Molecular characterization of 'Bhut Jolokia' the hottest chilli

J PURKAYASTHA[1,*], SI ALAM[2], HK GOGOI[1], L SINGH[3] and V VEER[1]

[1]*Defence Research Laboratory, Tezpur 784001, Assam, India*

[2]*Biotechnology Division, Defence Research & Development Establishment, Gwalior 474 002, India*

[3]*Directorate of Life Sciences, Defence Research and Development Organization, New Delhi 110 105, India*

**Corresponding author (Fax, +91-3712-258534; Email, jubilee_p@rediffmail.com)*

The northeast region of India, considered as 'hot spot' of biodiversity, having unique ecological environment with hot and high-humidity conditions, has given rise to the world's hottest chilli, 'Bhut Jolokia', which is at least two times hotter than Red Savina Habanero in terms of Scoville heat units (SHU). This study was undertaken to determine the distinctiveness of 'Bhut Jolokia' from *Capsicum frutescens* or *Capsicum chinense* through sequencing of the ribosomal RNA (rRNA) gene-internal transcribed (ITS) region along with its phylogenetic analysis. Although a compensatory base change (CBC) in the ITS2 region was not observed between the closely related species of *C. frutescens* and *C. chinense* when compared with *Bhut Jolokia*; phylogenetic analysis using ITS1, 5.8S and ITS2 sequences indicated a distinct clade for all the accessions of 'Bhut Joloikia', while *C. frutescens* and *C. chinense* occupied discrete lineages. Further, a unique 13-base deletion was observed in all the representative accessions of 'Bhut Jolokia', making it distinct from all other members within the genus and beyond. The degree of genetic variations along with its extreme pungency might be related to ambient environmental factors of northeastern India.

[Purkayastha J, Alam SI, Gogoi HK, Singh L and Veer V 2012 Molecular characterization of 'Bhut Jolokia' the hottest chili. *J. Biosci.* **37** 757–768]
DOI 10.1007/s12038-012-9249-8

1. Introduction

The genus *Capsicum* is believed to be one of the earliest domesticated plant genera and has been dated to around 7000 years back based on archaeological data (Basu and De 2003). Chilli peppers are popular spices in many parts of the world, valued for their sensory attributes of colour, pungency and flavour (Pino *et al.* 2007). Peppers are economically important because of the vast consumption of their diverse varieties. The food industry is the largest user of capsicums, where they are used as colouring and flavouring agents in sauces, soups, processed meats, snacks, candies, soft drinks and alcoholic beverages (Pino *et al.* 2007). Pungency, the unique characteristics of this genus, is due to the biosynthesis of capsaicinoids and has driven the rapid domestication of several *Capsicum* spp. in different parts of the world (Andrews 1984; Walsh and Hoot 2001).

Since the time of Darwin (1845), geographical and climatic variations have served as living laboratories for speciation. Chilli peppers are tropical plants and thus are ideally suited to hot and humid conditions. However, they are very adaptable and do well even in semi-arid regions. They love nitrogen, hot and sunny weather and well-draining soil. Their degree of pungency varies and depends not only on genetic makeup but also on weather and soil conditions. Apparently, environmental stress factors increase pungency levels (Lindsay and Bosland 1995). Chilli plants may become more, or less, pungent under different environmental stress. Two types of stresses that affect the pungency of peppers are high average temperature outside and insufficient watering. Another form of stress is overwatering or

Keywords. Bhut Jolokia; environmental factors; ITS region; northeastern India; plant phylogeny; rDNA

Supplementary materials pertaining to this article are available on the *Journal of Biosciences* Website at *http://www.ias.ac.in/jbiosci/sep2012/supp/Purkayastha.pdf*

Published online: 14 August 2012

water clogging. It does not matter what type of stress the plant goes through; any type of environmental stress will increase or decrease the pungency of the fruit. *Capsicum* growth is very much dependent on temperature. Particularly, the root temperature must be sufficiently high (20–22°C). The key factor affecting fruit setting in *Capsicum* is night temperature, which ideally should be between 65–80°F.

The northeast region of India, considered a 'hot spots' of biodiversity, having unique ecological environment with high-humidity conditions acting as the hub of speciation, has given rise to the world's hottest chilli, 'Bhut Jolokia' (*Guinness Book of World Records 2006*) displacing the Red Savina. The name 'Bhut Jolokia' is locally popular probably due to its ghostly bite ('Bhut' means 'ghost' and 'Jolokia' means 'chilli pepper') or introduction by the Bhutias from Bhutan as poison chilli. The other names mentioned include 'Bih Jolokia', 'Borbih Jolokia', 'Nagahari', 'Nagajolokia', 'Naga Morich', 'Naga Moresh', 'Raja Mirchi' and 'Dorset Naga' (its derivative from Bangladesh). Regardless of the nomenclature, they all refer to the same plant. 'Bhut Jolokia' is being grown and consumed in different states of Northeast India including Assam, Nagaland, Manipur and Mizoram since ancient times. It is used as a spice in food or eaten raw along with the staple food. One seed from a 'Bhut Jolokia' can sustain intense pain sensations in the mouth for up to 30 min before subsiding. Nevertheless, because of its refreshing aroma, palatability and medicinal properties, people have been using it for pickle preparation, flavouring curries and for home remedies of ailments like gastritis, arthritis and chronic indigestion problems. It is also used as a remedy to summer heat, presumably by inducing perspiration. It was demonstrated that capsaicin binds to a protein TRPV1, which activates sensory neurons as a transducer of painful thermal stimuli and the compound was linked to thermogenesis (Caterina *et al.* 1997). In northeastern India the peppers are smeared on fences or used in smoke bombs as a safety precaution to keep wild elephants away.

Capsaicin, the major ingredient of pungent *Capsicum* fruit, shows numerous bioactive properties in humans. For instance, capsaicin is used to relieve the pain of peripheral neuropathy (Backonja *et al.* 2008), and can inhibit a variety of cancer cells (Mori *et al.* 2006; Baek *et al.* 2008). It has also been shown to possess anti-inflammatory and antioxidant activities (Surh 2002; Baek *et al.* 2008). Further, capsaicin inhibits obesity by decreasing energy intake (Reinbach *et al.* 2009), adipose tissue weight and serum triglyceride through stimulation of lipid mobilization (Kawada *et al.* 1986). In addition, capsaicin prevents adipogenesis and obesity by activation of TRPV1 channels (Zhang *et al.* 2007).

Since its discovery, disagreement has prevailed on the systematic position of 'Bhut Jolokia' and its heat content. Mathur *et al.* (2000) reported the 'Bhut Jolokia' to be a variety of *C. frutescens* L. and to have a very high heat level (855,000 SHUs). An independent analysis, at the Chile Pepper Institute, New Mexico University, in 2006, revealed a rating of 1,001,304 SHU for 'Bhut Jolokia'; almost double that of Red Savina (577,000 SHUs), previously considered to be the hottest chilli in the world. However, 'Bhut Jolokia' was recorded in *Guinness Book of World Records* as belonging to the species of *C. chinense*.

