Genotyping of Russian isolates of fungal pathogen *Trichophyton rubrum*, based on simple sequence repeat and single nucleotide polymorphism

Running title: Population structure in *Trichophyton rubrum*

Ivan M. Pchelin[1], Yuri V. Mochalov[1], Daniil V. Azarov[2], Sofya A. Romanyuk[3], Galina A. Chilina[1], Irina V. Vybornova[1], Tatiyana V. Bogdanova[4], Vasily V. Zlatogursky[5], Svetlana V. Apalko[3], Natalya V. Vasilyeva[1,4], Anastasia E. Taraskina[1]

[1]Kashkin Research Institute of Medical Mycology, North-Western State Medical University named after I.I. Mechnikov, Saint Petersburg, Russia

[2]Department of Epidemiology, Parasitology and Disinfectology, North-Western State Medical University named after I.I. Mechnikov, Saint Petersburg, Russia

[3]City Hospital No. 40, Saint Petersburg, Russia

[4]Department of Medical Microbiology, North-Western State Medical University named after I.I. Mechnikov, Saint Petersburg, Russia

[5]Department of Invertebrate Zoology, Faculty of Biology, St. Petersburg State University, Saint Petersburg, Russia

Correspondence

Ivan M. Pchelin, Kashkin Research Institute of Medical Mycology, North-Western State Medical University named after I.I. Mechnikov, Saint Petersburg, Russia

Email: arcella.oraia@gmail.com

Funding information

Russian Foundation for Basic Research, Grant Number: 18-34-00153

Highlights

- Russian population of fungal pathogen *Trichophyton rubrum* is formed by two global genetic lineages
- Population structure in *T. rubrum* can be studied by improved microsatellite typing assay along with SNP typing and whole genome sequencing
- In *T. rubrum*, there is seemingly no association between genetic lineages and types of infection, places of geographic origin, aldolase gene expression and urease activity
- In *T. rubrum*, microevolution is driven by strong directional selection of yet unknown nature

Abstract

The *Trichophyton rubrum* species group consists of prevalent causative agents of human skin, nail and hair infections, including *T. rubrum* sensu stricto and *T. violaceum*, as well as other less well established or debatable taxa like *T. soudanense*, *T. kuryangei* and *T. megninii*. We aimed to describe genetic lineages in Russian *T. rubrum* sensu stricto, to find whether these lineages possess specific characteristics and to identify factors, shaping the population structure. We assessed polymorphism of 12 simple sequence repeat (SSR, or microsatellite) loci and TERG_02941 protein-coding gene in 70 *T. rubrum* isolates and also performed phylogenomic analysis. In both cases, two genetic lineages were found. The analysis of molecular variance (AMOVA) on the basis of SSR typing data indicated that 22-48% of the variability was between groups within *T. rubrum*. The pattern of phylogenetic changes in polymorphic protein-coding loci and SSR typing results revealed the pressure of strong directional selection, but its sources remained elusive. Our results suggest that Russian

population of *T. rubrum* consists of two cosmopolitan genetic lineages, with no clear connection of population structure with types of infection, places of geographic origin, aldolase gene expression and urease activity.

Keywords

microsatellites, population structure, dermatophyte, *Trichophyton rubrum*

1. Introduction

Trichophyton rubrum is an ascomycete from the order Onygenales. It is considered to be prevalent fungus of medical importance in most countries due to its major contribution to etiology of toenail onychomycosis and foot skin mycosis (Sigurgeirsson, Baran 2014). *Trichophyton rubrum* sensu stricto is a part of species group, including also *T. violaceum* and *T. soudanense* along with several forms, potentially deserving species status. All of these can be differentiated by ribosomal DNA internal transcribed spacer (ITS) region sequences (Summerbell et al. 1999; Li et al. 2008; Zhan et al. 2018; Packeu et al. 2019; Su et al. 2019). Statistical analysis of genetic polymorphism in simple sequence repeats (SSR, or microsatellites) revealed clonal mode of reproduction in the species (Gräser et al. 2007; Gong et al. 2016). It is expectedly accompanied by low genetic polymorphism (Gräser et al. 1999). A number of studies of distribution of mating type locus idiomorphs in *T. rubrum* population found only one mating type locus variant, from two needed for mating (Li et al. 2010; Kano et al. 2013; Persinoti et al. 2018). A certain proportion of *T. rubrum* isolates, belonging to the so-called "raubitschekii" morphotype, possess urease activity. This trait is believed to be ancestral, but has unclear value for taxonomy (Arabatzis et al. 2005).

To the best of our knowledge, by the moment there are two papers, reporting bipartite population structure in *T. rubrum* sensu stricto. In previous study, we reported a concordance between ribosomal non-transcribed spacer region (NTS) genotypes and nonsynonymous single nucleotide polymorphisms (SNPs) in two protein-coding loci, TERG_02941 and TERG_03298 in *T. rubrum* isolates from a city in North-Western Russia (Pchelin et al. 2018a). The two genetic lines were named ST1 and ST2. Also, Zheng et al. (2019) studied a sample of *T. rubrum* isolates, covering major part of China, by whole genome sequencing

(WGS). A number of other studies obtained *T. rubrum* WGS data from wide variety of geographic locations, covering Europe, Asia and North America (Martinez et al. 2012; Latka et al. 2015; Persinoti et al. 2018; Zhan et al. 2018; Zheng et al. 2019). However, there have been no studies, bringing together data on SNP, SSR, WGS and NTS polymorphism in *T. rubrum*.

Our study aimed to describe genetic lineages of *T. rubrum* sensu stricto in Russia, to find whether these lineages possess specific characteristics and, finally, to identify factors, shaping the population structure. We argue that Russian *T. rubrum* population is represented by ST1 and ST2 lineages sensu Pchelin et al. (2018a). This structure could not explain observed differences in types of infection, places of geographic origin, aldolase gene expression and urease activity. The population structure in *T. rubrum* was seemingly shaped by strong directional selection, but its sources remained elusive.

2. Materials and Methods

2.1. Fungal isolates and species identification

Clinical isolates of *T. rubrum* have been obtained in the clinic of Kashkin Research Institute of Medical Mycology, Saint Petersburg (n = 44) and in Sverdlovsk Regional Dermatovenereologic Dispensary, Yekaterinburg (n = 23) from October 2015 to January 2018 (Supplement S1) and subcultured on Sabouraud agar medium, with the addition of 2% dextrose. We also used two reference strains RCPF 1280 and RCPF 1408 and the type strain CBS 392.58 up to a total of 70 strains in the sample. This sample shared 19 isolates with our earlier study (Pchelin et al. 2018a). The following clinical *T. rubrum* isolates were deposited in the Russian Collection of Pathogenic Fungi (RCPF): D15P33, RCPF 1864; D15P44, RCPF 1865; D15P50, RCPF 1866; D15P62, RCPF 1867; D15P71, RCPF 1868; Y1013, RCPF 1869;

D15P139, RCPF 1870. The urease activity was tested twice in Christensen's urea broth, supplemented with chloramphenicol to the final concentration of 40 µg/ml.

Species identification was performed by phenotypic methods and confirmed by DNA sequencing of ITS region. To extract genomic DNA, we used GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Fisher Scientific, Lithuania). Aerial mycelium and fungal spores were scraped by pipette tips, placed into microtubes containing Lysis Buffer A, beaten with two glass beads in Minilys tissue homogenizer (Bertin Technologies, France), supplemented with Lysis Buffer B and RNase A and further treated according to the kit manufacturer's protocol. Amplification of ITS region with ITS5 and ITS4 primers (White et al. 1990) was performed in 50-µl volumes, using standard PCR diagnostics reaction kit (Syntol, Russia). Reaction mixture consisted of 100 mM Tris/HCl (pH 9.0), 250 mM KCl, 10 mM Tween-20, 2.5 mM $MgCl_2$, 250 µM each dNTP and contained 5 pmol of primers (Table 1), 1 U Taq polymerase, and 2.5 µl template DNA. The thermal cycler C1000 Touch (Bio-Rad, USA) was programmed as follows: initial denaturation step for 5 min at 95 °C, 34 cycles of amplification: denaturation for 30 s at 95 °C, annealing 30 s at 56 °C, extension for 1 min at 72 °C, and a final extension step for 10 min at 72 °C. The products of PCRs were purified by isopropanol precipitation in the presence of glycogen. PCR products were sequenced on both strands using BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, USA) on a 3500 Genetic Analyzer (Thermo Fisher Scientific, USA).

Table 1. Primers, used in the study.

Locus	Description	Appli-cation	Sequences, 5'-3' direction	Amplicon length, bp	Reference
ITS	rDNA ITS region	DNA seq	F: GGAAGTAAAAGTCGTAACAAGG R: TCCTCCGCTTATTGATATGC	716	White et al. (1990)
TERG_02941 §	Hypothetical protein	— " —	F: CGGCCTCGTTCTGCGTTAAC R: CTTGCCGGCTTCCAAGTCTTT	449	Pchelin et al. (2018a)
— " —	— " —	TMA	F: TACCACGCCCGTCAAAGG R: CCACTTGTTCTTGCCGGCTT Probe A: **ROX**-TCTTCCGTTTCACTTTGCGAGTG-**BHQ1** Probe G: **FAM**-TCTTCCGTTTCACTTTGCGAGCG-**BHQ2**	71	This paper
TERG_05742	DNA-dependent RNA polymerase II, *rpb2*	qPCR	F: TGCAGGAGCTGGTGGAAGA R: GCTGGGAGGTACTGTTTGATCAA	59	Jacob et al. (2012)
TERG_07904	tubulin beta chain, β-*tub*	— " —	F: AACATGATGGCTGCCACTGA R: AAGATGGCAGAGCAGGTAAGGT	253	— " —
TERG_06744	fructose 1,6-bisphosphate aldolase	— " —	F: TACTTTGCCGGAAAGGGAGT R: TTCTTAGCACAGTGGTCGGT	139	This paper

DNA seq, Sanger DNA sequencing; TMA, TaqMan SNP genotyping assay; qPCR, quantitative real-time PCR. § Pchelin et al. (2018a) referred to this locus as A7C99_6411.

2.2. Antifungal susceptibility testing

Antifungal susceptibility testing was performed according to EUCAST E.Def 9.3.1 document (Arendrup et al. 2017), in RPMI-1640 medium with the addition of glucose to a final concentration of 2%. We tested the activity of terbinafine, fluconazole and itraconazole against six *T. rubrum* isolates with newly obtained WGS sequences. Terbinafine was tested at concentration range 0.007-4 μg/ml, fluconazole and itraconazole were tested at 0.015-8 μg/ml.

2.3. Molecular strain typing

We employed two strain typing approaches. Firstly, we probed the G793A nucleotide substitution in protein-coding locus TERG_02941 (A7C99_6411). The amplification was carried out under following conditions: 5 min at 95 °C; 39 cycles of 30 s at 95 °C, 1 min 5 s at 64 °C; and final elongation for 5 min at 72 °C (Pchelin et al. 2018a). All other steps were performed as described in the section 2.1 for ITS region sequencing. To optimize the genotyping process, we developed TaqMan SNP genotyping assay with two allele-specific

probes (Table 1). Samples were run using a C1000 Touch thermal cycler equipped with a CFX96 real-time PCR detection system (Bio-Rad, USA) according to the following program: initial denaturation step for 3 min at 95 °C; and 39 cycles of 10 s at 95 °C and 30 s at 59 °C.

Secondly, we performed microsatellite typing. We used eight microsatellite loci from the papers by Gräser et al. (2007) and Gong et al. (2016) and developed four additional loci (Table 2). To find new polymorphic loci, we searched genome contigs for repeat sequences and made nucleotide alignments. In total, 170 loci were screened. Primers for amplification of nine candidate loci were designed with the use of Primer-BLAST tool (Ye et al. 2012). Five loci had uniform lengths in 21-56 isolates and were excluded from further experiments. Amplification of microsatellite loci was performed in 25-µl volumes. The primers used are listed in the Table 2.

Table 2. Primers, used for SSR amplification in *Trichophyton rubrum* strains. Fluorescent modifications are given in bold.

