifungal agents, sounds the alarm for revisiting the traditional molecular workflow and a need for change in traditional thinking in treatment of dermatophytosis. Availability of few whole genome sequences (wgs), across genomic resources of EMBL/GenBank/DDBJ has limited a genome-wide detailed analysis approach so far. We provide here the whole genome sequence of two clinical isolates from India belonging to *T. mentagrophytes/T. interdigitale* genotype VIII (Taghipour et al. 2019; Nenoff et al. 2019) and recently reannotated as *Trichophyton indotineae* (Kano et al. 2020; Tang et al. 2021). The wgs of *T. indotineae* clinical isolates from India and their comparison with available genomes of *T. interdigitale/T. mentagrophytes* species complex will aid in our understanding of pathogenesis of these dermatophytes, in better management of the disease and further evaluation of available therapeutic options.

The clinical isolates were collected from skin scrapings around the lesions and examined under the microscope using 10% KOH. A portion of the sample was cultured on Sabouraud's dextrose agar with chloramphenicol (0.05 g/L), gentamicin (20 mg/L) and cycloheximide (0.5 g/L) at 25 °C for 3–4 weeks. After growth, the etiological agent was confirmed by the characteristic morphology of the colony and by studying the microscopic appearance of the fungus on Lacto Phenol Cotton Blue mount. Genotypic identification of the isolates was carried out by DNA extraction and PCR amplification of the region spanning nuclear ribosomal internal transcribed spacer (ITS) regions 1 to 2 of 18S rRNA using panfungal primers; ITS1 (5′-TCCGTAGGTGAACCTGCGG-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) (White et al. 1990) followed by Sanger sequencing and comparison with available ITS sequences of dermatophytes in GenBank using BLAST. A 100% sequence similarity of the amplified 18S rRNA region was obtained with the ITS sequence of *T. mentagrophytes* genotype VIII (GenBank accession number: MH517560.1, Nenoff et al. 2019). *T. mentagrophytes* genotype VIII has recently been reclassified as a new species, i.e. *T. indotineae*, with characteristic three single nucleotide polymorphisms (SNPs) at position 94 (C), 125 (T) and 462 (T) in the ITS region. The presence of all three characteristic SNPs instead of 94A, 125C and 462C compared to the reference *T. interdigitale* 428.63 strain (Kano et al. 2020) (Fig. 1) confirmed the two isolates as *T. indotineae* and were designated as *T. indotineae* UCMS-IGIB-CI12 (abbreviated here as TiCI12) and *T. indotineae* UCMS-IGIB-CI14 (abbreviated here as TiCI14), respectively. The sequence of the internal transcribed spacer regions 1 to 2 of 18S rRNA of the two isolates TiCI12 and TiCI14 are deposited in GenBank with Accession numbers MW600527.1 and MW600653.1.

Several dermatophytes from the *T. interdigitale/T. mentagrophytes* species complex from India, including the recently reannotated *T. indotineae*, exhibit high level of terbinafine resistance associated with mutations in its molecular target, squalene epoxidase (*SQLE/erg1*) (Khurana et al. 2018; Rudramurthy et al. 2018; Singh et al. 2018, 2019, 2021; Shaw et al. 2020; Ebert et al. 2020; Gaurav et al. 2021). The presence of any mutations in *erg1* of TiCI12 or TiCI14 was probed by sequencing of PCR amplified product of *erg1* of TiCI12 and TiCI14 using forward primer, FP (5′ ATGGTTGTAGAGGCTCCTCCCTGC 3′) and reverse primer, RP (5′ CTAGCTTTGAAGTTCGGCAAATA 3′). A c.1342G>A mutation corresponding to Ala448Thr amino acid substitution was identified in *erg1* of both the clinical isolates. Any change in minimum inhibitory concentration (MIC) due to these substitutions was assessed by an antifungal susceptibility test (AFST) to terbinafine as well as fluconazole by broth microdilution method as per Clinical and Lab Standards Institute (CLSI) M38-A2 guidelines, as defined earlier (Gaurav et al. 2021). *T. mentagrophytes* ATCC 18748 was included as a quality control strain for AFST. The in vitro antifungal susceptibility profile of TiCI12 and TiCI14 to these agents is summarized in Table 1. While MIC_{90} to fluconazole (16 µg/mL) was same in the control and clinical isolates, and in the range reported earlier for *T. interdigitale/T. mentagrophytes* (Rudramurthy et al. 2018; Singh et al. 2019), TiCI12 and TiCI14 showed variable tolerance to terbinafine, highlighting the need for detailed analysis of genomic features likely to be associated with virulence, pathogenicity and variability of these Indian strains.