Based on their RAPD analysis, Bosland and Baral (2007) concluded 'Bhut Jolokia' as having a taxonomic position between *C. frutescens* and *C. chinense* and as a natural hybrid of the two. The RAPD technique, however, suffers from several limitations and is not suitable for inter-population or inter-species comparisons and often results in incorrect genotype and species interpretations. Non-reproducibility of results, competition between different DNA fragments for amplification (Williams *et al.* 1993; Halldén *et al.* 1996) and identical mobility of fragments with different sequences (Xu 2006) are confounding issues for applying the technique for species delineation.

The nuclear ribosomal RNA (rRNA) gene complex is a tandem repeat unit of one to several thousand copies. This complex has several domains that evolve at varying rates (Jorgenson and Cluster 1988), and thus have different phylogenetic utilities. The 18S and 28S rRNA genes evolve relatively slowly and are used for phylogenetic analysis of broad range of organisms (Maidak *et al.*1997). The internal transcribed spacers (ITS) are known to evolve relatively quickly and are used for determining inter-specific (Jorgenson and Cluster 1988) and sometimes intra-specific relationships (Baura *et al.* 1992). For instance, compensatory base changes (CBCs) in the internal transcribed spacer 2 region (ITS2) of the nuclear rRNA cistron have been suggested as molecular classifier to indicate that two organisms belong to different species (Muller *et al.* 2007). CBCs occur in a paired region of a primary RNA transcript when both nucleotides of a paired site mutate, while the pairing itself is maintained (e.g. G-C mutates to A-U). Although the CBC is now accepted as criterion to distinguish species (Coleman 2007), a lack of CBCs in ITS2 secondary structures is not an indicator of two organisms belonging to the same species. Within the ITS is a coding 5.8S rRNA gene that evolves relatively slowly, but has been largely ignored in the context of molecular systematics of taxa. It has been shown, however, that the 5.8S rRNA gene contains considerable phylogenetic information particularly with respect to deep basal branches (Hershkovitz and Lewis 1996; Cullings and Vogler 1998).

Thus, the main objective of the study was to determine the relationship between 'Bhut Jolokia' and other related species of the genus *Capsicum* on the basis of genetic characterization results. The localities and the prevailing environmental conditions of the different 'Bhut Jolokia' accessions used in this study are highlighted here. Also, the variations in hotness of 'Bhut Jolokia' in different climatic conditions of India and the conducive climate of Northeast India are cited here. To the

best of our knowledge, this is the first detailed characterization of 'Bhut Jolokia', the world's naturally occurring hottest chilli.

2. Materials and methods

2.1 *Plant material*

Plant specimens including fruits and flowers of 'Bhut Jolokia' were collected from different parts of Northeast India during the month of May 2010. The mean environmental conditions of the study locations are shown in table 1. At least 10 plant specimens for each of the six accessions of Bhut Jolokia and one accession, each of *C. frutescens* and *C. chinense* (table 2), were evaluated for their genetic characteristics. The accessions were collected locally from different locations of Northeast India, differing in local climatic conditions in terms of altitude, average humidity and rainfall. The cultivars have been maintained since ancient times by the native population. Only gene sequences from the database were used for the additional plant species reflected in table 2.

2.2 *DNA Extraction*

Fresh fruits and shoot tips were used for isolation of total cellular DNA from different accessions of 'Bhut Jolokia' and also from *C. frutescens* and *C. chinense* with the use of the DNA-Easy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. Quality of DNA preparations were checked on 1% agarose gel using lambda DNA as marker (Fermentas). Nucleic acid concentrations were estimated by determining optical density at 260 nm using a UV-visible spectrophotometer (Specord200, USA).

2.3 *DNA amplification*

ITS1–5.8S rDNA–ITS2 region of *C. frutescens*, *C. chinense* and six accessions of 'Bhut Jolokia' was amplified using previously described universal primers ITS4 and ITS5 or ITS4 and ITSleu1, which target conserved regions in the 18S and 26S rRNA genes (White *et al.* 1990; Bohs and Olmstead 2001). The ITS1 and ITS2 regions were also independently amplified from the total DNA of plant tissues using primer pairs ITS5/ITS1R and ITS2F/ITS4 as illustrated in supplementary figure 1. The reaction mixture contained 1× PCR buffer (Tris-HCl, 100 mM; KCl, 500 mM; pH 8.3), 0.2 mM dNTPs, 2.0 µM of each primer, 1.5 mM $MgCl_2$, 1U of Taq Polymerase (Qiagen, Germany) and 50 ng of plant DNA in a total volume of 25 µL. Amplifications were carried out in a thermal cycler (T-professional Gradient Cycler, Biometra, Germany) according to the following amplification programme: an initial denaturation step of 95°C for 5 min followed by 30 cycles including denaturation for 1 min at 95°C, annealing for 1 min at 55°C and extension for 1 min at 72°C. Amplification was terminated by a final extension step of 10 min at 72°C.

2.4 *Cloning and sequencing*

The amplified DNA of approximately 750 bp corresponding to 18S partial, ITS1, 5.8S, ITS2 and 26S partial sequence was purified using a commercial PCR purification kit (Qiagen, Germany) following the manufacturer's instructions. Amplicons were sequenced directly in both sense and antisense directions using primers used for PCR amplification. The sequencing of ITS region for each accession was carried out twice, one with amplicons generated from DNA preparation from fruit and the other with those obtained from the shoot tips. A consensus sequence was created from the duplicates.

All the aforementioned amplified products (from *C. frutescens*, *C. chinense* and six accessions of 'Bhut Jolokia') were cloned in pTZ57R/T vector using InStaClone kit (Fermentas, USA) as per the manufacturer's instructions. The PCR products were ligated with 50 ng of vector DNA at 1:4 molar ratio of vector:insert. The ligated products were transformed into *E. coli* JM109 cells by heat-shock and the transformed clones were identified through blue/white colour selection. Plasmids were isolated from 10 randomly selected clones from each group and tested for the presence of insert by size determination on agarose gel (1.5%) and PCR

Table 1. Average environmental conditions in different locations of plant accessions studied

Location	Rainfall (mm)	Mean temperature (°C)		Relative humidity (%)		Sunshine hours
		Max	Min	08:30 h	17:30 h	
Tezpur/Sonitpur	170.61±51.24	29.74±0.70	19.95±1.5	83.91±1.37	77.41±2.2	149.5±10.47
Dibrugarh	260.09±65.13	28.44±0.74	19.40±1.59	82±1.91	77.58±2.13	142.10±16.45
Imphal	133.35±30.10	28.21±0.66	16.04±1.94	80.50±2.0	70.66±3.31	146.07±15.50
Kohima	156.78±52.46	22.50±0.95	12.60±1.35	79.34±1.23	70.33±2.31	145.37±12.41
Jorhat	188.88±47.27	28.42±0.84	19.57±1.73	88.66±0.76	80±2.05	133.47±12.30
Karbianglong	84.65±21.59	25.5±0.89	14.5±1.53	94.32±0.78	77.54±2.33	137.34±11.45