Locus	Sequences, 5'-3' direction	Amplicon length, b.p.	Repeat unit	Length of flanking regions, b.p.	Reference
Tr001	F: **FAM**-GTGGGCTCGAACCGAGACAAAC R: GGTGACTGGAGGGTGAAGGG	263–286	GA	212	Gong et al. 2016
Tr002	F: **HEX**-CTCGCATTTGGCCTTGAGTGT R: CTGGCTGGTGGAAGGACATAG	248–281	CTT, ACT	138	— " —
Tr003	F: **FAM**-GCCAGTGATCCTGCATCGTCC R: CGGGCTTCTCGTCTTGCTCTTC	233–242	ACT	187	— " —
Tr004	F: **HEX**-CAGGAGGGTAGAAGACGAGAAG R: AAAGAAAAACAAAGTCAAGCACGC	295–300	GCA, GAA	228	— " —
Tr005	F: **FAM**-AAAGTACCTCCTGTTTACCATC R: GGCATTCCGGGCAGTTGAAGC	304–310	CT	276	— " —
Tr006	F: **HEX**-TAAGACAGGGATCTCAGAATG R: CGGAACCGCTACAAGCAGAAT	191–214	GTT	133	— " —
Tr007	F: **TAMRA**-TGGGATACAGCCATGCCGTT R: CTCAGCCCTACCAGACGACC	361–364	AG	325	This paper
Tr008	F: **TAMRA**-AGGCCCCCAAGGTAGAGGAT R: GTGCCTGCGCGAAATACGAA	195–204	GA	151	— " —
Tr009	F: **HEX**-ACGACGATACTGGCTGGACAA R: GTTATGACTTCCACGGTGTCACG	348–368	AAG	318	— " —
$Tr(CT)_{20}$	F: **FAM**-GCGGTGGCGTCTTCTATC R: CAGCAGACAACATAGCAGTCG	134–139	GA	101	Gräser et al. 2007
$Tr(GA)_{25}$	F: **HEX**-GCGATGGTTGGAGGAGTTG R: GCCTGTCGCTGCTTACTTG	143–150	GA	92	— " —
STR*Tr*9	F: **TAMRA**-ACACTGCTCTTCGCATCTTGA R: GGCTGTTGTTGTTCGCTGCT	240–342	CAAG	197	This paper

The amplification program included initial denaturation at 95 °C for 10 min, followed by 30 cycles of 50 s at 95 °C, 1 min at 58 °C, and 1 min at 72 °C, and a terminal extension step of 72 °C for 10 min. The set of fluorescent dyes used in this work appeared to be suboptimal because of the need to check data for issues of spectral overlapping between TAMRA and ROX channels.

2.4. Whole genome sequencing

Genomic DNA of six *T. rubrum* and one *T. benhamiae* isolates was extracted using GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Fisher Scientific, Lithuania). Preparation of DNA sample library was done with the use of Kapa HTP Library Preparation Kit (Roche, USA), according to the manufacturer's instructions. Paired end sequencing (250 × 250 cycles) was performed using the MiSeq sequencer (Illumina, USA). The reads obtained were prepared for the analyses as described in the paper by Pchelin et al. (2019a). Quality control of the reads was performed using FastQC 0.11.8 (Andrews 2010). It was followed by trimming low-quality reads and TruSeq2 adapters with Trimmomatic 0.39 (Bolger et al. 2014). De novo assembly was performed with SPAdes 3.9.0 with the "--careful option" (Nurk et al. 2013). Raw reads were submitted to the Sequence Read Archive of National Center for Biotechnology Information (BioProject ID PRJNA552357).

2.5. Phylogenomic reconstruction

The whole genome dataset included six genomes of Russian isolates, as well as the genomes of isolate CBS 118892, BioProject accession ACPH00000000 (Martinez et al. 2012), isolates CBS 202.88 AOKX00000000, CBS 288.86 AOKU00000000, CBS 289.86 AOKV00000000, CBS 100081 AOKT00000000, D6 AOLB00000000, MR850 AOKR00000000, MR1448 AOKZ00000000, MR1459 AOLC00000000 (Persinoti et al. 2018), isolate CMCC(F)T1i LHPM00000000 (Zhan et al. 2018), IGIB-SBL-CI1

JPGR00000000 (Latka et al. 2015). The raw read data for eight strains from the project of Zheng et al. (2019) have been downloaded from NCBI SRA database and assembled *de novo* for the analysis. This number included the strains LZ164 (NCBI SRA accession SRX5814844), GZ0580 (SRX5814815), NJ785 (SRX5814822) from one cluster, and the strains LZ155 (SRX5814845), NJ786 (SRX5814823), NJ823 (SRX5814824), DL7506 (SRX5814836) from another cluster, revealed in the original publication. The eighth strain GZ0663 (SRX5814819) from the study of Zheng et al., analyzed here, had unclear cluster affiliation. The sequences of *T. rubrum* morphotype "megninii" CBS 735.88 JHQM00000000, *T. soudanense* CBS 452.61 AOKW00000000 (Persinoti et al. 2018) and *T. violaceum* CMCC(F)T3l LHPN00000000 (Zhan et al. 2018) were used as an outgroup. Single-nucleotide polymorphism analysis in core genome was done in Harvest Parsnp 1.2 software (Treangen et al. 2014). Parsnp run with *T. rubrum* CBS 118892 assembly GCA_000151425.1 as a reference and with default parameters. Whole-genome phylogenetic tree was reconstructed by Maximum Likelihood method in FastTree2 (Price et al. 2010), as implemented in Harvest suite. Branch support values were calculated by performing Shimodaira-Hasegawa test on three alternate topologies around particular splits, with a total of 1000 resamples.

2.6. Analysis of genetic polymorphism

A matrix of mutation probability-based genetic distances (Bruvo et al. 2004) was calculated in polysat 1.7-3 package for R (Clark, Jasieniuk 2011) using microsatellite polymorphism data. Distances in terms of weighted divergence of microsatellite loci were calculated in Microsoft Excel 2010 according to the formula, modified for haploids from Rozenfeld et al. (2007). The dissimilarity d_i between haplotypes *A* and *B* at particular locus *i*

was calculated as absolute value of arithmetic difference between the lengths of PCR products A_i and B_i

$$d_i(A, B) = |A_i - B_i|$$

The genetic distance *D* between haplotypes *A* and *B* was based on the sum of dissimilarities *d* between haplotypes at all loci, weighted by dividing this sum by the number of loci taken into analysis, *k*

$$D(A, B) = \frac{1}{k}\sum_{i=1}^{k} d_i$$

Both distance matrices were further employed in the construction of phylogenetic network in SplitsTree 4.14.2 program (Huson, Bryant 2006). Neighbor-Joining tree on the basis of Bruvo distances was calculated in Mega 5.2 (Tamura et al. 2013).

The allele association index was calculated from the lengths of microsatellite loci in poppr 2.8.1 package for R (Kamvar et al. 2014), using the Multilocus Style Permutation method, after deleting one non-informative locus and performing clonal correction. The null hypothesis was the absence of association between the lengths of microsatellite loci. Simpson's index of genotype diversity within upper and lower bounds for a 95% confidence interval was calculated in the package polysat on the basis of Bruvo genetic distances, in equal groups of 21 isolates for ST1 and ST2 lineages. The analysis of molecular variation (AMOVA) was done on the basis of Bruvo and Rozenfeld distance matrices in GenAlEx 6.5 package (Peakall, Smouse 2006, 2012) with 999 standard permutations.

TERG_03298 genotypes, MAT idiomorphs and TERG_05790 urease gene variants were determined through Blast searches in whole genome sequences (Zhang et al. 2000). For

mating type determination, TERG_02406 gene coding for α-domain protein was probed by AB759084 sequence (Kano et al. 2013). The absence of HMG protein-coding gene was verified by the searches with *T. simii* HMG gene, GenBank accession number AB605766 (Kano et al. 2012). The boundaries of the minisatellite in TERG_08771 gene were determined with the use of Tandem Repeats Finder 4.09 (Benson 1999).

To address evolutionary pattern in *T. rubrum* complex, we prepared a phylogenetic tree on the basis of seven protein-coding genes sequences with predicted nonsynonymous mutations among *T. rubrum* complex isolates. This number included genes, coding for hypothetical proteins TERG_02941 (A7C99_6411), TERG_03298 (A7C99_6714), TERG_03748, copper radical oxidase TERG_02718, LIM domain-containing protein TERG_02913, mismatch repair protein TERG_03581, serine/threonine protein kinase TERG_03779. The same sequences were extracted by blast searches from the genomes of *T. mentagrophytes*, *T. interdigitale*, *T. tonsurans*, *T. equinum*, *T. benhamiae*, *T. verrucosum* and *Nannizzia gypsea* (Martinez et al. 2012; Gallo et al. 2017; Alshahni et al. 2018; Persinoti et al. 2018; Pchelin et al. 2019b).

2.7. Quantitative RT-PCR

The sample of six isolates was grown on Sabouraud agar medium for two weeks at 28 °C. RNA was isolated with the use of GeneJet Plant RNA Purification Mini Kit (Thermo Fisher Scientific, Lithuania). The quantitative RT-PCR experiments were performed in triplicates on two independently grown colonies for each isolate in the sample with the use of OneTube RT-PCR SYBR Kit (Evrogen, Russia). The reactions were carried out on CFX96 Touch Real-Time PCR Detection System (Bio-Rad, USA) with the following cycling conditions: reverse transcription step at 55 °C for 15 min, polymerase activation and reverse transcriptase inactivation step at 95 °C for 1 min, and main amplification stage for 40 cycles,

including denaturation at 95 °C for 15 s, annealing at 60 °C for 20 s and elongation at 72 °C for 15 s. Melting curve analysis was performed to check whether one PCR product was obtained. Relative gene expression levels were assessed using beta-tubulin and DNA-dependent RNA polymerase II for normalization. The sequences of primers used are given in the Table 1.

3. Results

The majority of 69 *T. rubrum* isolates from Russia and the type strain CBS 392.58 had the same ribosomal ITS region sequence, identical to KT285224. There were two exceptions, the isolate D15P139 had 247insC mutation, and another isolate had additional TA element in the microsatellite near the 3'-end of the sequence (Fig. S1). Our preliminary study showed discrepancies between the results of strain typing by TERG_02941 locus polymorphism and the microsatellite typing assay of Gong et al. (2006), based on eight loci (Pchelin et al. 2018b). We hypothesised that this discrepancy was due to insufficient informativeness of the assay and developed four new polymorphic loci, up to a total of 12. The stability of microsatellite loci was assessed in 20 isolates for the time span of 1.5 years. Observed differences of less than one repeat length were considered insignificant, since those differences could be particularly associated with the rounding of measurements, fluctuating near the half-integers. At the endpoint, mutations were observed in five loci. The most extensive changes were noticed in $Tr(CT)_{20}$ locus, with 30% of isolates with mutations. The most pronounced indels were in Tr006 and Tr008 loci, covering up to 3 repeat lengths.

Our first approach to calculating microsatellite polymorphism-based genetic distances employed the probability of mutation events, described with the use of stepwise mutation model (Bruvo et al. 2004). The split network, inferred using mutation probability-based distances, was divided into three major partitions. One partition contained TERG_02941

793A isolates and another two partitions contained TERG_02941 793G isolates (Fig. 1). The second approach was based on the comparison of the lengths of microsatellite loci (Rozenfeld et al. 2007). The resulting phylogenetic network also had pronounced division into TERG_02941 793A and 793G parts (Fig. 2). In Rozenfeld distances-based network, further subdivision of TERG_02941 793G partition was less evident, but still discernable. From 22 TERG_02941 793G isolates, 18-21 located to the same subgroups in both networks.

For the two genetic lines, the average lengths of PCR products of Tr001, Tr002, Tr003, Tr004, Tr005, Tr007, Tr008, Tr009, Tr$(CT)_{20}$ and Tr$(GA)_{25}$ loci differed by less than one repeat unit length. Average Tr006 lengths differed by a unit length. For STR*Tr*9 locus, this difference was 7.5 unit lengths (Fig. S2A). This difference should be obviously explained by the pressure of natural selection. Since the analysis of intraspecific divergence in most cases is carried out using neutral loci (Leliaert et al. 2014), we performed calculation of genetic distances without STR*Tr*9 length data. This broke clear grouping of isolates in both phylogenetic networks. However, a certain tendency to clustering of TERG_02941 793G isolates persisted (Figs. 3, 4).