For whole genome sequencing, the purified gDNA was end-repaired, adapter-ligated and enriched by following manufacturer's protocol to make the library for sequencing using Nextra XT DNA Library Prep Kit. Library quantification was carried out using Agilent Bioanalyzer with high sensitivity DNA kit. The enriched libraries were pooled according to the unique adapter barcodes and sequencing was carried out on Illumina platform by 2 × 100 bp paired-end sequencing. Quality assessment of raw reads obtained in paired-end reads was done using FastQC-v.0.11.8 (Release Date: 04-October-2018) (Andrews 2010) followed by trimming to remove any duplicates and low quality reads using Trimmomatic v0.39 tools (Bolger et al. 2014). Further de novo assembly of filtered reads was done using SPAdes assembler v-3.13 (Bankevich et al. 2012) followed by Pilon correction (Walker et al. 2014) in order to make improvement within the draft assemblies. Assembly statistics and completeness was estimated using Assemblathon 3 and BUSCO v4.0.3 (Seppey et al. 2019) with lineage dataset i.e. fungi_odb10 (creation date: 2019-12-13). Finally, assembled

Fig. 1 Comparison of ITS sequences of *Trichophyton* spp. Sequence alignment of internal transcribed spacer region (ITS1–ITS2) of 18S rRNA gene of *T. indotineae* UCMS-IGIB-CI12 (GenBank accession number: MW600527.1, this study) and *T. indotineae* UCMS-IGIB-CI14 (GenBank accession number: MW600653.1, this study), *T. interdigitale* NCCPF 800062 (*T. mentagrophytes* type VIII reference strain, Nenoff et al. (2019), GenBank accession number: MH517560.1) and *T. interdigitale* CBS 428.63 (*T. interdigitale* Type II reference strain, Nenoff et al. (2019), GenBank accession number: KT155896.1) showing the SNPs for *T. indotineae* (in red). The numbering convention for ITS sequence according to Kano et al. (2020) has been used for clarity and for maintaining reproducibility with their analysis

fasta files were carried forward for structural and functional annotation using web-interface of AUGUSTUS (Stanke and Morgenstern 2005) and PANNZER2 (Protein ANNotation with Z-scoRE) (Törönen et al. 2018), respectively. The general features of TiCI12 and TiCI14 are summarized in Table 2 following the Minimal Information about any (X) Sequence (MIxS) standard checklist.

The total estimated size of the draft assembled genomes of TiCI12 and TiCI14 is 22.06 Mb and 22.04 Mb in 824 and 904 contigs with > 130-fold coverage for each genome and an N_{50} value of 57.4 kb and 54.0 kb for TiCI12 and TiCI14, respectively (Table 2). Despite the recent classification of members of *T. interdigitale*/*T. mentagrophytes* species complex into different genotype groups representing their geographical distribution (Nenoff et al. 2019; Taghipour et al. 2019) and the subsequent recent reannotation of *T. mentagrophytes* genotype VIII as *T. indotineae* (Kano et al. 2020; Tang et al. 2021), species boundaries between the two is difficult phylogenetically as *T. mentagrophytes* and *T. interdigitale* species are conspecific (Pchelin et al. 2019). Further, *T. interdigitale* and *T. indotineae* have been considered as anthropophilic clonal offshoots of *T. mentagrophytes* and the two names have been suggested to be retained primarily to maintain the epidemiological source of infection only; *T. interdigitale* for true anthropophilic and *T. mentagrophytes* for zoophilic infections (Symoens et al. 2011; de Hoog et al. 2017; Pchelin et al. 2019; Taghipour et al. 2019). Hence, considering the overall relatedness of members of *T. interdigitale*/*T. mentagrophytes* species complex, we decided to compare genomic features of all the available *T. interdigitale*/*T. mentagrophytes* strains in NCBI. In order to avoid heterogeneity among genomic datasets that may arise due to different sequencing platforms or different gene prediction pipelines, the complete genome assembly of all available *T. interdigitale*/*T. mentagrophytes* genomes were downloaded from NCBI and a structural and functional annotation was carried out using the same pipeline as described for TiCI12 and TiCI14, before comparative genomics analysis.

Fig. 1 (continued)

Table 1 MIC profile of indicated dermatophytes against different antifungal agents

Strains	MIC_{90} (μg/mL)	
	Terbinafine	Fluconazole
T. mentagrophytes ATCC 18748	0.03	16
TiCI12	0.06	16
TiCI14	> 8	16

The achieved level of genome assembly for TiCI12 or TICI14 is in excellent agreement with the representative genome of RefSeq strain, *T. interdigitale* MR816 (Table 3). A total of 7581 and 7575 protein coding DNA sequences (cds) were predicted and annotated in the genomes of TiCI12 and TiCI14, respectively. A similar number, varying between 7579 and 7823 cds, were predicted for the other genomes except D15P152 (8488 cds) (Table 3). The somewhat higher predicted cds in D15P152, could possibly be due to the low coverage (5X), lower N50 (~ 5 kb) and large number of scaffolds (7988) for this genome (Table 3).

Presence of virulence genes encoded in the genomes of dermatophytes enables them to establish and maintain infection and survive on the outer cornified layers of skin of the host. A comparative analysis of virulence factors in the genomes using online prediction tools can reveal genomic features related to high pathogenicity of strains belonging to different genotypes and different geographical locations. PHI-base v.410 (The Pathogen-Host Interaction database, http://www.phi-base.org/) (Urban et al. 2017) is commonly used to predict the presence of broad spectrum of genes associated with virulence or pathogenicity. PHI-base v.410 did not identify any major difference in predicted pathogenicity, virulence or other effector genes in any of the dermatophytes and revealed a similar number of genes, varying between twenty to twenty seven, in all the genomes, with an identical number of twenty one pathogenicity and/or virulence factors in both TiCI12 or TiCI14 (Table 4).