Table 2. Plant species and source of ITS sequences included in this study

Species/Accession no.	Location/Country	Latitude/Longitude/Altitude	Soil Type	Accession no.	Source/ Reference
Bhut Jolokia accn.1	Dibrugarh, Assam, India	27°28′N /94°55′E/108 m	Lateritic	HQ705983	This study
Bhut Jolokia accn.2	Jorhat, Assam, India	$26^0$44′N/$94^0$10′E/91 m	Young and Old Alluvial	HQ705984	This study
Bhut Jolokia accn.3	Sonitpur, Assam, India	26°30′N-27°01′N/92°16′E-93°43'E/21 m	Old Alluvial	HQ705985	This study
Bhut Jolokia accn.4	Karbianglong, Assam, India	25.33′-26.35′N/ 92.10′-93.50′E/300-1600 m	Red Loamy	HQ705986	This study
Bhut Jolokia accn.M	Ukhrul, Manipur, India	24°-25.41°N/94°-94.47°E/ 913-1763 m	Alluvium, Lateritic Black Regur and Red Ferruginous	HQ705987	This study
Bhut Jolokia accn.N	Kohima, Nagaland, India	25°40′N/94°08′E/1500 m	Inceptisols	HQ705988	This study
Capsicum frutescens	Tezpur, Assam, India	26°38′N/92°48′E/157 m	Old Alluvial	HQ705989	This study
Capsicum chinense	Sonitpur, Assam, India	26°30′N-27° 01′N/92°16′E-93°43′E/105 m	Old Alluvial	HQ705990	This study
Capsicum eximium	Mexico, North America			AY665841	Whitson and Manos 2005
Capsicum annum	Badajoz, Spain			GU944973	Hernández *et al.* 2010
Capsicum baccatum	Utah, USA			AF244708	Bohs and Olmstead 2001
Capsicum pubescens	Madison, USA			AY875749	Spooner *et al.* 2005 (UP)*
Capsicum lycianthoides	Madison, USA			DQ314158	Smith and Baum 2006
Solanum nigrum	Beijing, China			FJ980391	Chen and Han 2009 (UP)*

* UP=Unpublished submission

amplification of target gene. Two clones positive for insert were sequenced using the commercial services employing an automated sequencer (ABI PRISM, Model 3730, USA). Vector-specific forward and reverse sequencing primers were used for sequencing, and the deduced sequences (~700 bases) were subjected to BLAST search for closest match in the database (*http://www.ncbi.nlm.nih.gov*). DNA sequences have been deposited to NCBI database under the accession numbers HQ705983 to HQ705990 (table 2).

2.5 *Phylogenetic analysis*

The related sequences were retrieved and aligned by ClustalW alignment program. Sequences were truncated to the length of shortest sequence in the alignment data before constructing the tree. A maximum parsimony tree was reconstructed showing phylogenetic relationship between Bhut Jolokia and other related reference species of *Capsicum* using alignment data of concatenated ITS1–5.8S-ITS2 sequences (668 bases) with 1000 bootstrap replicates using tree construction software MEGA version 3.1 (Kumar *et al.* 2004).

2.6 *Secondary structure prediction*

The secondary structures of ITS2 rDNA from 'Bhut Jolokia' accessions were obtained by online secondary structure determination tool at ITS2 database v1.0 (Wolf *et al.* 2005; Schultz *et al.* 2005, 2006; Koetschan *et al.* 2010). The predicted structure of the closest match (*C. eximium*) was used for structural comparison (Eddy 1998; Keller *et al.* 2009) and a representative analysis with one of the 'Bhut Jolokia' accessions (accession 2) has been shown. The ITS2 sequences with homologous structures were synchronously aligned using 4SALE (Seibel *et al.* 2006, 2008). CBC and hemi-CBC analyses were carried out using CBC Matrix feature in 4SALE and CBC Analyzer version 1.1 (Wolf *et al.* 2005).

3. Results

3.1 *ITS-rDNA sequence analysis*

Nucleotide sequence of approximately 750 bp corresponding to 18S partial, ITS1, 5.8S, ITS2 and 28S partial sequence

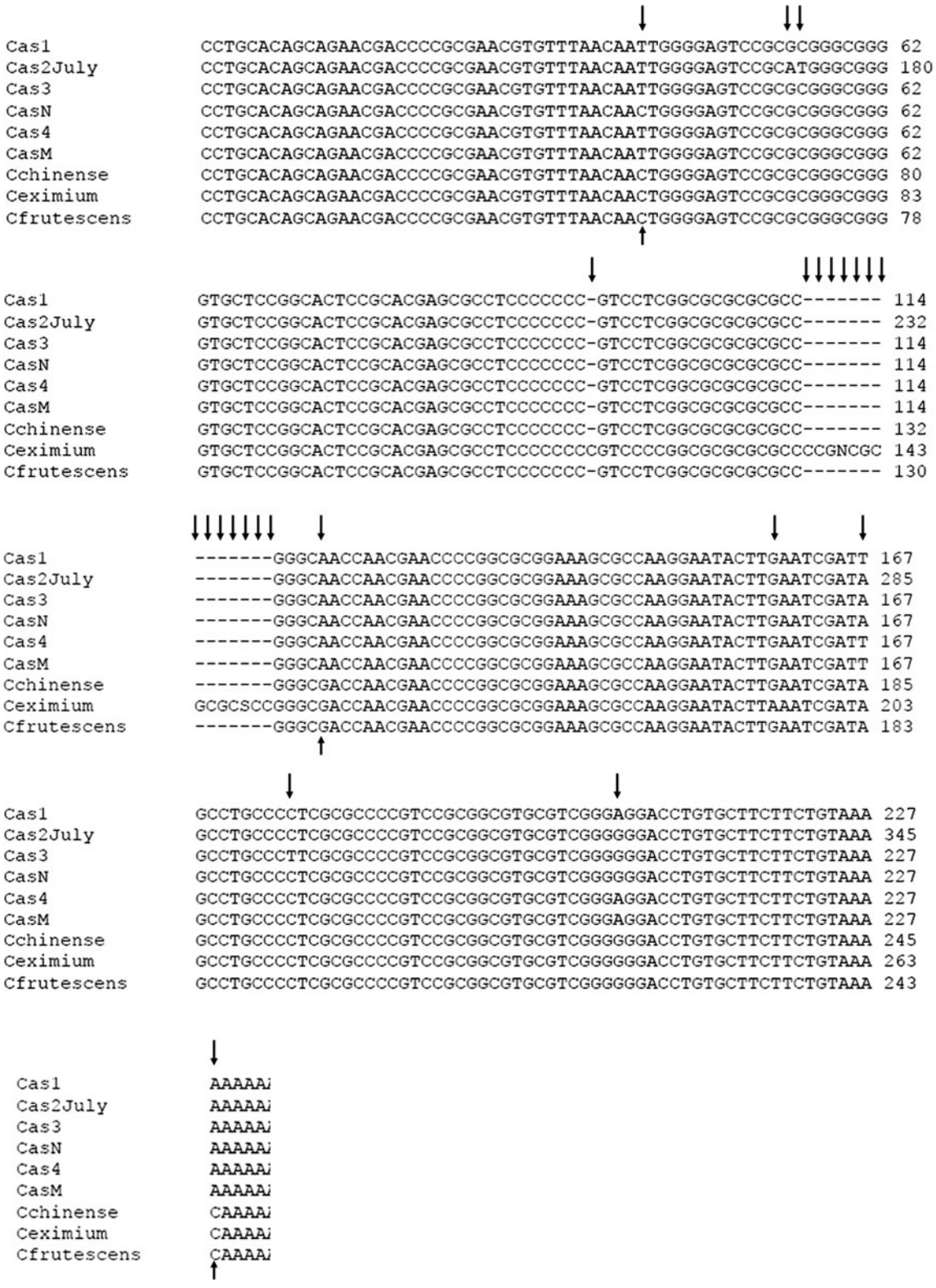

Figure 1. ClustalW alignment of ITS1 gene sequence (227 bases) from Bhut Jolokia and other related reference species of *Capsicum*. The polymorphic sites are shown with arrows above, and the consensus base substitutions in all the accessions of Bhut Jolokia in comparison to *C. frutescens* and/or *C. chinense* are shown with arrows below the alignment. The annotation 'Cas' in the alignment represents sequences from Bhut Jolokia accessions.