The analysis of molecular variance with 12 microsatellite loci revealed neglectable contribution of TERG_02941 793G subdivision to intergroup and overall variability (Table 3). With the removal of STR*Tr*9 from the analysis, the percent of variability, explained by the differences between groups, dropped drastically. But, this was accompanied by significant increase in explanatory power of TERG_02941 793G group subdivision. A half of the between-group variability in the dataset with 11 loci depended on it.

Table 3. The analysis of molecular variance (AMOVA) in 70 *T. rubrum* isolates, performed on microsatellite polymorphism-based genetic distances. The numbers reflect percent of variability among TERG_02941 793A and 793G groups of isolates from the total amount of variability between and within groups. To obtain three groups, TERG_02941 793G isolates were further divided according to the topology of Bruvo distances-based network (left-hand numbers in columns with data on 3 lineages) or by Rozenfeld distances-based network (right-hand numbers).

	Bruvo et al. (2004)		Rozenfeld et al. (2007)	
	2 groups	3 groups	2 groups	3 groups
12 loci	22	27-24	48	48-48
11 loci	9	18-14	10	19-20

The CAAG repeats of STR*Tr*9 ended at the position -278 in the 5'-region of fructose 1,6-bisphosphate aldolase gene TERG_06744, wherein the adenine of the ATG translation initiation codon is considered +1. Therefore, we hypothesized that STR*Tr*9 locus promotes microevolution in *T. rubrum* through regulation of aldolase gene expression, which in its turn is one of the key enzymes of glycolysis (Nelson, Cox 2008). We determined relative TERG_06744 gene expression in six *T. rubrum* isolates. The isolate D15P44 with maximal length of STR*Tr*9 locus demonstrated the lowest level of TERG_06744 transcripts (Fig. S2B). The highest levels of TERG_06744 transcripts were observed in D15P62 isolate with the microsatellite length, very close to minimal one. In the same time, three isolates with the same STR*Tr*9 length demonstrated significantly lower level of TERG_06744 gene expression. Also, there were no clear differences between TERG_06744 transcript levels in isolates of different sequence types. Since urease activity has been regarded as phylogenetically important trait in the dermatophytes (Summerbell 2000), we assessed urease activity in our sample and found three urease-positive strains (Supplement S1).

In the sample of 70 isolates, the association index value $\bar{r}_d$ = 0.0248 lay beyond the right boundary of the distribution of expected frequencies. The probability $\bar{r}_d$ was estimated as

0.002, which made it possible to reject the null hypothesis of the absence of connections between the markers. Hence, the studied *T. rubrum* population was considered to be clonal. Also, six Russian isolates, subjected here to whole genome sequencing, harbored *MAT1-1* idiomorph of mating type locus. Simpson index was calculated on the basis of data from 11 neutral SSR loci. Its value within 95% confindence interval was estimated as 0.59-1.03 for ST1 lineage and 0.38-0.91 for ST2 lineage. The Fisher exact test showed no significant association between TERG_02941 genotypes and the localization of the lesion on patient's body. Also, Yates-corrected χ^2 test showed no connection between TERG_02941 genotype and a city of origin.

Russian sample of *T. rubrum* isolates was placed in global context by whole genome sequencing in six *T. rubrum* isolates. To get a representative sample, we selected isolates D15P33, D15P44, D15P50, D15P62, D15P71 and Y1013, covering the whole span of microsatellite distances-based Neighbour-Joining tree (not shown). These isolates were susceptible to tested antifungals (Table 4).

Table 4. Annotation and antifungal susceptibility data for six *T. rubrum* isolates, subjected to whole genome sequencing.

Isolate	Year of isolation	Place of isolation	Clinical picture	MIC_{50} of antifungals (μg/ml)		
				TER	FCZ	ICZ
D15P33	2015	St. Petersburg	OM, TP	≤ 0.01	0.125	0.06
D15P44	2015	St. Petersburg	OM	≤ 0.01	0.25	0.125
D15P50	2016	St. Petersburg	OM	≤ 0.01	0.125	0.06
D15P62	2016	St. Petersburg	OM	≤ 0.01	0.25	0.06
D15P71	2016	St. Petersburg	OM	≤ 0.01	0.06	0.125
Y1013	2017	Yekaterinburg	OM	≤ 0.01	0.25	0.125

Clinical picture: OM, toenail onychomycosis; TP, tinea pedis. Antifungals: TER, terbinafine; ICZ, itraconazole; FCZ, fluconazole.

The phylogenomic tree included *T. rubrum* sensu stricto genomes from Europe, Asia and North America (Fig. 5) and contained well-supported monophyletic branch of ST1

isolates and basal paraphyletic ST2 branch. Both branches contained European and Asian isolates. In a representative selection of seven isolates from the study of Zheng et al. (2019), cluster affiliation, determined in the original study, matched their TERG_02941/03298 genotypes. Depending on the inclusion or the exclusion of *T. violaceum* CMCC(F)T3l and *T. rubrum* “megninii” CBS 735.88 from the phylogenomic analysis, the depth of branches within *T. rubrum* sensu stricto clade related to *T. rubrum* sensu stricto — *T. soudanense* CBS 452.61 branch as 1:63 or 1:337.

The study of Zheng et al. (2019) has associated two SNPs in TERG_08771 with voriconazole resistance. Unfortunately, these mutations were located in continuous degenerated minisatellite, spanning 1699 bp in long read-based assembly LHPM01 and 2248 bp in reference assembly ACPH01. We were unable to assemble the whole span of TERG_08771 gene unambiguously and perform the alignment, needed to determine the mutations. Also, in the same study the position NW_003456428.1 1020017, residing in TERG_00393 gene, has been associated with griseofulvin resistance. In our dataset, this position fell into a microsatellite and wasn’t polymorphic.

There is an observation of the 68-bp intron and the 9-bp insertion in a urease-positive *T. rubrum* strain (Hiruma et al. 2013). However, similarly to the results of Adamski et al. (2014), we found identical urease gene sequences in all *T. rubrum* sensu stricto isolates, shown on the Fig. 5, including urease-positive D15P50 and CBS 202.88. *Trichophyton rubrum* "megninii" CBS 735.88 (Persinoti et al. 2018) harbored two point mutations in the urease gene, A685G in the intron and synonymous substitution G803A in the second exon. All our three isolates, positive for urease activity, came from the cases of toenail and foot skin infections and belonged to both major genetic lineages of *T. rubrum*.

We studied seven protein-coding loci with predicted nonsynonymous SNPs, differentiating *T. rubrum* strains: TERG_02718, TERG_02913, TERG_02941, TERG_03298, TERG_03581, TERG_03748 and TERG_03779 (Fig. 6). The nonsynonymous SNPs in loci TERG_02718, TERG_02941, TERG_03298, TERG_03581 and TERG_03748 allowed distinguishing groups of isolates, whereas mutations in TERG_02913 and TERG_03779 occurred in single isolates. Unlike *T. mentagrophytes* / *T. interdigitale* species group isolates, *T. rubrum* isolates demonstrated consecutive branching pattern, suggestive of strong directional selection of discriminative SNPs (Ito, Dieckmann 2007; Edwards 2009).

In our previous work, we associated single-banded TRS-1 pattern with the length of approximately 450 bp with particular TERG_02941/03298 genotype, substantiating the presence of ST1 and ST2 genetic lines in Russian population of *T. rubrum* (Pchelin et al. 2018a). Later on, Zheng et al. 2019 observed the subdivision of Chinese *T. rubrum* population into two groups. Here through phylogenomic reconstruction we showed that those groups not only match Russian ones, but also have cosmopolitan distribution. Therefore we believed that mapping global distribution of *T. rubrum* sequence types sensu Pchelin et al. (2018a) is possible (Fig. 7). The map of global distribution was prepared with the use of strain typing data from the following publications: Jackson et al. (2000), Iceland, Japan, the UK; Baeza et al. (2006), Brazil; de Assis Santos et al. (2007), Brazil; Rad et al. (2008), Iran; Hryncewicz-Gwóźdź et al. (2011), Poland; Takahashi et al. (2014), Japan; Ramaraj et al. (2017), India; Pchelin et al. (2018a), Russia; Suzuki et al. (2018), Japan; Zheng et al. (2019), China.

4. Discussion

Until recently, low genetic diversity of *T. rubrum* precluded persuasive description of its population structure. For this species, our earlier work (Pchelin et al. 2018a) seemingly represented the first successful attempt to obtain congruent results by two independent

molecular strain typing techniques. However, we have not performed DNA sequencing of ribosomal ITS region to clarify whether the sample contained only *T. rubrum* sensu stricto isolates or not. Here we filled this gap by demonstrating that Russian population of the fungus is represented by only «classic» *T. rubrum* genotype. The two genetic lines of Russian *T. rubrum* population, found here, demonstrated significant divergence. The AMOVA analysis showed relatively high proportion of variability, explained by the division of *T. rubrum* sample into ST1 and ST2 groups, 22-48% depending on the distance metrics. It was not less, than in other studies on clonal filamentous ascomycetes with 19-21% of variability between groups (Ahmadpour et al. 2017; Singh et al. 2019) or filamentous ascomycetes with mixed reproduction mode, with 2-13% of total variability explained by differences between intraspecific groups (Kashyap et al. 2020; Sexton & Howlett, 2004).

Intraspecific genetic lines in fungi sometimes demonstrate restricted geographic ranges (Abdel-Rachman et al. 2010; Mironenko et al. 2016; Gultyaeva et al. 2017; Taghipour et al. 2019). Here, we compiled the map of global distribution of *T. rubrum* genetic lineages (Fig. 7). We largely built upon existing literature on NTS typing in *T. rubrum* and departed from the fact that single 450 bp PCR product of TRS-1 locus is associated with ST1 genetic lineage (Pchelin et al. 2018a). We didn't observe a specific distribution pattern of the two genetic lineages, but in most parts of the World, the derived group of *T. rubrum*, called ST1, presumably had higher prevalence and therefore demonstrated evolutionary success (Fig. 7). Therefore, we studied the pattern of evolutionary changes in *T. rubrum*. The branching pattern of phylogenetic tree, based upon seven polymorphic protein-coding loci consisted of consequential branching events (Fig. 6). There were no inner branches, as in the *T. mentagrophytes* / *T. interdigitale* species group. This pattern can be explained by clonal divergence, driven by directional selection. Intriguingly, this pattern obviously can't be

explained neither by the selective pressure of climatic factors, nor by the adaptation to the host.

The affiliation of *T. rubrum* isolates to genetic lineages can be probed not only by NTS, SNP and WGS-based typing techniques, but also by our modified scheme of microsatellite analysis. But, this result may be of limited interest for healthcare practitioners, since there is seemingly no association of lineage affiliation and common isolate descriptors, like geography of origin and clinical picture of respective infection. Also, in the study of Zheng et al. (2019), there was no association between cluster affiliation and antifungal susceptibility levels. It is widely accepted that truly independent species of clonal microbes occupy distinct ecological niches (Tibayrenc, 2006; Achtman & Wagner, 2008; Giraud et al. 2008). In present study, there seemingly have been no clinical differences between the subpopulations, contrary to what we has claimed earlier (Pchelin et al. 2018a). In that previous study, the association of STs and clinical forms of dermatophytosis might have occurred due to some random factors. From this perspective, the two genetic lines described here belong to the same fungal species.

The STR*Tr*9 locus was paramount for the analysis of the structure of the *T. rubrum* population by microsatellite analysis. The fact that its length is significantly different between the isolates of two genetic lines meant that it was under the pressure of natural selection. The regulatory role of this microsatellite could not be confirmed experimentally. But, there were significant differences in relative aldolase gene expression levels between different isolates. Also, in our sample there were three urease-positive isolates, a proportion of 4% from the total. We did not find peculiarities in urease gene sequences of urease-positive isolates D15P50 and CBS 202.88. Positive strains belonged to both ST1 and ST2 groups. Therefore,

our results suggest that urease-positive strains present in certain proportion in the population of the fungus, and represent natural variation of *T. rubrum* sensu stricto, not an independent entity. Lackner et al. (2019) obtained comparable results, studying virulence potential of closely related *Aspergillus* species from the section *Terrei*. They concluded that virulence is a strain-specific rather than species-specific feature. Lineage-specific differences in STR*Tr*9 locus length may still hallmark different adaptation strategies, given that glycolysis pathway regulation is connected with mechanisms of keratin degradation (Martins et al. 2020).