Secreted lipases and proteases are other important enzymes that not only serve as key virulence factors but also aid the dermatophytic fungi in surviving on the terminally differentiated keratinized layers of skin by deriving nutrients from it (Burmester et al. 2011; Achterman and

Table 2 Standard checklist of the minimal information about any (X) sequence (MIxS) for *Trichophyton indotineae* UCMS-IGIB-CI12 and *Trichophyton indotineae* UCMS-IGIB-CI14

	Trichophyton indotineae UCMS-IGIB-CI12	*Trichophyton indotineae* UCMS-IGIB-CI14
Classification	Kingdom: Fungi Division: Ascomycota Class: Eurotiomycetes Order: Onygenales Family: Arthrodermataceae Genus: Trichophyton Species: *T. indotineae* Strain: UCMS-IGIB CI12	Kingdom: Fungi Division: Ascomycota Class: Eurotiomycetes Order: Onygenales Family: Arthrodermataceae Genus: Trichophyton Species: *T. indotineae* Strain: UCMS-IGIB CI14
Environment		
Collection date	2014-09-25	2014-09-30
Geographical location (latitude and longitude)	28.7041 N 77.1025 E Delhi, India	28.7041 N 77.1025 E Delhi, India
Geographical location (country)	Homo sapiens	Homo sapiens
Environment (biome)	Skin	Skin
Environment (feature)	Lesion	Lesion
Environmental (material)		
MIxS data investigation		
Submitted to insdc	JAATJQ000000000.1 (GenBank)	JAAQVJ000000000.1 (GenBank)
Investigation type	Fungi	Fungi
Project name	Whole genome sequence of *Trichophyton indotineae* CI12 from India	Whole genome sequence of *Trichophyton indotineae* CI14 from India
BioProject	PRJNA604098	PRJNA604102
BioSample	SAMN13951834	SAMN13951939
Depth	143.7×	136.3×
Ploidy	Diploid	Diploid
Isolation and growth condition	SDA with chloramphenicol, gentamicin and cycloheximide	SDA with chloramphenicol, gentamicin and cycloheximide
Sequencing method	Illumina HiSeq SPAdes v. 3.13 assembler	Illumina HiSeq SPAdes v. 3.13 assembler
Assembly	Draft, over 143.7-fold genome coverage, 824 contigs	Draft, over 136.3-fold genome coverage, 904 contigs
Finishing strategy		

White 2012; Martinez et al. 2012) and were next predicted in the genomes. A nearly identical number of lipases were predicted in all the genomes including TiCI12 and TiCI14 with help of lipase prediction server, LED v4.0.0 (https://led.biocatnet.de/sequence-browser) (Table 4). Dermatophytes also encode a large number of proteases in their genomes, with a key role of secretory subtilases in adherence and in initial stages of infection in the host cells (Kaufman et al. 2007; Latka et al. 2015). Among the 174 or 175 total peptidases predicted in the genomes of TiCI12 or TICI14 by the MEROPS server (https://www.ebi.ac.uk/merops/download_list.shtml) (Rawlings et al. 2018), a similar subset of 16 or 18 secreted subtilases were identified, which compares well with the 16–18 secreted subtilases among 167–191 predicted peptidases in all the strains (Table 4).

Among other key virulence factors, LysM domain-containing proteins help in binding of the fungi to N-linked oligosaccharides on the human skin glycoproteins (Kar et al. 2019) and have been reported to be enriched in pathogenic fungi. Genome annotation and function prediction identified eight LysM domain-containing proteins each in TiCI12 and TiCI14, compared to seven in the RefSeq *T. mentagrophytes* MR816 strain. Among other genes associated with virulence, proteins associated with carbohydrate metabolism were predicted as a set of 'carbohydrate-active enzymes' using dbCAN server (http://bcb.unl.edu/dbCAN/index.php) (Lombard et al. 2014). However, no clear distinguishing factor could be identified across different dermatophytes as the predicted set of genes was similar among all dermatophyte genomes (for instance, 163 proteins for TiCI12 or TiCI14 and 165 for RefSeq *T. interdigitale* MR816).

Cytochrome P450 family is another family of proteins that are abundant in fungi due to their roles not only in ergosterol biosynthesis but also in production of secretory secondary metabolites and detoxification of drugs and xenobiotics (Shin et al. 2018). Several cytochrome P450 enzymes (29 each in TiCI12 and TiCI14 and 30 in *T. interdigitale* MR816) were predicted by the Cytochrome P450 domain-containing proteins server (CYPED v6.0, CYtochrome P450 Engineering Database; https://cyped.biocatnet.de/sequence-browser) (Fischer et al. 2007), highlighting the key role of this family of enzymes in dermatophytic fungi.