was obtained from the amplified DNA, for *C. frutescens*, *C. chinense*, and the six accessions of 'Bhut Jolokia'. The double-pass sequencing results from both leaf and fruit tissues from each accession were edited and aligned with related sequences of *Capsicum* spp. The alignment result (figures 1–3) indicated a 13-base deletion in the 5.8S rDNA region for all the six accessions of 'Bhut Jolokia' but not in *C. chinense* and *C. frutescens* (supplementary figure 2). To rule out a possibility of contamination and sequencing error, the PCR-amplified fragments were cloned in pTZ57R/T vector and four randomly selected clones were sequenced using vector-specific forward and reverse primers. The deduced sequences (~750 bases) were subjected to BLAST search for closest match and aligned with related sequences. The BLAST search results confirmed that the sequences were from the ITS region and a characteristic conserved GGCG sequence, which is presumed to be a recognition site for processing of a primary transcript into the structural rRNA (Liu and Schardl 1994), was observed in the ITS1 region from all the accessions of 'Bhut Jolokia' studied here. All the sequences from the four clones of any one accession were identical.

The sequences obtained from multiple clones (4) from each accession ruled out the possibility of contamination and sequencing error and further confirmed the presence of a 13-base deletion in 5.8S region of all 'Bhut Jolokia' accessions but not in the related species of *C. chinense* and *C. frutescens*. We further synthesized primers corresponding to the 13-base deletion region and performed PCR with it, using ITS2R as reverse primer. Amplification of expected length was observed using recombinant plasmids or total DNA from *C. chinense* and *C. frutescens* but not with those from accessions of 'Bhut Jolokia' (data not shown).

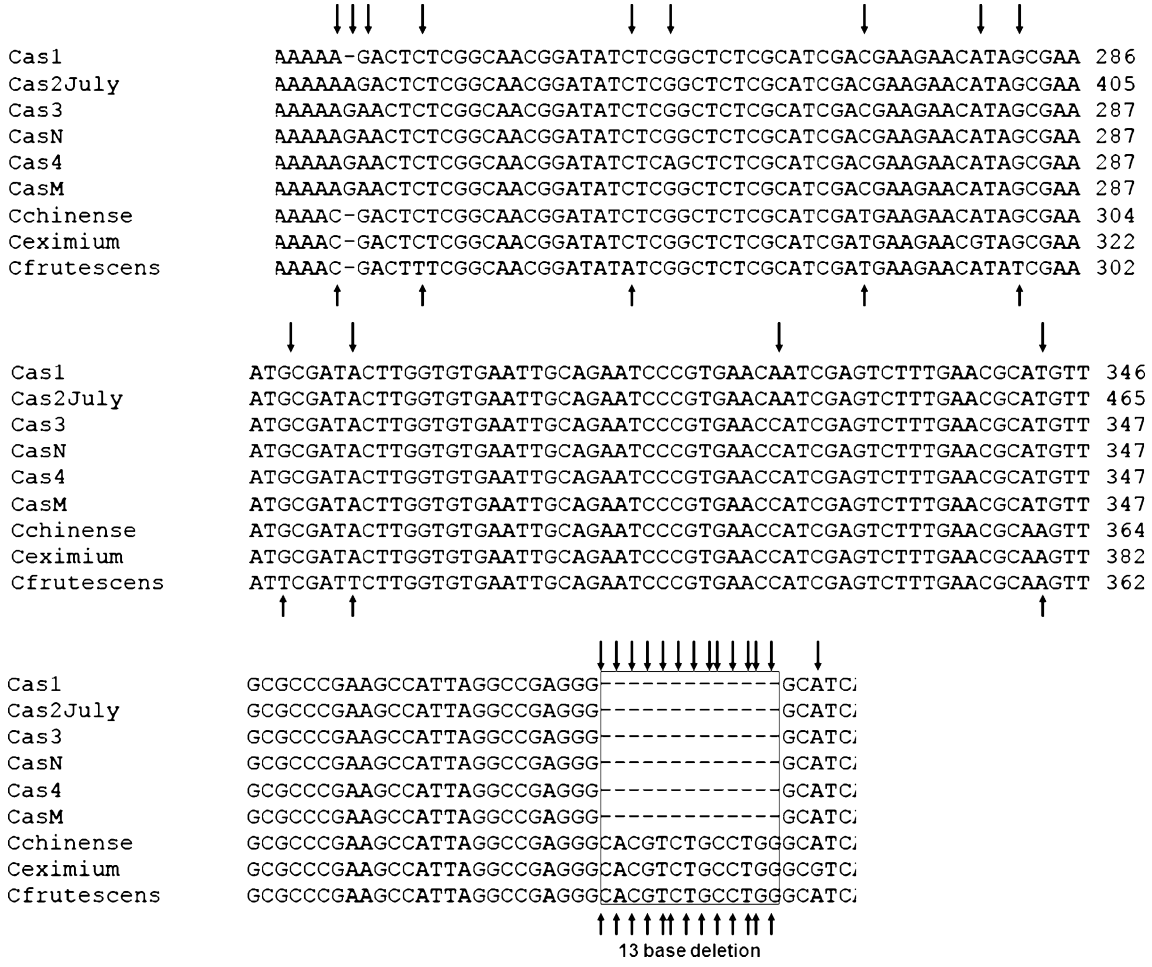

Figure 2. ClustalW alignment of 5.8S rRNA gene sequence (160 bases) from Bhut Jolokia and other related reference species of *Capsicum*. The polymorphic sites are shown with arrows above ,and the consensus base substitutions in all the accessions of Bhut Jolokia in comparison to *C. frutescens* and/or *C. chinense* are shown with arrows below the alignment. The annotation 'Cas' in the alignment represents sequences from Bhut Jolokia accessions.