Urease activity is an ancestral trait, commonly observable in geophilic dermatophytes (Summerbell 2000) and we expected to find it only in basal ST2 lineage. However, in our sample one of three urease-positive isolates belonged to ST1 lineage. Also, genetic diversity in the two lineages proved to be the same, which is contrary to what might be expected if ST2 is an ancestral lineage. It can be probably due to a sort of bottleneck, occurred in the population on the way from South-East Asia to Russia. In future studies, it is necessary to sample isolates from South and South-East Asia, in order to find the territory with highest genetic diversity and thereby conclusively determine the geographic region of the origin of the species.

Thus, we provided the description of the two intraspecific global lineages of *T. rubrum* sensu stricto. They can be revealed by SNP, SSR and WGS-based typing techniques (and NTS, with a certain amount of ambiguity). This population structure is shaped by pressure of natural selection of yet unknown nature. Also, the species is polymorphic in terms of ecology and physiology, but this polymorphism is not associated with genetic population structure.

Acknowledgements

The authors would like to thank the CBS-KNAW culture collection of Westerdijk Fungal Biodiversity Institute, Utrecht, the Netherlands, for providing the strain CBS 392.58. We also appreciate valuable comments from Adéla Čmoková and Ali Rezaei-Matehkolaei. This work was supported by Russian Foundation for Basic Research [grant number 18-34-00153].

Author contributions

I.P. designed the study. I.P., Y.M., I.V., G.C., S.R. and T.B. performed the experiments. I.P., D.A. and Y.M. analyzed the data with the participation of A.T. N.V., A.T. and S.A. performed supervision and funding acquisition. I.P. wrote the manuscript with the help of V.Z.

Declarations of interest:

none

Data accessibility

Raw WGS reads were submitted to NCBI SRA database, BioProject PRJNA552357: the isolate D15P33, accession SRX6389058; D15P44, SRX6389060; D15P50, SRX6389057; D15P62, SRX6389059; D15P71, SRX6389062; Y1013, SRX6389061.

ORCID

Svetlana V. Apalko https://orcid.org/0000-0002-3853-4185
Daniil V. Azarov https://orcid.org/0000-0003-2483-5144
Ivan M. Pchelin https://orcid.org/0000-0003-3062-3316
Vasily V. Zlatogursky https://orcid.org/0000-0002-2688-3900
Anastasia E. Taraskina https://orcid.org/0000-0002-1725-8433
Natalya V. Vasilyeva https://orcid.org/0000-0003-3693-5468

References

Abdel-Rahman, S. M., Sugita, T., González, G. M., Ellis, D., Arabatzis, M., Vella-Zahra, L., ... Preuett, B. (2010). Divergence among an international population of *Trichophyton tonsurans* isolates. *Mycopathologia, 169*(1), 1–13. https://doi.org/10.1007/s11046-009-9223-7

Achtman, M., & Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology, 6*(6), 431–440. https://doi.org/10.1038/nrmicro1872

Adamski, Z., Kowalczyk, M. J., Adamska, K., Kubisiak-Rzepczyk, H., Bowszyc-Dmochowska, M., Banaszak, A., … Zaba, R. (2014). The first non-African case of *Trichophyton rubrum* var. *raubitschekii* or a urease-positive *Trichophyton rubrum* in Central Europe? *Mycopathologia, 178*(1-2), 91–96. https://doi.org/10.1007/s11046-014-9751-7

Ahmadpour, A., Castell-Miller, C., Javan-Nikkhah, M., Naghavi, M. R., Dehkaei, F. P., Leng, Y., ... Zhong, S. (2018). Population structure, genetic diversity, and sexual state of the rice brown spot pathogen *Bipolaris oryzae* from three Asian countries. *Plant Pathology, 67*(1), 181–192. https://doi.org/10.1111/ppa.12714

Alshahni, M. M., Yamada, T., Yo, A., Murayama, S. Y., Kuroda, M., Hoshino, Y., … Makimura, K. (2018). Insight into the draft whole-genome sequence of the dermatophyte *Arthroderma vanbreuseghemii*. *Scientific Reports, 8*(1), 15127. https://doi.org/10.1038/s41598-018-33505-9

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. accessed: 22.02.2020.

Arabatzis, M., Velegraki, A., Kantardjiev, T., Stavrakieva, V., Rigopoulos, D., & Katsambas, A. (2005). First report on autochthonous urease-positive *Trichophyton rubrum* (*T. raubitschekii*) from South-east Europe. *British Journal of Dermatology, 153*(1), 178–182. https://doi.org/10.1111/j.1365-2133.2005.06615.x

Arendrup, M. C., Meletiadis, J., Mouton, J. W., Lagrou, K., Hamal, P., Guinea, J., … Rogers, T. (2017). EUCAST definitive document E.Def 9.3.1 method for the determination of broth dilution minimum inhibitory concentrations of antifungal agents for conidia forming moulds. *EUCAST*.

Baeza, L. C., Matsumoto, M. T., Almeida, A. M., & Mendes-Giannini, M. J. (2006). Strain differentiation of *Trichophyton rubrum* by randomly amplified polymorphic DNA and analysis of rDNA nontranscribed spacer. *Journal of Medical Microbiology, 55*(4), 429–436. https://doi.org/10.1099/jmm.0.46236-0

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acid Research, 27*(2), 573–580. https://doi.org/10.1093/nar/27.2.573

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bruvo, R., Michiels, N. K., D'Souza, T. G., & Schulenburg, H. (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology, 13*(7), 2101–2106. https://doi.org/10.1111/j.1365-294X.2004.02209.x

Clark, L., & Jasieniuk, M. (2011). Polysat: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources, 11*(3), 562–566. https://doi.org/10.1111/j.1755-0998.2011.02985.x

de Assis Santos, D., de Carvalho Araújo, R. A., Kohler, L. M., Machado-Pinto, J., Hamdan, J. S., & Cisalpino, P. S. (2007). Molecular typing and antifungal susceptibility of *Trichophyton rubrum* isolates from patients with onychomycosis pre- and post-treatment. *International Journal of Antimicrobial Agents, 29*(5), 563–569. https://doi.org/10.1016/j.ijantimicag.2006.09.028

Edwards, S. V. (2009). Natural selection and phylogenetic analysis. *Proceedings of the National Academy of Sciences, 106*(22), 8799–8800. https://doi.org/10.1073/pnas.0904103106

Gallo, J. G., Woods, M., Graham, R. M., & Jennison, A. V. (2017). A severe transmissible Majocchi's granuloma in an immunocompetent returned traveler. *Medical Mycology Case Reports, 18*, 5–7. https://doi.org/10.1016/j.mmcr.2017.07.003

Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D. M., & Hood, M. E. (2008). Speciation in fungi. *Fungal Genetics and Biology, 45*(6), 791–802. https://doi.org/10.1016/j.fgb.2008.02.001

Gong, J., Wu, W., Ran, M., Wang, X., Liu, W., Wan, Z., … Li, R. (2016). Population differentiation and genetic diversity of *Trichophyton rubrum* as revealed by highly discriminatory microsatellites. *Fungal Genetics and Biology, 95*, 24–29. https://doi.org/10.1016/j.fgb.2016.08.002

Gräser, Y., Fröhlich, J., Presber, W., & de Hoog, S. (2007). Microsatellite markers reveal geographic population differentiation in *Trichophyton rubrum*. *Journal of Medical Microbiology, 56*(8), 1058–1065. https://doi.org/10.1099/jmm.0.47138-0

Gräser, Y., Kühnisch, J., & Presber, W. (1999). Molecular markers reveal exclusively clonal reproduction in *Trichophyton rubrum*. *Journal of Clinical Microbiology, 37*(11), 3713–3717.

Gultyaeva, E. I., Aristova, M. K., Shaidayuk, E. L., Mironenko, N. V., Kazartsev, I. A., Akhmetova, A. & Kosman, E. (2017). Genetic differentiation of *Puccinia triticina* Erikss. in Russia. *Russian Journal of Genetics, 53*(9), 998–1005. https://10.1134/S1022795417070031

Hiruma M, Kano R, Sugita T, Mochizuki T, Hasegawa A, Hiruma M. Urease gene of *Trichophyton rubrum* var. *raubitschekii*. J Dermatol. 2013; 40: 111–113. https://10.1111/1346-8138.12017

Hryncewicz-Gwóźdź, A., Jagielski, T., Sadakierska-Chudy, A., Dyląg, M., Pawlik, K., Baran, E., & Szepietowski, J. C. (2011). Molecular typing of *Trichophyton rubrum* clinical isolates from Poland. *Mycoses, 54*(6), e726–e736. https://doi.org/10.1111/j.1439-0507.2010.02007.x

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution, 23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Jackson, C. J., Barton, R. C., Kelly, S. L., & Evans, E. G. V. (2000). Strain identification of *Trichophyton rubrum* by specific amplification of subrepeat elements in the ribosomal DNA nontranscribed spacer. *Journal of Clinical Microbiology, 38*(12), 4527–4534.

Jacob, T. R., Peres, N. T., Persinoti, G. F., Silva, L. G., Mazucato, M., Rossi, A., & Martinez-Rossi, N. M. (2012). rpb2 is a reliable reference gene for quantitative gene expression analysis in the dermatophyte *Trichophyton rubrum*. *Medical Mycology, 50*(4), 368–377. https://doi.org/10.3109/13693786.2011.616230

Ito, H. C., & Dieckmann, U. (2007). A new mechanism for recurrent adaptive radiations. *The American Naturalist, 170*(4), E96–E111. https://doi.org/10.1086/521229

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ, 2*, e281. http://dx.doi.org/10.7717/peerj.281

Kano, R., Isizuka, M., Hiruma, M., Mochizuki, T., Kamata, H., & Hasegawa, A. (2013). Mating type gene (*MAT1-1*) in Japanese isolates of *Trichophyton rubrum*. *Mycopathologia, 175*(1-2), 171–173. https://doi.org/10.1007/s11046-012-9603-2

Kano, R., Kawasaki, M., Mochizuki, T., Hiruma, M., & Hasegawa, A. (2012). Mating genes of the *Trichophyton mentagrophytes* complex. *Mycopathologia, 173*(2-3), 103–112. https://doi.org/10.1007/s11046-011-9487-6

Kashyap, P. L., Kumar, S., Kumar, R. S., Tripathi, R., Sharma, P., Sharma, A., ... Singh, G. P. (2020). Identification of novel microsatellite markers to assess the population structure and genetic differentiation of *Ustilago hordei* causing covered smut of barley. *Frontiers in Microbiology, 10*, 2929. https://doi.org/10.3389/fmicb.2019.02929

Lackner, M., Obermair, J., Naschberger, V., Raschbichler, L. M., Kandelbauer, C., Pallua, J., ... Binder, U. (2019). Cryptic species of *Aspergillus* section *Terrei* display essential physiological features to cause infection and are similar in their virulence potential in *Galleria mellonella*. *Virulence, 10*(1), 542–554. https://doi.org/10.1080/21505594.2019.1614382

Latka, C., Dey, S. S., Mahajan, S., Prabu, R., Jangir, P. K., Gupta, C., … Taneja, B. (2015). Genome sequence of a clinical isolate of dermatophyte, *Trichophyton rubrum* from India. *FEMS Microbiology Letters, 362*(8), fnv039. https://doi.org/10.1093/femsle/fnv039

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J. M., Zuccarello, G. C. & De Clerck, O. (2014). DNA-based species delimitation in algae. *European Journal of Phycology, 49*(2), 179–196. http://dx.doi.org/10.1080/09670262.2014.904524

Li, H. C., Bouchara, J. P., Hsu, M. M., Barton, R., Su, S., & Chang, T. C. (2008). Identification of dermatophytes by sequence analysis of the rRNA gene internal transcribed

spacer regions. *Journal of Medical Microbiology, 57*(5), 592–600. https://doi.org/10.1099/jmm.0.47607-0

Li, W., Metin, B., White, T. C., & Heitman, J. (2010). Organization and evolutionary trajectory of the mating type (MAT) locus in dermatophyte and dimorphic fungal pathogens. *Eukaryotic Cell, 9*(1), 46–58. https://doi.org/10.1128/EC.00259-09