Table 3 Overall genomic features of *T. indotineae* UCMS-IGIB-CI12 and *T. indotineae* UCMS-IGIB-CI14 and other available genomes in NCBI, of *T. interdigitale/T. mentagrophytes* species complex

Strain name	Species[a]	GenBank accession number (wgs)	Total Sequence length (Mbp)	Depth of coverage (×)	N50 (bp)	Total number of scaffolds	Predicted cds (as per this study)	Reference
MR816 (Ref-Seq)	*T. interdigitale*[a]	AOKY00000000	22.47	56	58,103	956	7621	Persinoti et al. (2018)
H6	*T. mentagrophytes*[a]	AOKS00000000	21.94	91	18,239	2831	7823	Persinoti et al. (2018)
M8436	*T. mentagrophytes*[a]	FUFL00000000	22.59	35	27,151	3285	7674	Gallo et al. (2017)
TiCI12	*T. indotineae*	JAATJQ000000000	22.06	143.7	57,434	824	7581	This study
TiCI14	*T. indotineae*	JAAQVJ010000000	22.04	136.3	54,025	904	7575	This study
D15P127	*T. mentagrophytes*[a]	QQSR00000000	23.71	15	72,043	1842	7585	Pchelin et al. (2019)
D15P135	*T. mentagrophytes*[a]	QQSQ00000000	22.49	12	49,567	1462	7633	Pchelin et al. (2019)
D15P152	*T. mentagrophytes*[a]	QQSP00000000	23.12	5	5,051	7988	8488	Pchelin et al. (2019)
D15P156	*T. mentagrophytes*[a]	QQSO00000000	23.23	12	67,448	1442	7579	Pchelin et al. (2019)
TIMM2789[b]	*T. mentagrophytes*[a]	BFBS00000000	24.07	160	81,080	16,543	7619	Unpublished[c]

[a]Species classification is according to Pchelin et al. (2019)

[b]This entry is flagged by NCBI as 'Anomalous assembly' due to contaminated reads (https://ncbi.nlm.nih.gov/assembly/GCA_003118255.1/). Hence was not considered further in this study

[c]Alshahni et al. (2018), Unpublished work

Table 4 Comparative distribution of predicted virulence factors in the genomes of *T. interdigitale/T. mentagrophytes* species complex

Strain name	Virulence/pathogenicity factors (PHI-base)	Lipases[a]	Peptidase	Subtilisin	LysM	Cytochrome-P450	Carbohydrate active enzymes (dbCAN)
MR816	23	23 (5)	167	16	6	30	140
H6	22	23 (5)	175	16	6	30	138
M8436	22	22 (5)	174	17	6	29	137
TiCI12	21	24 (6)	174	16	8	29	163
TiCI14	21	24 (6)	175	18	8	29	163
D15P127	22	22 (5)	170	17	6	29	139
D15P135	21	23 (5)	172	16	4	28	137
D15P152	27	21 (4)	191	18	6	35	117
D15P156	20	22 (5)	168	16	5	29	139

[a]The predicted secretory subset of lipases are indicated in parenthesis

Although a large number of potential factors associated with pathogenicity and/or virulence were identified in TiCI12 and TiCI14, there was no obvious difference in the predicted proteins in any group when compared to the purported zoophilic *T. mentagrophytes* or the anthropophilic *T. interdigitale* genomes (Table 4). Single nucleotide polymorphisms (SNPs) is a key method to identify any population/geographical/lineage/drug response or pathogenicity variability among different available genomes. The low number of nine wgs available from the genomic resources of EMBL/GenBank/DDBJ, however, limits a statistically significant analysis for SNPs across the genomes of *T. interdigitale/T. mentagrophytes* species complex. Among different types of SNPs in the genomes, the non-synonymous SNPs, designated as single amino-acid polymorphisms (or SAPs), (Kumar et al. 2009) would result in amino-acid

changes and have a direct affect on function of the protein, providing important clues into mechanistic aspects of the disease. Owing to the key role of subtilisins and lipases in nutrient acquisition and survival on host cells, we analyzed SAPs in these two enzyme families in all the genomes. A multiple sequence alignment using CLUSTAL omega (Sievers et al. 2011), for the predicted subtilisins and lipases in all genomes helped identify SAPs in each protein when compared to *T. interdigitale* MR816 (Fig. 2). Predicted subtilisins with GenBank accession numbers KDB23473.1, KDB24326.1, KDB23788.1 and KDB27845.1 (Fig. 2A) and predicted lipases with GenBank accession numbers KDB26859.1 and KDB25940.1 (Fig. 2B) show largest number of SAPs in dermatophytes, as compared to RefSeq *T. interdigitale* MR816. Interestingly, TiCI12, TiCI14 and D15P135 (belonging to *T. mentagrophytes* genotype VIII, (Pchelin et al. 2019)) cluster together for both subtilase and lipase families on the basis of single amino acid change analysis (Fig. 2).