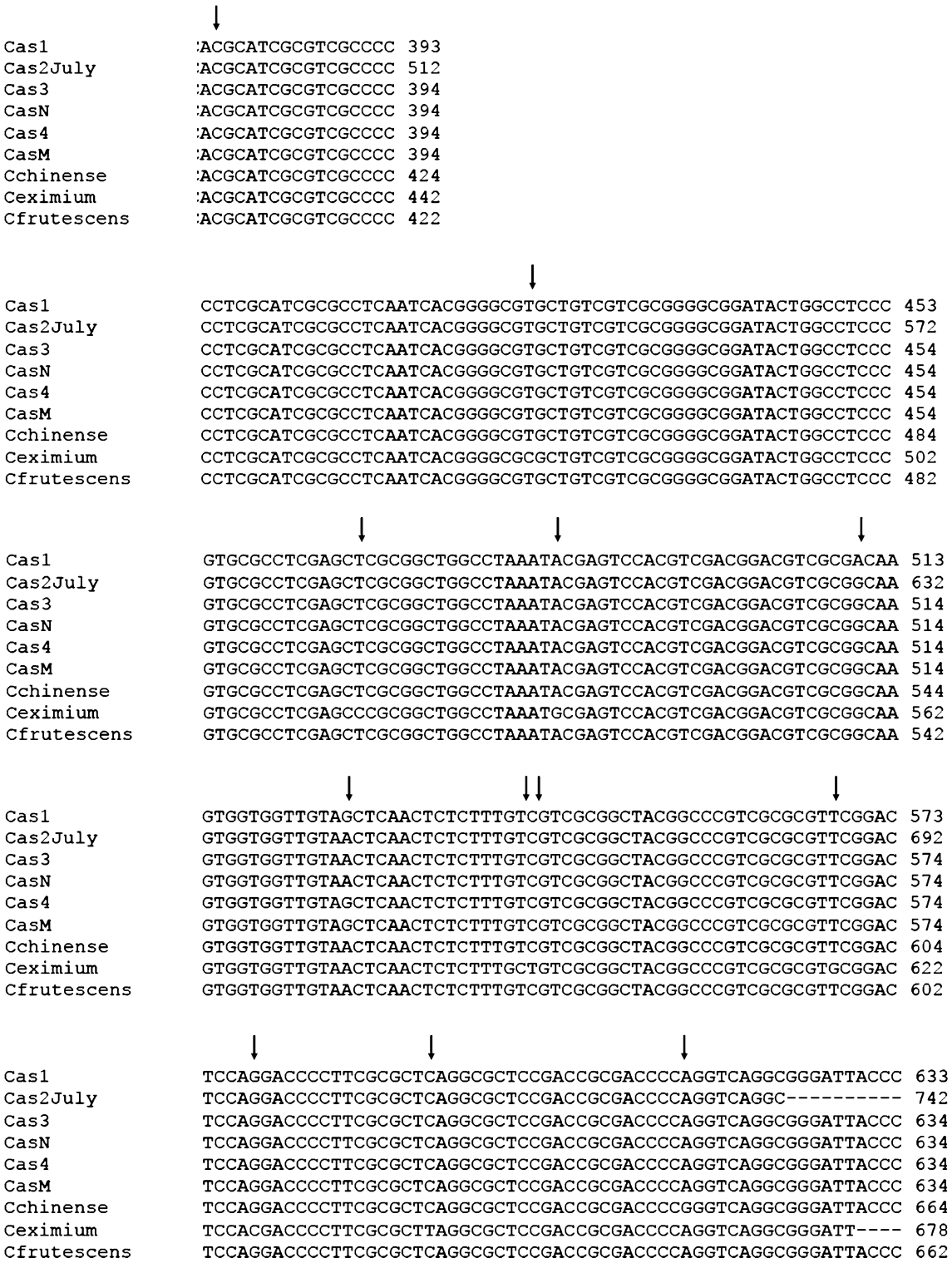

Figure 3. ClustalW alignment of ITS2 gene sequence (230 bases) from Bhut Jolokia and other related reference species of *Capsicum*. The polymorphic sites are shown with arrows above, and the consensus base substitutions in all the accessions of Bhut Jolokia in comparison to *C. frutescens* and/or *C. chinense* are shown with arrows below the alignment. The annotation 'Cas' in the alignment represents sequences from Bhut Jolokia accessions.

Table 3. Sequence characteristics of ITS-rDNA in plant the species/accessions used in this study

Species /Accessions	ITS1			5.8S			ITS2			Total no. of polymorphic sites
	Length (bp)	G+C (%)	Polymorphic sites	Length (bp)	G+C (%)	Polymorphic sites	Length (bp)	G+C (%)	Polymorphic sites	
C. eximum	242	70.24	-	160	54.37	-	230	69.56	-	-
C. chinense	227	69.16	16	160	53.12	2	230	68.26	8	26
C. frutescens	227	69.16	16	160	50.62	7	230	68.26	7	30
Bhut Jolokia acc.1	227	67.84	21	147	51.00	19	230	68.26	10	50
Bhut Jolokia acc.2	227	67.40	21	148	50.67	20	230	68.26	8	49
Bhut Jolokia acc.3	227	67.84	20	148	51.35	20	230	68.26	8	48
Bhut Jolokia acc.4	227	67.84	21	148	50.67	21	230	68.69	9	51
Bhut Jolokia acc.M	227	67.84	21	148	51.35	20	230	68.69	9	50
Bhut Jolokia acc.N	227	68.72	18	148	51.35	20	230	68.26	8	46

The number of polymorphic sites is with respect to the reference sequence of *C. eximum.*

Length of different regions and the base composition (%G+C) of the sequences for six 'Bhut Jolokia' accessions along with *C. chinense* and *C. frutescens* have been reflected in table 3. The length and G+C content of ITS region of 'Bhut Jolokia' was typically corresponding to the region of other species within the genus *Capsicum*. Among the different accessions of 'Bhut Jolokia', the total length of the entire ITS-rDNA varied from 604 to 605 bp and that of 5.8S region varied from 147 to 148 bp, while 227 bp ITS1 and 227 bp ITS2 sequence was observed for all the accessions sequenced (table 3). The presence of a 14 bp motif in 5.8S (5′-GAA TTG CAG AAT CC-3′), which is seed-plant specific (Jobes and Thien 1997), was located in all the sequences obtained in this study.

3.2 *Secondary structure predictions and phylogenetic analysis*

The homology modelling algorithm at ITS2 database yielded the lowest-energy secondary structures of ITS2 rDNA from 'Bhut Jolokia' accessions (supplementary table 1). The predicted structure had closest match with *C. eximium* and was used for structural comparison (supplementary figure 3 and supplementary table 1). A CBC was not observed in 'Bhut Jolokia' accessions when compared with *C. frutescens* or *C. chinense*; however, a hemi-CBC was located in the third helix with respect to *C. eximium* reference structure of ITS2 region (supplementary figure 3).