Martinez, D. A., Oliver, B. G., Gräser, Y., Goldberg, J. M., Li, W., Martinez-Rossi, N. M., … White, T. C. (2012). Comparative genome analysis of *Trichophyton rubrum* and related dermatophytes reveals candidate genes involved in infection. *MBio, 3*(5), e00259–12. https://doi.org/10.1128/mBio.00259–12

Martins, M. P., Rossi, A., Sanches, P. R., Bortolossi, J. C., & Martinez-Rossi, N. M. (2020). Comprehensive analysis of the dermatophyte *Trichophyton rubrum* transcriptional profile reveals dynamic metabolic modulation. *Biochemical Journal,* pii: BCJ20190868. https://doi.org/10.1042/BCJ20190868

Mironenko, N. V., Baranova, O. A., Kovalenko, N. M., Mikhailova, L. A., & Rosseva, L. P. (2016). Genetic structure of the Russian populations of *Pyrenophora tritici-repentis*, determined by using microsatellite markers. *Russian Journal of Genetics, 52*(8), 771–779. https://doi.org/10.1134/S1022795416080093

Nelson, D. L. & Cox, M. M. (2008). Lehninger principles of biochemistry. WH Freeman. New York.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., ... Stepanauskas, R. (2013). Assembling genomes and mini-metagenomes from highly chimeric reads. In M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), *Annual International Conference on Research in Computational Molecular Biology* (pp. 158-170). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-37195-0_13

Packeu, A., Stubbe, D., Roesems, S., Goens, K., Van Rooij, P., de Hoog, S., & Hendrickx, M. (2019). Lineages within the *Trichophyton rubrum* complex. *Mycopathologia.* https://doi.org/10.1007/s11046-019-00386-z

Pchelin, I. M., Azarov, D. V., Chilina, G. A., Dmitriev, K. A., Vasilyeva, N. V., & Taraskina, A. E. (2018a). Single-nucleotide polymorphism in a local population of *Trichophyton rubrum*. *Medical Mycology, 56*(1), 125–128. https://doi.org/10.1093/mmy/myx009

Pchelin, I. M., Azarov, D. V., Churina, M. A., Ryabinin, I. A., Vibornova, I. V., Apalko, S. V., … Vasilyeva, N. V. (2019a). Whole genome sequence of first *Candida auris* strain, isolated in Russia. *Medical Mycology*. https://doi.org/10.1093/mmy/myz078

Pchelin, I. M., Azarov, D. V., Churina, M. A., Scherbak, S. G., Apalko, S. V., Vasilyeva, N. V., & Taraskina, A. E. (2019b). Species boundaries in the *Trichophyton mentagrophytes* / *T. interdigitale* species complex. *Medical Mycology, 57*(6), 781–789. https://doi.org/10.1093/mmy/myy115

Pchelin, I. M., Kryuchkova, M. A., Bogdanova, T. V., Chilina, G. A., Boronina, L. G., Olina E. S., … Taraskina A. E. (2018b). We need more powerful microsatellite assay for population genetic studies of *Trichophyton rubrum*. *Medical Mycology, 56*(S2), S119.

Peakall, R. & Smouse P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6*(1), 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R. & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics, 28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Persinoti, G. F., Martinez, D. A., Li, W., Döğen, A., Billmyre, R. B., Averette, A., … Cuomo, C. A. (2018). Whole-genome analysis illustrates global clonal population structure of the ubiquitous dermatophyte pathogen *Trichophyton rubrum*. *Genetics, 208*(4), 1657–1669. https://doi.org/10.1534/genetics.117.300573

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One, 5*, e9490. https://doi.org/10.1371/journal.pone.0009490

Rad, M. M., Mohammadi, A. M. A., & Barton, R. C. (2008). PCR typing of *Trichophyton rubrum* isolates by specific amplification of subrepeat elements in ribosomal DNA nontranscribed spacer. *Iranian Journal of Dermatology, 11*(1), 17–20.

Ramaraj, V., Vijayaraman, R. S., Elavarashi, E., Rangarajan, S., & Kindo, A. J. (2017). Molecular strain typing of clinical isolates, *Trichophyton rubrum* using non transcribed spacer (NTS) region as a molecular marker. *Journal of Clinical and Diagnostic Research, 11*(5), DC04–DC09. https://doi.org/10.7860/jcdr/2017/21994.9843

Rozenfeld, A. F., Arnaud-Haond, S., Hernández-García, E., Eguíluz, V. M., Matías, M. A., Serrão, E., & Duarte, C. M. (2007). Spectrum of genetic diversity and networks of clonal organisms. *Journal of the Royal Society Interface, 4*(17), 1093–1102. https://doi.org/10.1098/rsif.2007.0230

Sexton, A. C., & Howlett, B. J. (2004). Microsatellite markers reveal genetic differentiation among populations of *Sclerotinia sclerotiorum* from Australian canola fields. *Current Genetics, 46*(6), 357–365. https://doi.org/10.1007/s00294-004-0543-3

Sigurgeirsson, B., & Baran, R. (2014). The prevalence of onychomycosis in the global population: a literature study. *Journal of the European Academy of Dermatology and Venereology, 28*(11), 1480–1491. https://doi.org/10.1111/jdv.12323

Singh, N., Anand, G., & Kapoor, R. (2019). Virulence and genetic diversity among *Fusarium oxysporum* f. sp. *carthami* isolates of India using multilocus RAPD and ISSR markers. *Tropical Plant Pathology, 44*(5), 409–422. https://doi.org/10.1007/s40858-019-00303-1

Su, H., Packeu, A., Ahmed, S. A., Al-Hatmi, A. M. S., Blechert, O., İlkit, M., … de Hoog, S. (2019). Species distinction in the *Trichophyton rubrum* complex. *Journal of Clinical Microbiology, 57*(9), e00352-19. https://doi.org/10.1128/jcm.00352-19

Summerbell, R. C. (2000). Form and function in the evolution of dermatophytes. In: Kushwaha RKS, Guarro J, editors. *Biology of dermatophytes and other keratinophilic fungi*. *Revista Iberoamericana de Micología*.17 Suppl: 30–43.

Summerbell, R. C., Haugland, R. A., Li, A., & Gupta, A. K. (1999). rRNA gene internal transcribed spacer 1 and 2 sequences of asexual, anthropophilic dermatophytes related to *Trichophyton rubrum. Journal of Clinical Microbiology, 37*(12), 4005–4011.

Suzuki, S., Mano, Y., Furuya, N., & Fujitani, K. (2018). Molecular epidemiological analysis of the spreading conditions of *Trichophyton* in long-term care facilities in Japan. *Japanese Journal of Infectious Diseases, 71*(6), 462–466. https://doi.org/10.7883/yoken.jjid.2018.090

Taghipour, S., Pchelin, I. M., Zarei Mahmoudabadi, A., Ansari, S., Katiraee, F., Rafiei, A., … Rezaei-Matehkolaei, A. (2019). *Trichophyton mentagrophytes* and *T. interdigitale* genotypes are associated with particular geographic areas and clinical manifestations. *Mycoses, 62*(11), 1084–1091. https://doi.org/10.1111/myc.12993

Takahashi, I., Fukushima, K., Miyaji, M., Nishimura, K., Asano, K., & Iizuka, H. (2014). Species identification and strain typing of dermatophytes by single-strand conformation polymorphism (SSCP) analysis of the ribosomal DNA and polymerase chain reaction analysis of subrepeat elements in the intergenic spacer region of *Trichophyton rubrum*. *Asahikawa Medical University Repository, 15*, 27–36.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution, 30*(12), 2725–2729. https://doi.org/10.1093/molbev/mst197

Tibayrenc, M. (2006). The species concept in parasites and other pathogens: a pragmatic approach? *Trends in Parasitology, 22*(2), 66–70. https://doi.org/10.1016/j.pt.2005.12.010

Treangen, T. J., Ondov, B. D., Koren, S., & Phillippy, A. M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biology, 15*(11), 524. https://doi.org/10.1186/s13059-014-0524-x

White, T. J., Bruns, T., Lee, S., & Taylor, J. W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. Innis, D. Gelfand, J. Sninsky, & T. White (Eds.), *PCR protocols: A guide to methods and applications* (pp. 315–322). Orlando, FL: Academic Press.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics, 13*, 134. https://doi.org/10.1186/1471-2105-13-134

Zhan, P., Dukik, K., Li, D., Sun, J., Stielow, J. B., Gerrits van den Ende, B., … de Hoog, G. S. (2018). Phylogeny of dermatophytes with genomic character evaluation of clinically distinct *Trichophyton rubrum* and *T. violaceum*. *Studies in Mycology, 89*, 153–175. https://doi.org/10.1016/j.simyco.2018.02.004

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, *7*(1-2): 203–214. https://doi.org/10.1089/10665270050081478

Zheng, H., Blechert, O., Mei, H., Ge, L., Liu, J., Tao, Y., … Liu, W. (2019). Whole-genome resequencing of *Trichophyton rubrum* provides insights into population differentiation and drug resistance. *Mycopathologia*. https://doi.org/10.1007/s11046-019-00384-1

Supplementary data

Figures S1 and S2.

Supplement S1. Molecular strain typing in 70 *T. rubrum* isolates.

Figures

Figure 1. A phylogenetic network, calculated from the matrix of mutation probability-based genetic distances (Bruvo et al. 2004), 12 microsatellite loci. Urease-positive isolates are marked by asterisks.

Figure 2. A phylogenetic network, calculated from the matrix of distances, based on the lengths of microsatellite loci (Rozenfeld et al. 2007), 12 microsatellite loci. Urease-positive isolates are marked by asterisks.

Figure 3. A phylogenetic network, calculated from the matrix of mutation probability-based genetic distances (Bruvo et al. 2004), 11 microsatellite loci. Urease-positive isolates are marked by asterisks.

Figure 4. A phylogenetic network, based on Rozenfeld genetic distances, 11 microsatellite loci. Urease-positive isolates are marked by asterisks.

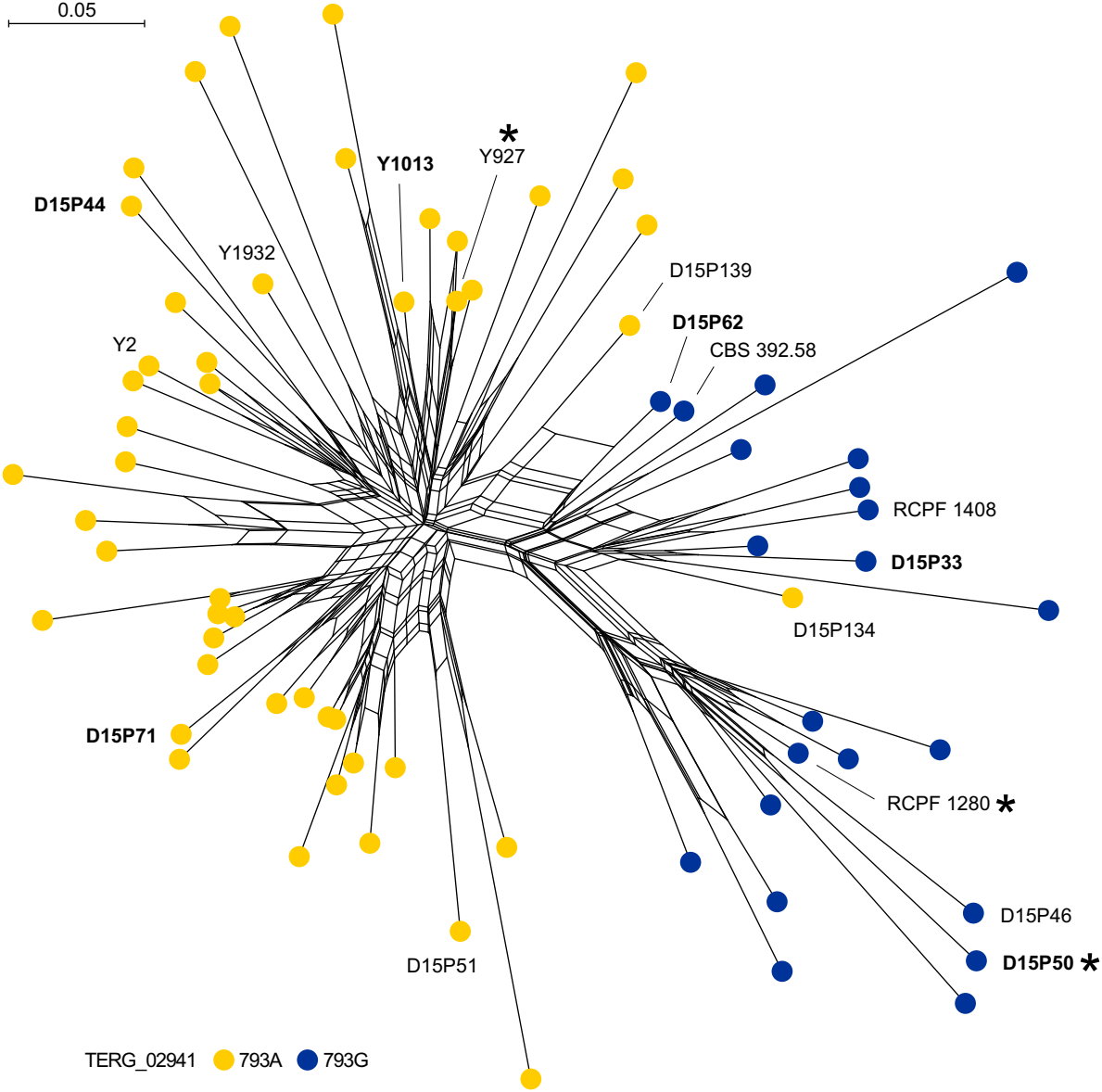
Figure 5. Phylogenomic tree of *T. rubrum* isolates. ST1 and ST2 lineages are defined by non-synonymous substitutions in TERG_02941 and TERG_03298 protein-coding loci. Russian population shares both clades with West European and Asian countries. Single and double arrowheads mark Chinese isolates, fallen into two different clusters in the original publication by Zheng et al. (2019). Asterisks mark isolates with urease activity. Branch support values ≥ 75% are marked by open circles and values ≥ 95% are marked by solid circles. Newly sequenced isolates are given in bold. The tree is rooted by *T. soudanense* CBS 452.61 whole genome sequence (the branch isn't shown).