The clustering in subtilase and lipase families of these genomes prompted us to investigate the possible use of the respective subtilase or lipase members for multilocus phylogenetic classification. Functional annotation of KDB23788.1 and KDB24326.1, showing large number of SAPs, identified them as Sub3 and Sub6 subtilisins, respectively. Both proteins play key roles in pathogenesis of dermatophytes. While Sub3 is one of the major subtilases secreted by dermatophytes at alkaline pH, Sub6 is the major protease secreted during infection (Méhul et al. 2016; Gräser et al. 2018).

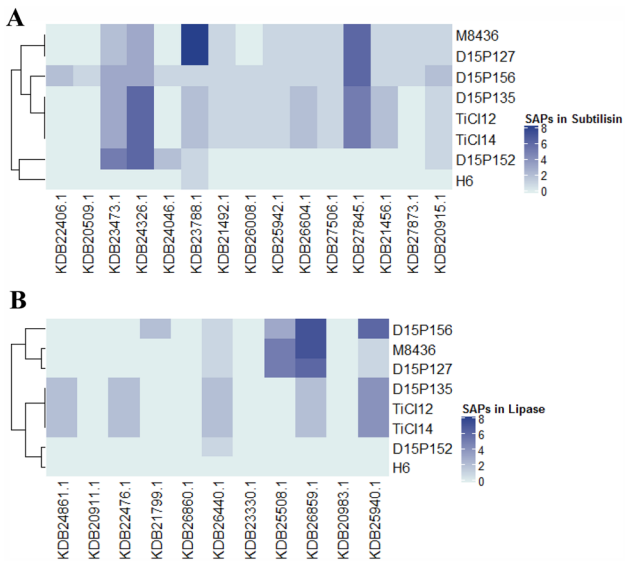

Fig. 2 Heatmap for SAP analysis of (**A**) subtilisin and (**B**) lipase families. Corresponding homologous protein sequences of subtilisins and lipases were identified by local BLAST with *T. interdigitale* MR816 sequences as query in all the indicated genomes. Only full length sequences were used for mapping amino acid polymorphisms in a CLUSTAL omega multiple sequence alignment based SAP analysis. Gene accession numbers of *T. interdigitale* MR816 proteins used as a reference are indicated for **A** subtilisins and **B** lipases. *T. interdigitale*/*T. mentagrophytes* species included in the analysis are as mentioned in Table 3

A maximum likelihood phylogenetic tree after 1000 bootstrap cycles was constructed for Sub3 (KDB23788.1), Sub6 (KDB24326.1) and transcription elongation factor 1-α (TEF1-α) and compared with an ITS-based phylogenetic tree. TiCI12, TiCI14 and D15P135 all cluster together in Sub3- and Sub6-based phylogenetic trees, similar to that of the ITS-based tree (Fig. 3). However, phylogenetic analysis with TEF1-α offered low diversity that could not distinguish among different members of *T. interdigitale*/*T. mentagrophytes* species complex, possibly due to only one identified SAP each in *T. interdigitale* M8436 and *T. mentagrophytes* D15P127 genomes when compared to *T. interdigitale* MR816 strain, suggesting low frequency of SAPs in housekeeping genes. Multilocus phylogenetic analysis earlier had also shown that although topologies of ITS and TEF1-α based trees were congruent, TEF1-α offers relative less diversity as compared to the ITS locus (Nenoff et al. 2019; Tang et al. 2021). It is hence provocative to propose that functionally important proteins (namely Sub3 and Sub6) offer higher diversity among amino acid sequences and may be considered in future multilocus studies for phylogenetic analysis and delineation of *T. interdigitale*/*T. mentagrophytes* species complex.

In conclusion, the overall architecture of the genomes of *T. indotineae* UCMS-IGIB-CI12 and *T. indotineae* UCMS-IGIB-CI14 from India were found to be similar to that of the RefSeq *T. interdigitale* MR816 strain with no major difference in the predicted gene families involved in virulence and/or infection. However, key members of proteases and lipases of TiCI12 and TiCI14 exhibited a higher frequency of SAPs. Phylogenetic analysis with Sub3 and Sub6 subtilases reveals a clustering of strains that was similar to that in ITS-based trees, and further analysis will help in their consideration in multilocus phylogenetic analysis in future. TiCI12 and TiCI14 harbor an Ala448Thr mutation in *erg1* but exhibit a variable response to terbinafine. The whole genome sequences of clinical isolates provided in this report can serve as a key reference point for a thorough and detailed investigation of clinical strains originating from different parts of the world and in devising disease management policies towards emerging resistance cases in India.