Relationships among different accessions of 'Bhut Jolokia' and its position with respect to the closely related species of

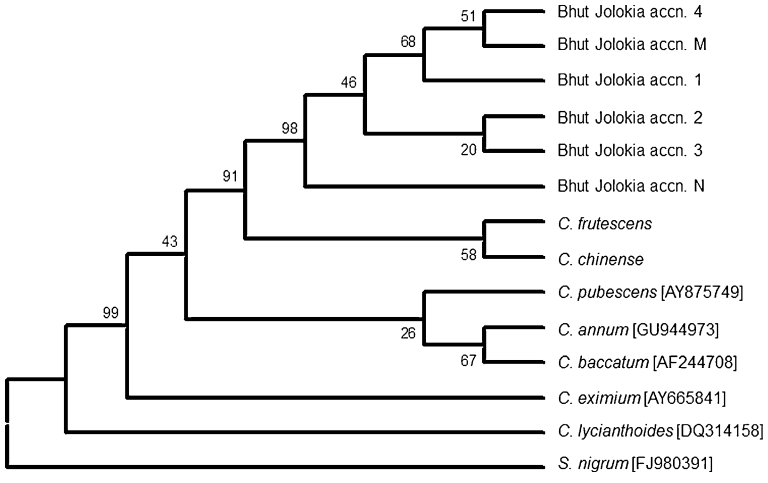

Figure 4. Maximum Parsimony trees showing phylogenetic relationship between Bhut Jolokia and other related reference species of *Capsicum* based on the ITS1–5.8S–ITS2 sequence (668 bases) analysis. Bootstrap values above 30 are given at the nodes. This is a rooted tree reconstructed with 1000 bootstrap replicates using tree construction software MEGA version 3.1.

C. frutescens and *C. chinense* were evaluated through phylogenetic analysis (figure 4) using ITS sequences (concatenated ITS1, 5.8S and ITS2). The topology of the ITS-sequence-based tree clearly indicated that the different accessions of 'Bhut Jolokia' could be resolved discretely in the form of a tightly clustering clade from the two related species of *C. frutescens* and *C. chinense* (figure 4). Pair-wise distance calculation using nucleotide maximum composite likelihood model for the entire ITS region used for phylogenetic analysis is reflected in supplementary table 2.

4. Discussion

The effect of climate on the scoville rating of 'Bhut Jolokia' is dramatic. Comparison of percentage availability of capsaicin and dihydrocapsaicin in 'Bhut Jolokia' grown in Tezpur (Assam) and Gwalior (Madhya Pradesh), India, showed that the heat of the pepper is decreased by over 50% in Gwalior's more arid climate (Tiwari *et al.* 2005). This clearly showed that 'hot and humid' climatic conditions of northeastern India are crucial for its extreme pungency (table 4). The heat and humidity of the Indian Northeast make it the ideal greenhouse for 'Bhut Jolokia'. Variation in capsaicinoid content is also known to be the directly related to the plant growth environment (Harvell and Bosland 1997; Zewdie and Bosland 2000; Blum *et al.* 2003). The diversified geological conditions, topographical characteristics, climatic situations and vegetation types favour the formation of different types of soil in the state of Assam and the Northeast. However, these minor variations do not seem to play any vital role in determining hotness level for 'Bhut Jolokia', as the major geoclimatic conditions remains more or less similar throughout the region. A radical change in climate to extreme dryness and to sandy loam soil, amply strewn with loose stones and pebbles, drastically reduces the hotness, as seen in Gwalior conditions (table 4). 'Bhut Jolokia' leaves get characteristic viral infection at a certain stage of growth, especially before fruit setting. This also may be one of the contributing factors towards pungency of 'Bhut Jolokia', as stress from virus infection is also known to increase pungency in *Capsicum* (Charles *et al.* 2005). As in other chillis (Estrada *et al.* 1999), water stress condition increases 'Bhut Jolokia' pungency.

In all cases, amplification of the ITS region produced a ~750 bp fragment, which is in agreement with the previous studies with other species of *Capsicum* (Ryzhova *et al.* 2002). Interestingly, a deletion of 13 bp in 5.8S rDNA region was found in all of the accessions of 'Bhut Jolokia'. ITS region sequencing and phylogenetic analysis of the different accessions of 'Bhut Jolokia' clearly indicate early delineation of this unique group of plant in the evolution of genus *Capsicum*. Homogenization of the ITS sequences among different accessions of 'Bhut Jolokia' but distinction from the other two related species of *C. frutescens* and *C. chinense* further indicate a divergent evolution after speciation of 'Bhut Jolokia'. Even if this speciation is the result of a hybridization event, it might have occurred in ancient times and concerted evolution appears to have homogenized the ITS sequences.

The nrDNA (including ITS) genes has the remarkable property that their paralogs, resulting from concerted evolution, within individuals, are quite homogenous. Unequal crossing over (Smith 1976) and gene conversion (Nagylaki 1984) are the proposed, underlying molecular processes involved in the homogenization. These paralogs will, however, display polymorphisms in individuals where concerted evolution is incomplete, for example, in cases where hybridization is involved (Muir *et al.* 2001), or when they are dispersed on non-homologous chromosomes in the genome (Wei and Wang 2004).

Reimer *et al.* (2007) sequenced the internal transcribed spacer of ribosomal DNA (ITS-rDNA) of *Symbiodinium*, the symbiotic dinoflagellates (zooxanthellae), from samples of designated *Zoanthus sansibaricus* Carlgren (Anthozoa: Hexacorallia) colonies. The authors observed consistent (145 unique sequences out of 153 total obtained sequences)

Table 4. Variation in pungency of Bhut Jolokia in different climatic localities of India and North East India

Locality	Climate and soil*	Latitude, longitude and altitude (meter)	Capsaicin (%w/w)	Pungency (SHU)	Reference
Tezpur	Hot humid climate with alluvial derived soil	26°38′N/92°48′E/157	4.28	855,000	Mathur *et al.* 2000; Tiwari *et al.* 2005
Pithoragarh	Warm to hot dry to moist subhumid climate with brown forest and podzolic soils	29°35′N/80°13′E/1525	0.97	254,896	Pandey *et al.* 2009
Imphal	Warm to hot perhumid climate with red and yellow soils	24°44′N/93°58′E/790	2.06	329,100	Sanatombi and Sharma 2008
Gwalior	Semi-arid climate with mixed red and black soils	26°14′N/78°15′E/205	1.5	NA	Tiwari *et al.* 2005

*From Velayutham *et al.* (1999); SHU: Scoville heat unit; NA: Not available.

genotypic microvariation in their obtained sequences including seven sequences exhibiting large deletions (6–66 bp) in ITS-rDNA and two sequences showing 6 bp and 25 bp deletions, respectively, in their 5.8S rRNA gene loci. The observed deletion in our study is restricted only to the accessions of 'Bhut Jolokia' and not observed in sequences from *C. frutescens* or *C. chinense*. It might be the result of multiple copies of the ITS-rDNA region within a single genome.

In the columnar cactus *Lophocereus*, Hartmann *et al.* (2001) showed that one of the operons encodes rDNA pseudogenes in a low copy-number (Truncated), whereas the second operon encodes an expressed rRNA (Functional). Extensive paralogy was observed not only in the ITS regions but also in the 5.8S coding regions in both within and between operons. Our sequencing results for multiple clones of the same accession do not indicate such a paralogous low-copy-number operon with the truncated 5.8S region. Moreover, pseudogenes are assumed to have escaped from functional constraints, accumulate many mutations and can cluster randomly across phylogenetic trees due to long-branch attraction (LBA), which confounds attempts to recover correct phylogenetic species relationships (Kita and Ito 2000). The formation of well-resolved clade for the accessions of 'Bhut Jolokia', and highly conserved ITS2 region (indicating non-random mutations), does not indicate the sequences reported here to be pseudogenes. However, the possibility that the 'Bhut Jolokia' genome harbours rDNA pseudogenes with truncated 5.8S region, in a high copy-number, needs to be tested using extensive sequencing of several clones or an alternative strategy. In that case, we do not know the physical origin of the pseudogenes isolated from the 'Bhut Jolokia'genomes here, whether they cluster together in one locus or occur interspersed among functional copies that we are unable to amplify.