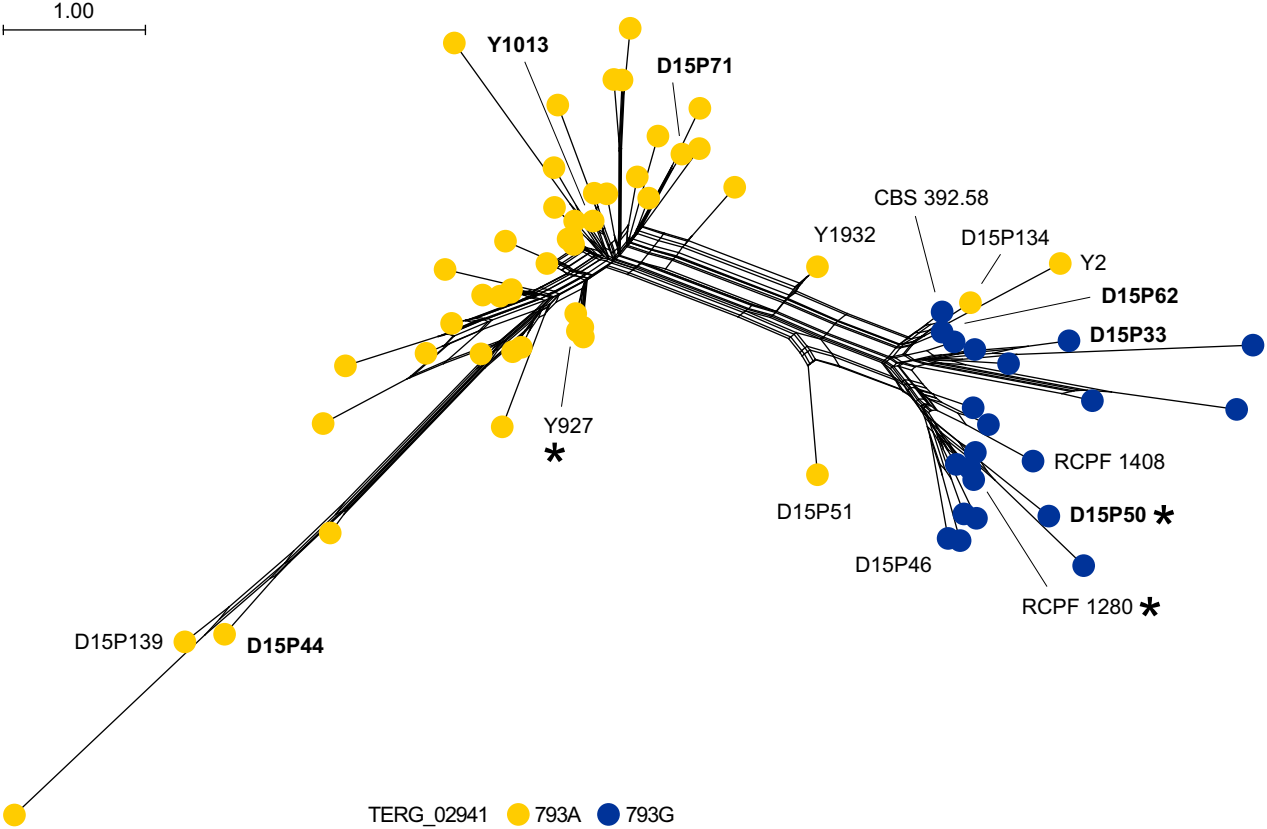
Figure 6. *Trichophyton* spp. A tree, based on concatenated sequences of TERG_02718, TERG_02913, TERG_02941, TERG_03298, TERG_03581, TERG_03748 and TERG_03778 loci. In the *T. mentagrophytes* / *T. interdigitale* species group, ITS region genotypes are indicated (Taghipour et al. 2019). Asterisks mark isolates with urease activity. Branch support values ≥ 75% are marked by open circles and values ≥ 95% are marked by solid circles. Two clades are shown separately at different scales.

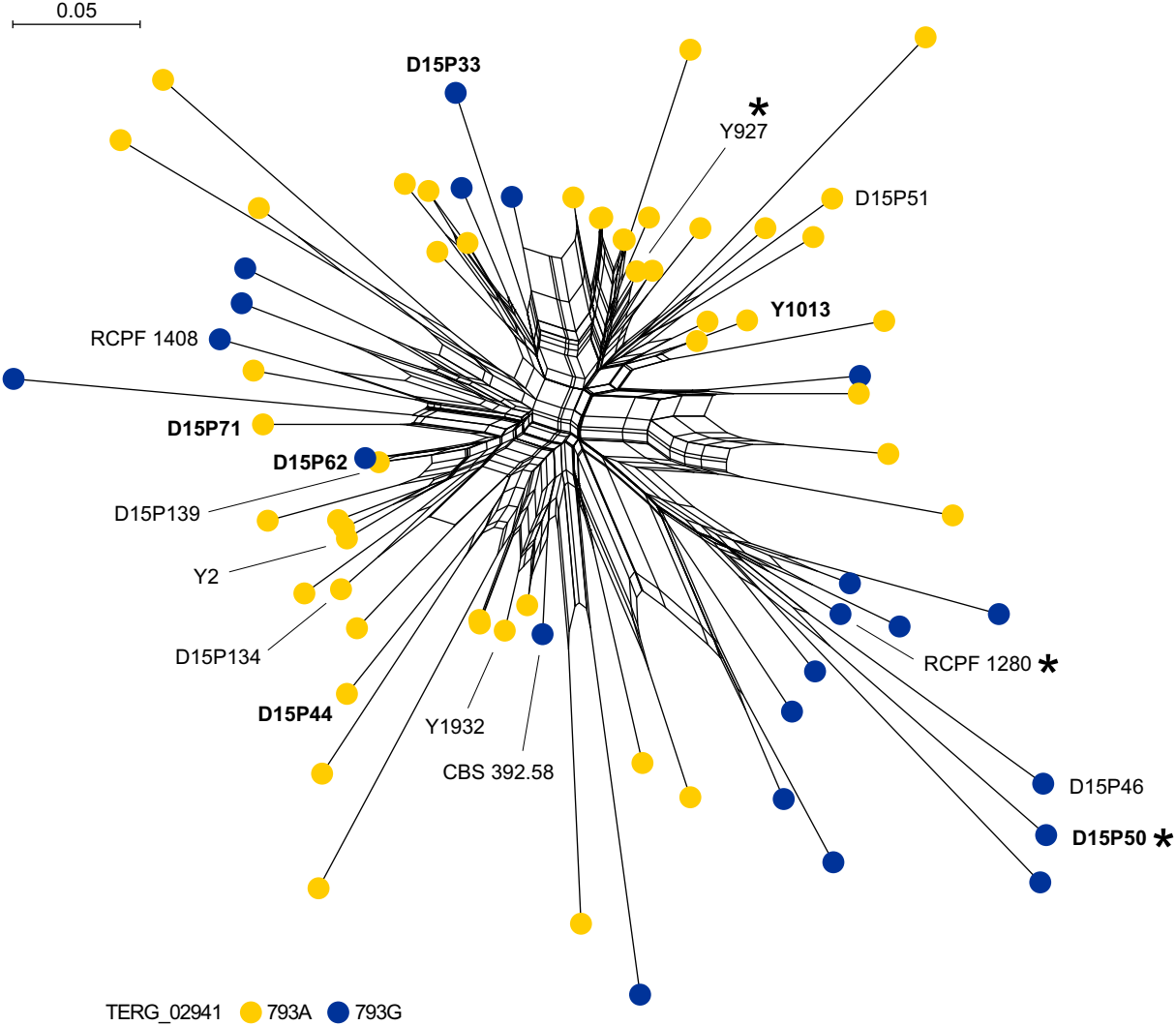
Figure 7. Hypothetical distribution of *Trichophyton rubrum* genetic lines. A total of 443 isolates, from 17 to 70 isolates for a country. The abbreviations above pie charts indicate the sources of variability: SNP, single nucleotide polymorphism of TERG_02941 and TERG_03298 genes; TRS-1, the length of TRS-1 locus of ribosomal non-transcribed spacer; WGS, whole genome sequence.