Accession numbers

The sequence of the internal transcribed spacer regions 1–2 of 18S rRNA of *Trichophyton indotineae* UCMS-IGIB-CI12 and *Trichophyton indotineae* UCMS-IGIB-CI14 were deposited in GenBank with Accession numbers

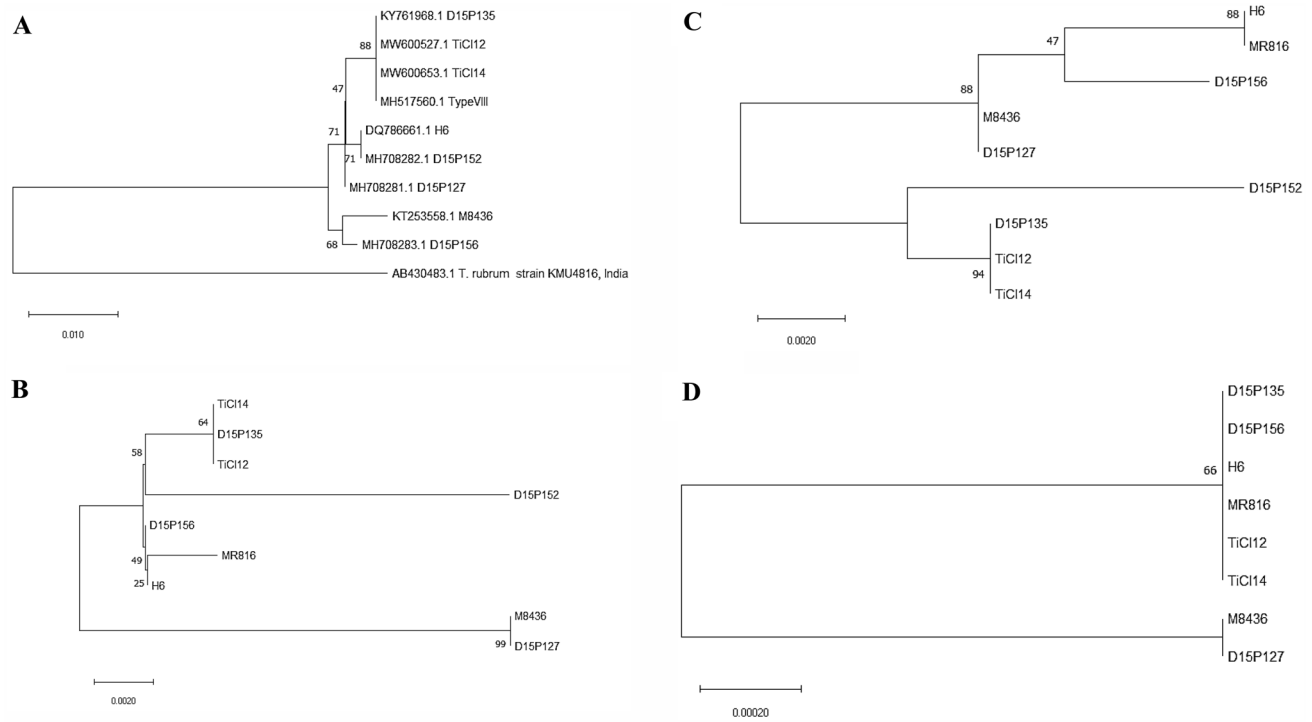

Fig. 3 Maximum likelihood phylogenetic tree different for *T. interdigitale/T.* mentagrophytes. A maximum likelihood phylogenetic tree was constructed in MEGA (Kumar et al. 2018) after 1000 bootstrap cycles for **A** ITS locus, and protein sequences of **B** Sub3, **C** Sub6 and **D** TEF1-α. Only full length sequences were used for in the construction of maximum likelihood phylogenetic trees. As only a partial sequence of TEF1-α of D15P152 was identified (possibly due to the low coverage of this genome), it is not included in the TEF1-α-based phylogenetic tree in (**D**). *T. interdigitale/T. mentagrophytes* species included in the analysis are as listed in Table 3. The *T. mentagrophytes* genotype VIII reference strain (GenBank Accession number: MH517560.1), as reported by Nenoff et al. (2019) is included and indicated as Type VIII in the ITS tree in (**A**)

MW600527 and MW600653, respectively. The whole genome project of *T. indotineae* UCMS-IGIB-CI12 and *T. indotineae* UCMS-IGIB-CI14 has been deposited into GenBank and are available under the accession numbers JAATJQ000000000.1 and JAAQVJ000000000.1 for TiCI12 and TiCI14, respectively.

Acknowledgements PK acknowledges Senior Research Fellowship from CSIR. The authors thank Vinod Scaria, S. Ramachandran and Debasis Dash for useful discussions. We thank CSIR-Institute of Genomics and Integrative Biology for providing the infrastructure and central instrument lab facility.

Author contributions All the authors contributed to this work. PK, SD, SNB and BT conceptualized the work; PK, RT and RP performed the experiments; PK, SD, RT, RP and BT analyzed the data; SD and BT supervised the work; PK and BT wrote the paper. All authors reviewed and approved the manuscript.

Funding This work was supported by funding from Department of Biotechnology (DBT), Government of India. (Grant number: BT/PR20790/MED/29/1130/2016).

Data availability The sequence of the internal transcribed spacer regions 1–2 of 18S rRNA of the isolates TiCI12 and TiCI14 were deposited in NCBI with Accession numbers MW600527 and MW600653, respectively. The complete annotated genome assemblies of TiCI12 and TICI14 have also been deposited at GenBank and are available under the accession numbers JAATJQ000000000.1 and JAAQVJ000000000.1, respectively.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no potential conflict of interest.