We do not rule out the possibility of alternative rRNA processing mechanisms in 'Bhut Jolokia' to cope with the observed 13-base deletion. Typically, the 18S, 5.8S and 25S rRNAs are co-transcribed in the nucleolus as a polycistronic transcript that undergoes a complex series of endonucleolytic cleavages and exonucleolytic processing steps to yield the mature rRNAs. Processing of pre-rRNAs occurs within pre-ribosomal particles that contain, in addition to the pre-rRNA and ribosomal proteins, some 200 processing, modification and assembly factors and 75 small nucleolar RNAs. Only one pair of primers was used in the present study to amplify the ITS sequences; thus, only genes where the primer sites were greatly preserved were obtained. There also exists a possibility that older and/or more diverged copies, where the primer sites have been lost, likely exist, but are not PCR-amplifiable and represent the typical ITS sequences without the 13-base deletions observed here.

The 5.8S gene is the most reliable indicator of the functionality of ITS paralogs within the ITS region (Hershkovitz *et al.* 1999), and Yokota *et al.* (1989) suggested that the GC content of 5.8S in functional copies in plants ranges from 50.6% to 59.3%. 'Bhut Jolokia' 5.8S copies showed 51.6% to 51.3% GC content (table 3), which is well in line with the values of Yokota *et al.* (1989). Values for the putative pesudogenes have been reported (Xiao *et al.* 2010) with an average of 44.8%, significantly lower than the data observed here.

Length variation between 230 to 248 bp in 5.8S rRNA gene in six species from the ancient extant seed plant *Cycas* has been reported by Xiao *et al.* (2010). Based on the distribution of a 14 bp deletion in ITS2 region, an early evolutionary origin of the pseudogenes was indicated, arguably predating the diversification of *Cycas*. The authors also reported a 61 bp deletion in 5.8S (categorized as a pseudogene) in *C. circinalis* 16, the largest one observed in seed plants. Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species of *Acropora* was demonstrated and based on the secondary structure predictions; it was argued that the divergent sequence variants were not pseudogenes (Odorico and Miller 1997).

5. Conclusion

The present investigation is the first detailed genetic characterization of Bhut Jolokia, the hottest chilli in the world. Although climatic conditions have been shown to affect pungency to a larger extent, they do seem to have a significant effect on the genetic makeup of the 'Bhut Jolokia' species as evidenced from the ribosomal sequence data of accessions collected from different environmental settings. Irrespective of the molecular mechanism (hybridization followed by divergent or concerted evolution) involved, the data indicates remarkable genetic differences between Bhut Jolokia and the species of *C. frutescens* and *C. chinense*. The results presented here warrant effective discrimination of 'Bhut Jolokia', from *C. frutescens* and *C. chinense,* thus providing support towards its claim as a new species.

Acknowledgements

We thank R Vijayaraghavan, Director, DRDE, Gwalior, India, for providing the facilities and support required for part of this study. This work was funded by the Defence Research and Development Organization, India.

References

Andrews J 1984 Peppers: The domesticated Capsicums (Texas, Austin: University of Texas Press)

Backonja M, Wallace MS, Blonsky ER, Cutler BJ, Malan PJr, Rauck R and Tobias J 2008 NGX-4010, a high-concentration

capsaicin patch, for the treatment of postherpetic neuralgia: a randomised, double-blind study. *Lancet Neurol*. **71** 106–1112

Baek YM, Hwang HJ, Kim SW, Hwang HS, Lee SH, Kim JA and Yun JW 2008 A comparative proteomic analysis for capsaicin-induced apoptosis between human hepatocarcinoma (HepG2) and human neuroblastoma (SK-N-SH) cells. *Proteomics* **8** 4748–4767

Basu KS and De AK 2003 The Capsicum. *Medicinal and aromatic plants-industrial profiles*. (London: Taylor and Francis)

Baura G, Szaro TM and Bruns TD 1992 Gastrosuillus laricinus is a recent derivative of *Suillus grevillei*: molecular evidence. *Mycologia* **84** 592–597

Blum E, Mazourek M, O'Connell MA, Curry J, Thorup T, Liu K, Jahn MM and Paran I 2003 Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. *Theor. Appl. Genet.* **108** 79–86

Bohs L and Olmstead G 2001 A reassessment of Normania and Triguera (Solanaceae). *Plant Syst. Evol.* **228** 33–48

Bosland PW and Baral JB 2007 'Bhut Jolokia'-the world's hottest known chile pepper is a putative naturally occurring interspecific hybrid. *HortScience* **42** 222–224

Caterina MJ, Schumacher MA, Tominaga M, Rosen TA, Levine JD and Julius D 1997 The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* **389** 816–824

Charles S Jr, Byoung-Cheorl K, Kede L, Michael M, Shanna LM, Eun YY, Byung-Dong K, Ilan P, *et al.* 2005 The Pun1 gene for pungency in pepper encodes a putative acyltransferase. *Plant J.* **42** 675–688

Chen SL and Han JP 2009 Identification of medicinal plants using DNA barcoding technique. Unpublished

Coleman AW 2007 Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res.* **35** 3322–3329

Cullings KW and Vogler DR 1998 A 5.8S nuclear ribosomal RNA gene sequence databse: applications to ecology and evolution. *Mol. Ecol.* **7** 919–923

Darwin C 1845 *The voyage of H.M.S. Beagle* 2nd edition (New York: The Heritage Press)

Eddy S 1998 Profile hidden Markov models. *Bioinformatics* **14** 755–763

Estrada B, Pomar F, Dı´az J, Merino F and Bernal MA 1999 Pungency level in fruits of the Padro´ n pepper with different water supply. *Sci. Hortic.* **81** 385–396

Halldén C, Hansen M, Nilsson NO, Hjerdin A and Sall T 1996 Competition as a source of errors in RAPD analysis. *Theor. Appl. Genet.* **93** 1185–1192

Hartmann S, Nason JD and Bhattacharya D 2001 Extensive ribosomal DNA genic variation in the columnar cactus Lophocereus. *J. Mol. Evol.* **53** 124–134

Harvell KP and Bosland PW 1997 The environment produces a significant effect on pungency of chiles. *Hortsci.* **32** 1292

Hernández A, Aranda E, Martín A, Benito MJ, Bartolomé T and de Gúa Córdoba M 2010 Efficiency of DNA typing methods for detection of smoked paprika "pimenton de la vera" adulteration used in the elaboration of dry-cured Iberian pork sausages. *J. Agric. Food Chem.* **58** 11688–11694

Hershkovitz MA and Lewis LA 1996 Deep-level diagnostic value of the rDNA-ITS region. *Mol. Biol. Evol.* **13** 1276–1295