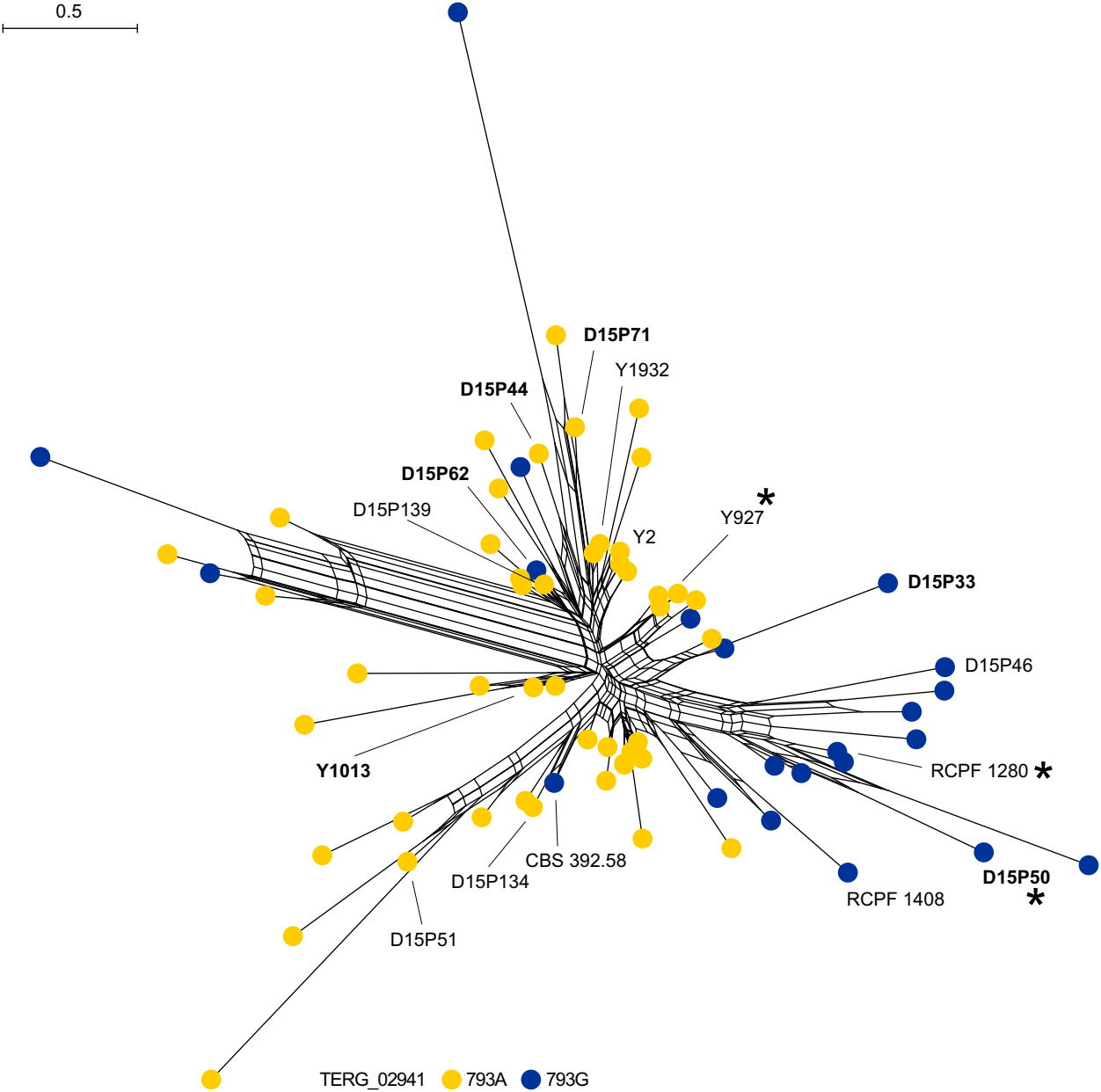
Figure S1. Nucleotide polymorphism of *Trichophyton rubrum* ITS region sequences. *A*, 247insC mutation in D15P139 isolate (asterisk). *B*, an additional TA repeat near the 5'-end of the sequence in RCPF 1280 strain, 595insTA (double asterisk).

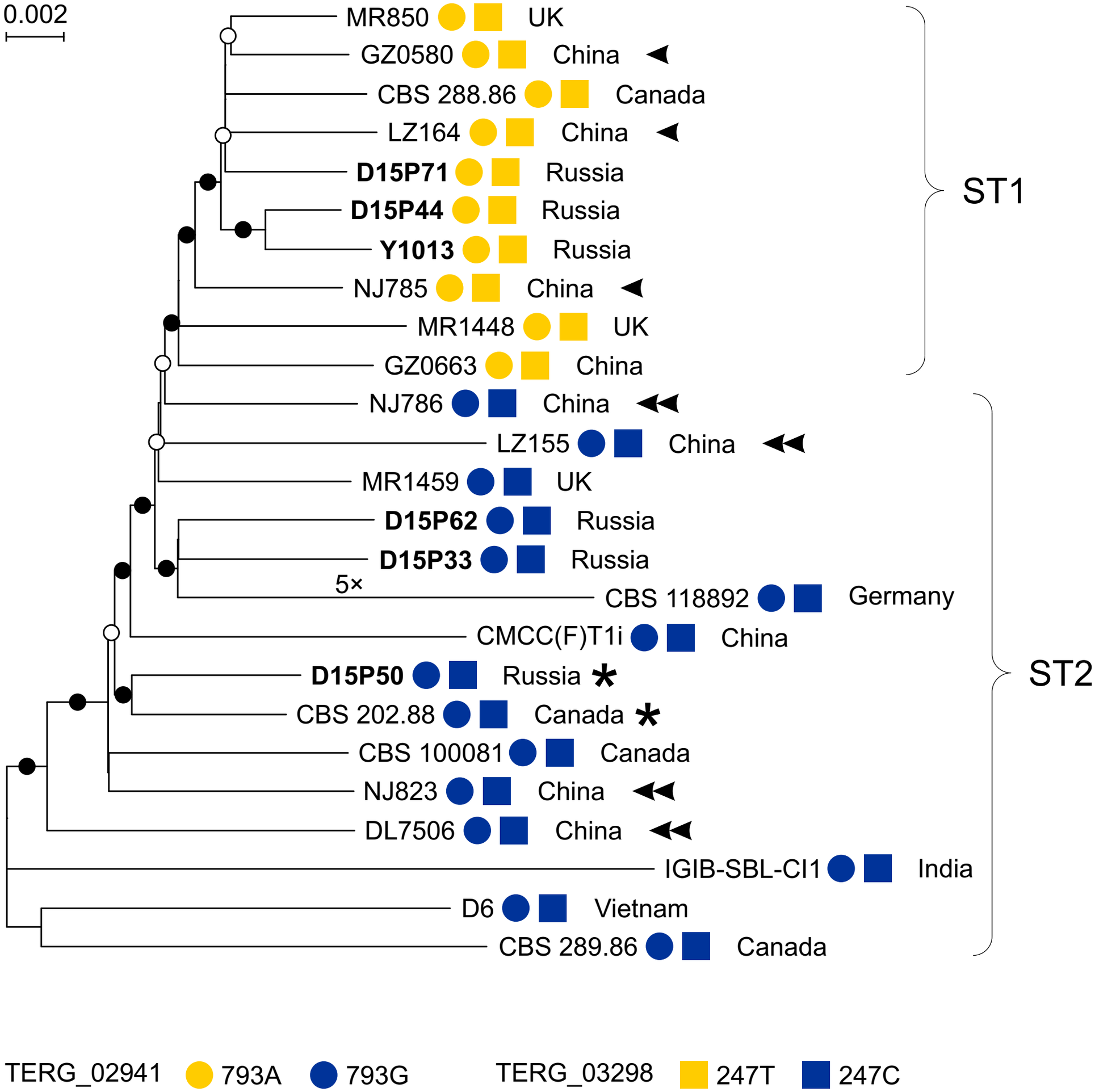
Figure S2. STR*Tr*9 microsatellite locus length and TERG_6744 gene expression in *T. rubrum* genetic lineages. *A*, Average length of STR*Tr*9 amplification products differs between ST1 and ST2 isolates. *B*, Relative expression of TERG_6744 gene, coding for fructose 1,6-bisphosphate aldolase, varies between strains irrespectively of their genetic lineage affiliation. From the left to the right: D15P44 and D15P71 (ST1), D15P46 and D15P50 ("ancestral" ST2), D15P33 and D15P62 ("derived" ST2). There is no immediate consequence of STR*Tr*9 locus length for TERG_06744 transcript level.

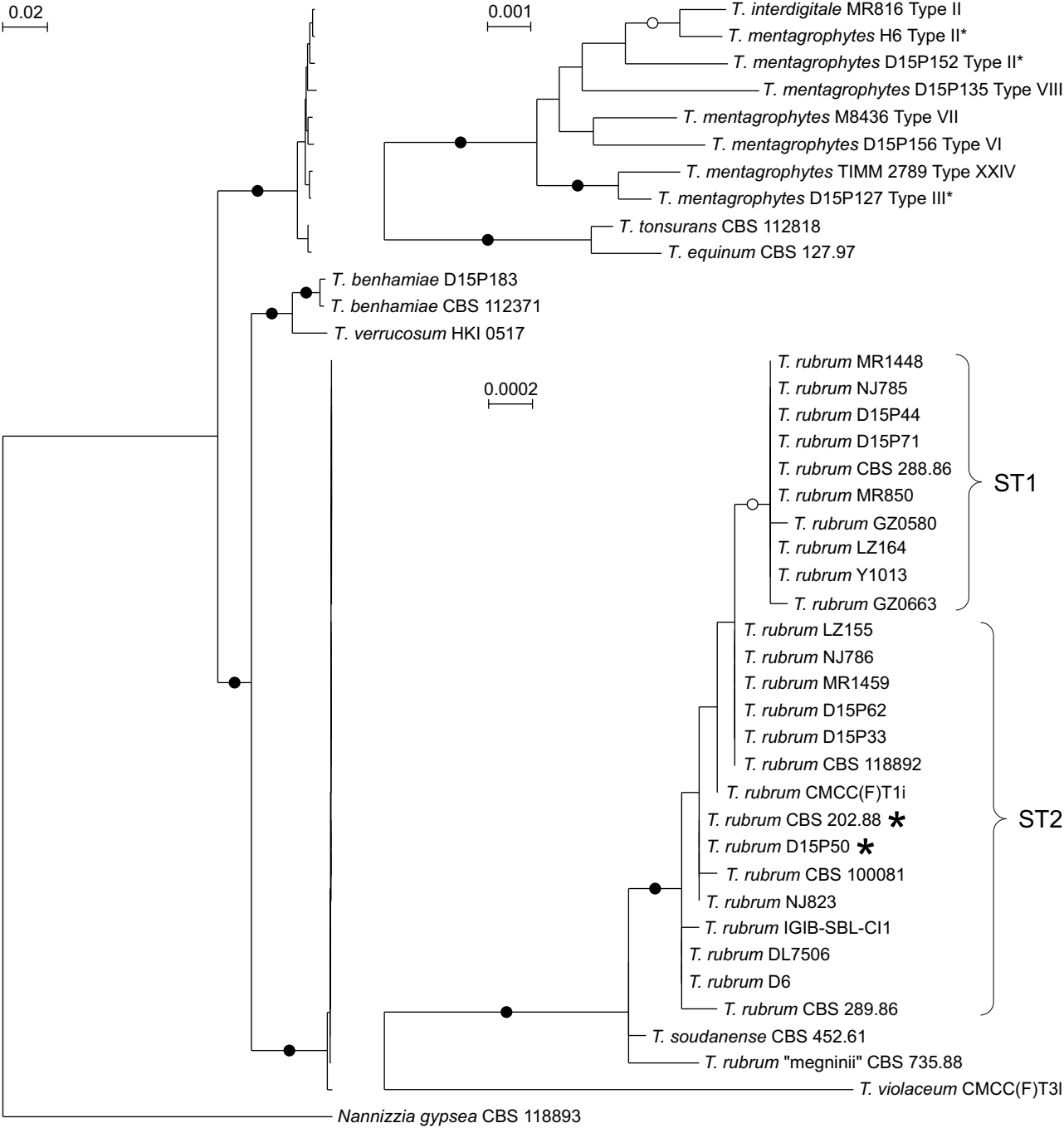

0.05
Y927
Y1013
D15P44
Y1932
D15P139
D15P62
CBS 392.58
Y2
RCPF 1408
D15P33
D15P134
D15P71
RCPF 1280
D15P46
D15P50
D15P51
TERG_02941
793A
793G

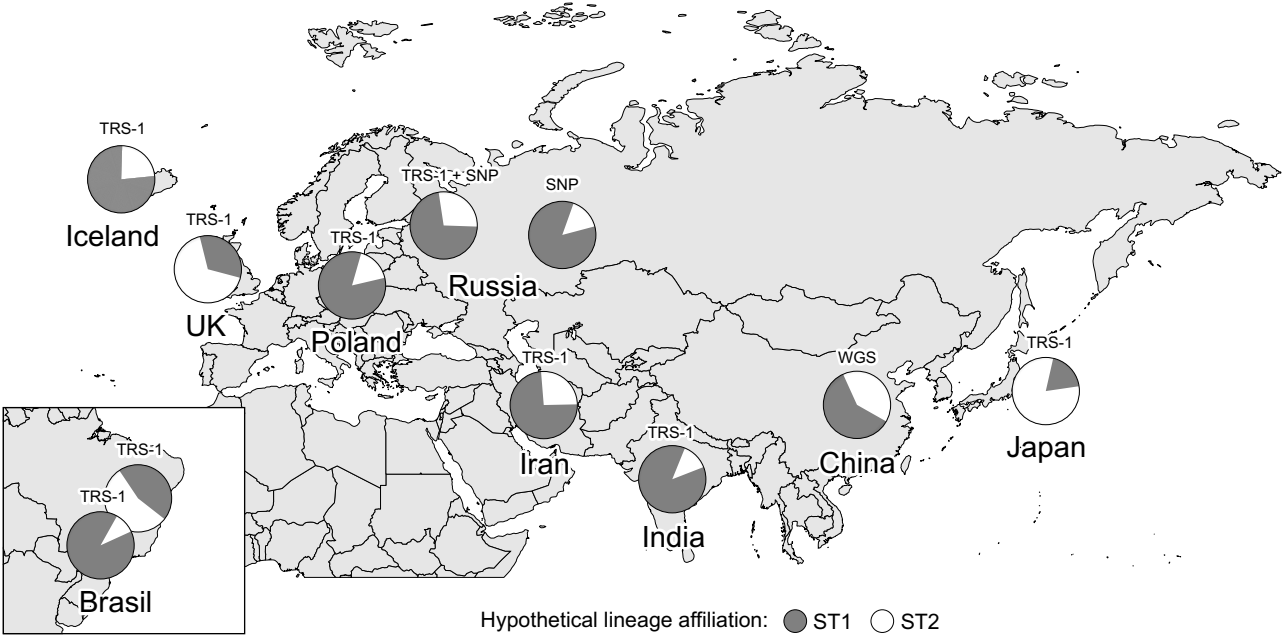

1.00
Y1013
D15P71
CBS 392.58
Y1932
D15P134
Y2
D15P62
D15P33
Y927
*
RCPF 1408
D15P51
D15P50 *
D15P46
RCPF 1280 *
D15P139
D15P44
TERG_02941
793A
793G