Ethical approval This study was approved by the Institutional Human Ethics Committee of UCMS-GTB (IECHR/2016/28/2 dated 27/12/2016) and CSIR-IGIB (CSIR-IGIB/IHEC/2017-18 dated 30/05/2017).

References

Achterman RR, White TC (2012) Dermatophyte virulence factors: identifying and analyzing genes that may contribute to chronic or acute skin infections. Int J Microbiol 2012:358305. https://doi.org/10.1155/2012/358305

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Version FastQC-v.0.11.8. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 4 Oct 2018

Baldo A, Monod M, Mathy A et al (2012) Mechanisms of skin adherence and invasion by dermatophytes. Mycoses 55:218–223. https://doi.org/10.1111/j.1439-0507.2011.02081.x

Bankevich A, Nurk S, Antipov D et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Burmester A, Shelest E, Glöckner G et al (2011) Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. Genome Biol 12:R7. https://doi.org/10.1186/gb-2011-12-1-r7

de Hoog GS, Dukik K, Monod M et al (2017) Toward a novel multilocus phylogenetic taxonomy for the dermatophytes. Mycopathologia 182:5–31. https://doi.org/10.1007/s11046-016-0073-9

Ebert A, Monod M, Salamin K et al (2020) Alarming India-wide phenomenon of antifungal resistance in dermatophytes: a multicentre study. Mycoses 63:717–728. https://doi.org/10.1111/myc.13091

Fischer M, Knoll M, Sirim D et al (2007) The Cytochrome P450 Engineering Database: a navigation and prediction tool for the cytochrome P450 protein family. Bioinformatics 23:2015–2017. https://doi.org/10.1093/bioinformatics/btm268

Gallo JG, Woods M, Graham RM, Jennison AV (2017) A severe transmissible Majocchi's granuloma in an immunocompetent returned traveler. Med Mycol Case Rep 18:5–7. https://doi.org/10.1016/j.mmcr.2017.07.003

Gaurav V, Bhattacharya SN, Sharma N et al (2021) Terbinafine resistance in dermatophytes: time to revisit alternate antifungal therapy. J Mycol Med 31:101087. https://doi.org/10.1016/j.mycmed.2020.101087

Gräser Y, Monod M, Bouchara J-P et al (2018) New insights in dermatophyte research. Med Mycol 56:2–9. https://doi.org/10.1093/mmy/myx141

Kano R, Kimura U, Kakurai M et al (2020) *Trichophyton indotineae* sp. nov.: a new highly terbinafine-resistant anthropophilic dermatophyte species. Mycopathologia 185:947–958. https://doi.org/10.1007/s11046-020-00455-8

Kar B, Patel P, Free SJ (2019) *Trichophyton rubrum* LysM proteins bind to fungal cell wall chitin and to the N-linked oligosaccharides present on human skin glycoproteins. PLoS One 14:e0215034. https://doi.org/10.1371/journal.pone.0215034

Kaufman G, Horwitz BA, Duek L et al (2007) Infection stages of the dermatophyte pathogen *Trichophyton*: microscopic characterization and proteolytic enzymes. Med Mycol 45:149–155. https://doi.org/10.1080/13693780601113618

Khurana A, Masih A, Chowdhary A et al (2018) Correlation of in vitro susceptibility based on MICs and squalene epoxidase mutations with clinical response to terbinafine in patients with tinea corporis/cruris. Antimicrob Agents Chemother 62:e01038-e1118. https://doi.org/10.1128/AAC.01038-18

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073–1081. https://doi.org/10.1038/nprot.2009.86

Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Latka C, Dey SS, Mahajan S et al (2015) Genome sequence of a clinical isolate of dermatophyte, *Trichophyton rubrum* from India. FEMS Microbiol Lett 362:fnv039. https://doi.org/10.1093/femsle/fnv039

Lombard V, Golaconda Ramulu H, Drula E et al (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490-495. https://doi.org/10.1093/nar/gkt1178

Martinez DA, Oliver BG, Gräser Y et al (2012) Comparative genome analysis of *Trichophyton rubrum* and related dermatophytes reveals candidate genes involved in infection. mBbio 3:e00259-e1212. https://doi.org/10.1128/mBio.00259-12

Méhul B, Gu Z, Jomard A et al (2016) Sub6 (Tri r 2), an onychomycosis marker revealed by proteomics analysis of *Trichophyton rubrum* secreted proteins in patient nail samples. J Investig Dermatol 136:331–333. https://doi.org/10.1038/JID.2015.367

Monod M (2008) Secreted proteases from dermatophytes. Mycopathologia 166:285–294. https://doi.org/10.1007/s11046-008-9105-4

Nenoff P, Verma SB, Vasani R et al (2019) The current Indian epidemic of superficial dermatophytosis due to *Trichophyton mentagrophytes*-a molecular study. Mycoses 62:336–356. https://doi.org/10.1111/myc.12878

Pchelin IM, Azarov DV, Churina MA et al (2019) Species boundaries in the *Trichophyton mentagrophytes* / *T. interdigitale* species complex. Med Mycol 57:781–789. https://doi.org/10.1093/mmy/myy115

Peres NT, Sanches PR, Falcão JP et al (2010) Transcriptional profiling reveals the expression of novel genes in response to various stimuli in the human dermatophyte *Trichophyton rubrum*. BMC Microbiol 10:39. https://doi.org/10.1186/1471-2180-10-39

Persinoti GF, Martinez DA, Li W et al (2018) Whole-genome analysis illustrates global clonal population structure of the ubiquitous dermatophyte pathogen *Trichophyton rubrum*. Genetics 208:1657–1669. https://doi.org/10.1534/genetics.117.300573

Rawlings ND, Barrett AJ, Thomas PD et al (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res 46:D624–D632. https://doi.org/10.1093/nar/gkx1134

Rudramurthy SM, Shankarnarayan SA, Dogra S et al (2018) Mutation in the squalene epoxidase gene of *Trichophyton interdigitale* and *Trichophyton rubrum* associated with allylamine resistance. Antimicrob Agents Chemother 62:e02522-e2617. https://doi.org/10.1128/AAC.02522-17

Saunte DML, Hare RK, Jørgensen KM et al (2019) Emerging terbinafine resistance in *Trichophyton*: clinical characteristics, squalene epoxidase gene mutations, and a reliable EUCAST method for detection. Antimicrob Agents Chemother 63:e01126-e1219. https://doi.org/10.1128/AAC.01126-19

Seppey M, Manni M, Zdobnov EM (2019) BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol 1962:227–245. https://doi.org/10.1007/978-1-4939-9173-0_14

Shaw D, Singh S, Dogra S et al (2020) MIC and upper limit of wild-type distribution for 13 antifungal agents against a *Trichophyton mentagrophytes-Trichophyton interdigitale* complex of Indian origin. Antimicrob Agents Chemother 64:e01964-e2019. https://doi.org/10.1128/AAC.01964-19

Shin J, Kim J-E, Lee Y-W, Son H (2018) Fungal cytochrome P450s and the P450 complement (CYPome) of *Fusarium graminearum*. Toxins 10:112. https://doi.org/10.3390/toxins10030112

Sievers F, Wilm A, Dineen D et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75

Singh A, Masih A, Khurana A et al (2018) High terbinafine resistance in *Trichophyton interdigitale* isolates in Delhi, India harbouring mutations in the squalene epoxidase gene. Mycoses 61:477–484. https://doi.org/10.1111/myc.12772

Singh A, Masih A, Monroy-Nieto J et al (2019) A unique multidrug-resistant clonal Trichophyton population distinct from *Trichophyton mentagrophytes*/*Trichophyton inter*digitale complex causing an ongoing alarming dermatophytosis outbreak in India: genomic

insights and resistance profile. Fungal Genet Biol 133:103266. https://doi.org/10.1016/j.fgb.2019.103266

Singh A, Singh P, Dingemans G et al (2021) Evaluation of DermaGenius® resistance real-time polymerase chain reaction for rapid detection of terbinafine-resistant *Trichophyton* species. Mycoses 64:721–726. https://doi.org/10.1111/myc.13271

Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res 33:W465-467. https://doi.org/10.1093/nar/gki458

Symoens F, Jousson O, Planard C et al (2011) Molecular analysis and mating behaviour of the *Trichophyton mentagrophytes* species complex. Int J Med Microbiol 301:260–266. https://doi.org/10.1016/j.ijmm.2010.06.001

Taghipour S, Pchelin IM, Mahmoudabadi AZ et al (2019) *Trichophyton mentagrophytes* and *T interdigitale* genotypes are associated with particular geographic areas and clinical manifestations. Mycoses 62:1084–1091. https://doi.org/10.1111/myc.12993

Tang C, Kong X, Ahmed SA et al (2021) Taxonomy of the *Trichophyton mentagrophytes*/*T. interdigitale* species complex harboring the highly virulent, multiresistant genotype *T. indotineae*. Mycopathologia. https://doi.org/10.1007/s11046-021-00544-2

Törönen P, Medlar A, Holm L (2018) PANNZER2: a rapid functional annotation web server. Nucleic Acids Res 46:W84–W88. https://doi.org/10.1093/nar/gky350

Urban M, Cuzick A, Rutherford K et al (2017) PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. Nucleic Acids Res 45:D604–D610. https://doi.org/10.1093/nar/gkw1089

Walker BJ, Abeel T, Shea T et al (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal rna genes for phylogenetics. PCR protocols. Elsevier, Amsterdam, pp 315–322

White TC, Oliver BG, Gräser Y, Henn MR (2008) Generating and testing molecular hypotheses in the dermatophytes. Eukaryot Cell 7:1238–1245. https://doi.org/10.1128/EC.00100-08

Yamada T, Maeda M, Alshahni MM et al (2017) Terbinafine resistance of *Trichophyton* clinical isolates caused by specific point mutations in the squalene epoxidase gene. Antimicrob Agents Chemother 61:e00115-e117. https://doi.org/10.1128/AAC.00115-17