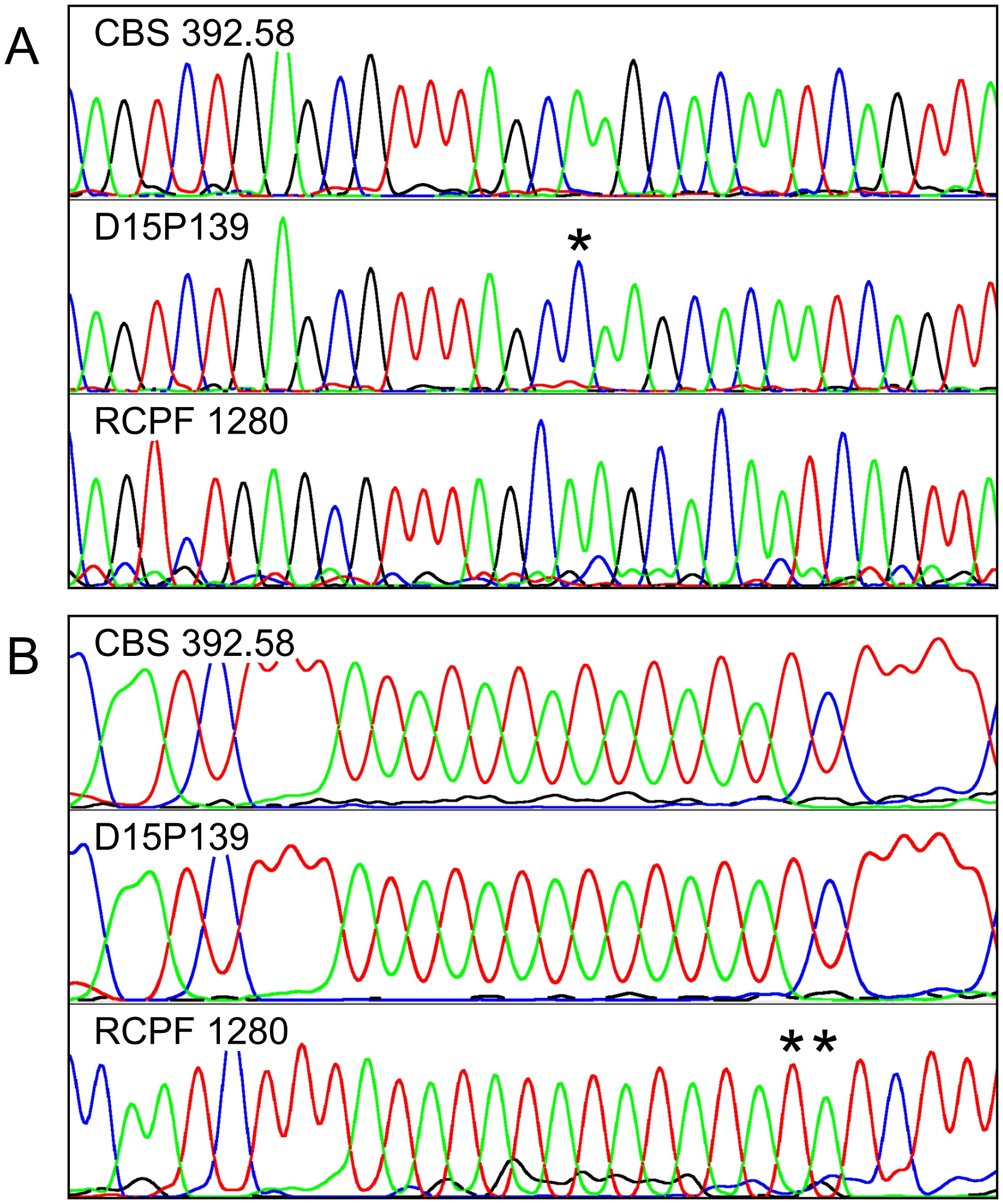

0.05
D15P33
Y927 *
D15P51
Y1013
RCPF 1408
D15P71
D15P62
D15P139
Y2
D15P134
D15P44
Y1932
CBS 392.58
RCPF 1280 *
D15P46
D15P50 *
TERG_02941
793A
793G

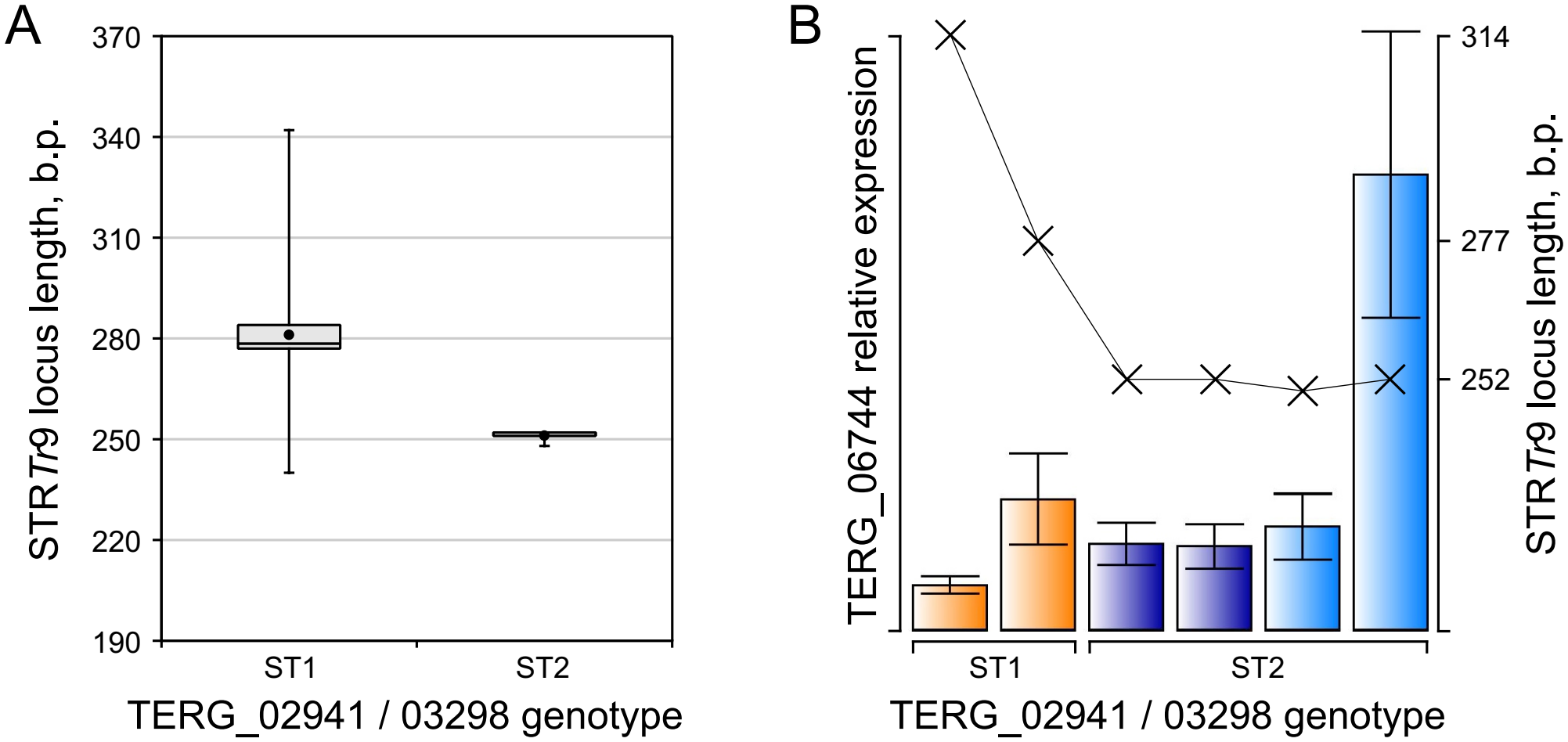

0.5
D15P71
Y1932
D15P44
D15P62
D15P139
Y927
*
Y2
D15P33
D15P46
Y1013
RCPF 1280 *
CBS 392.58
D15P134
D15P50
*
RCPF 1408
D15P51
TERG_02941
793A
793G


0.002
MR850 UK
GZ0580 China
CBS 288.86 Canada
LZ164 China
D15P71 Russia
D15P44 Russia
Y1013 Russia
NJ785 China
MR1448 UK
GZ0663 China
ST1
NJ786 China
LZ155 China
MR1459 UK
D15P62 Russia
D15P33 Russia
5×
CBS 118892 Germany
CMCC(F)T1i China
D15P50 Russia *
CBS 202.88 Canada *
CBS 100081 Canada
NJ823 China
DL7506 China
IGIB-SBL-CI1 India
D6 Vietnam
CBS 289.86 Canada
ST2
TERG_02941 793A 793G
TERG_03298 247T 247C


0.02
0.001
T. interdigitale MR816 Type II
T. mentagrophytes H6 Type II*
T. mentagrophytes D15P152 Type II*
T. mentagrophytes D15P135 Type VIII
T. mentagrophytes M8436 Type VII
T. mentagrophytes D15P156 Type VI
T. mentagrophytes TIMM 2789 Type XXIV
T. mentagrophytes D15P127 Type III*
T. tonsurans CBS 112818
T. equinum CBS 127.97
T. benhamiae D15P183
T. benhamiae CBS 112371
T. verrucosum HKI 0517
0.0002
T. rubrum MR1448
T. rubrum NJ785
T. rubrum D15P44
T. rubrum D15P71
T. rubrum CBS 288.86
T. rubrum MR850
T. rubrum GZ0580
T. rubrum LZ164
T. rubrum Y1013
T. rubrum GZ0663
ST1
T. rubrum LZ155
T. rubrum NJ786
T. rubrum MR1459
T. rubrum D15P62
T. rubrum D15P33
T. rubrum CBS 118892
T. rubrum CMCC(F)T1i
T. rubrum CBS 202.88 *
T. rubrum D15P50 *
T. rubrum CBS 100081
T. rubrum NJ823
T. rubrum IGIB-SBL-CI1
T. rubrum DL7506
T. rubrum D6
T. rubrum CBS 289.86
ST2
T. soudanense CBS 452.61
T. rubrum "megninii" CBS 735.88
T. violaceum CMCC(F)T3I
Nannizzia gypsea CBS 118893


TRS-1
Iceland
TRS-1
UK
TRS-1
Poland
TRS-1 + SNP
SNP
Russia
TRS-1
Iran
TRS-1
India
WGS
China
TRS-1
Japan
TRS-1
TRS-1
Brasil
Hypothetical lineage affiliation: ST1 ST2


A
CBS 392.58
D15P139
*
RCPF 1280
B
CBS 392.58
D15P139
RCPF 1280
**


A
370
340
310
280
250
220
190
STR *Tr9* locus length, b.p.
ST1
ST2
TERG_02941 / 03298 genotype
B
TERG_06744 relative expression
314
277
252
STR *Tr9* locus length, b.p.
ST1
ST2
TERG_02941 / 03298 genotype