Hershkovitz MA, Zimmer EA and Hahn WJ 1999 Ribosomal DNA sequences and angiosperm systematic; in *Molecular systematics and plant evolution* (London:Taylor and Francis) pp 268–326

Jobes DV and Thien LB 1997 A conserved motif in the 5.8S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequences. *Plant Mol. Biol. Rep.* **15** 326–334

Jorgenson RD and Cluster PD 1988 Modes and tempos in the evolution of ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. *Ann. Missouri Bot. Garden* **75** 1238–1247

Kawada T, Hagihara K and Iwai K 1986 Effects of capsaicin on lipid metabolism in rats fed a high fat diet. *J. Nutr.* **116** 1272–1278

Keller A, Schleicher T, Schultz J, Müller T, Dandekar T and Wolf M 2009 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. *Gene* **430** 50–57

Kita Y and Ito M 2000 Nuclear ribosomal ITS sequences and phylogeny in East Asian Aconitum subgenus *Aconitum* (Ranunculaceae), with special reference to extensive polymorphism in individual plants. *Plant Syst. Evol*. **225** 1–13

Koetschan C, Förster F, Keller A, Schleicher T, Ruderisch B, Schwarz R, Müller T and Wolf M *et al.* 2010 The ITS2 Database III - sequences and structures for phylogeny. *Nucleic Acids Res.* **38** D275–9

Kumar S, Tamura K and Nei M 2004 MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinfo.* **5** 150–163

Lindsay K and Bosland PW 1995 A field study of environmental interaction on pungency. *Capsicum Eggplant Newsletter* **14** 36–38

Liu JS and Schardl CL 1994 A conserved sequence in internal transcribed spacer of plant nuclear rRNA genes. *Plant Mol. Biol.* **26** 775–778

Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ and Woese CR 1997 The RDP (Ribosomal Database Project). *Nucleic Acid Res.* **25** 109–110

Mathur R, Dangi RS, Das SC and Malhotra RC 2000 The hottest chilli variety in India. *Curr. Sci.* **79** 287–288

Mori A, Lehmann S, O'Kelly J, Kumagai T, Desmond JC, Pervan M, McBride WH and Kizaki M *et al.* 2006 Capsaicin, a component of red peppers, inhibits the growth of androgenin-dependent, p53 mutant prostate cancer cells. *Cancer Res.* **66** 3222–3229

Muir G, Fleming CC and Schlötterer C 2001 Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. *Mol. Biol. Evol.* **18** 112–119

Muller T, Philippi N, Dandekar T, Schultz J and Wolf M 2007 Distinguishing species. *RNA* **13** 1469–1472

Nagylaki T 1984 The evolution of multigene families under intra-chromosomal gene conversion. *Genetics* **106** 529–548

Odorico DM and Miller DJ 1997 Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species of Acropora (Cnidaria; Scleractinia): patterns of variation consistent with reticulate evolution. *Mol. Biol. Evol.* **14** 465–434

Pandey V, Pandey HK, Dayal D, Joshi UC, Pant T and Ahmed Z 2009 Studies on antioxidant constituents of some domesticated capsicums in the middle hill conditions of western Himalayas. *HortScience* **36** 26–30

Pino J, Gonzalez M, Ceballos L, Centurio´n-Yah AR, MTrujillo-Aguirre J, Latournerie-Moreno L and Sauri-Duch E 2007 Characterization of total capsaicinoids, colour and volatile compounds of Habanero chilli pepper (*Capsicum chinense* Jack.) cultivars grown in Yucatan. *Food Chem.* **104** 1682–1686

Reimer JD, Ono S, Tsukahara J, Takishita K and Maruyama T 2007 Non-seasonal clade-specificity and subclade microvariation in symbiotic dinoflagellates (Symbiodinium spp.) in *Zoanthus sansibaricus* (Anthozoa: Hexacorallia) at Kagoshima Bay, Japan. *Phycol. Res.* **55** 58–65

Reinbach HC, Smeets A, Martinussen T, Moller P and Westerterp-Plantenga MS 2009 Effects of capsaicin, green tea and CH-19 sweet pepper on appetite and energy intake in humans in negative and positive energy balance. *Clin. Nutr.* **28** 260–265

Ryzhova NN, Goryunova SV, Tomilo AA and Kochieva EZ 2002 Two types of rDNA Internal Transcribed Spacers (ITSs) in the *Capsicum* genome. *Doklady Biol. Sci.* **387** 539–541

Sanatombi K, Sharma GJ 2008 Capsaicin content and pungency of different Capsicum spp. cultivars. *Not. Bot. Hort. Agrobot. Cluj-Napo.* **36** 89–90

Schultz J, Maisel S, Gerlach D, Müller T and Wolf M 2005 A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA* **11** 361–364

Schultz J, Müller T, Achtziger M, Seibel PN, Dandekar T and Wolf M 2006 The internal transcribed spacer 2 database-a web server for (not only) low level phylogenetic analyses. *Nucleic Acid Res.* **34** W704–W707

Seibel PN, Müller T, Dandekar T, Schultz J and Wolf M 2006 4SALE - A tool for synchronous RNA sequence and secondary structure alignment and editing. *BMC Bioinformat*. **7** 498

Seibel PN, Müller T, Dandekar T, Wolf M 2008 Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. *BMC Res. Notes* **1** 91

Smith GP 1976 Evolution of repeated DNA sequences by unequal crossover. *Science* **191** 528–535

Smith SD and Baum DA 2006 Phylogenetics of the florally diverse Andean clade Iochrominae (Solanaceae). *Am. J. Bot.* **93** 1140–1153

Spooner DM, Ballard H, Stephenson SA and Polgar Z 2005 Phylogeny of wild potatoes (Solanum sect. Petota) using the nuclear ITS ribosomal DNA region. Unpublished.

Surh YJ 2002 Anti-tumor promoting potential of selected spice ingredients with antioxidative and anti-inflammatory activities: a short review. *Food Chem. Toxicol.* **40** 1091–1097

Tiwari A, Kaushik MP, Pandey KS, Dangi RS 2005 Adaptability and production of hottest chilli variety under Gwalior agro-climatic conditions. *Curr. Sci.***88** 1545–1546

Velayutham M, Mandal DK, Mandal C, Sehgal J 1999 Agro-ecological subregions of India for planning and development. NBSS and LUP, Publ. No. 35

Walsh BM and Hoot SB 2001 Phylogenetic relationships of Capsicum (Solanaceae) using DNA sequences from two non-coding regions: the chloroplast atpB-rbcl spacer region and nuclear waxy introns. *Int. J. Plant Sci.* **162**1409–1418

Wei XX and Wang XQ 2004 Recolonization and radiation in Larix (Pinaceae), evidence from nuclear ribosomal DNA paralogs. *Mol. Ecol.* **13** 3115–3123

Whitson M and Manos PS 2005 Untangling Physalis (Solanaceae) from the physaloids: a two-gene phylogeny of the Physalinae. *Syst. Bot.* **30** 216–230

White TJ, Bruns T, Lee S and Taylor J 1990 Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics; in *PCR protocols: A guide to methods and applications* (CA, San Diego: Academic press) pp 315–324

Williams JGK, Hanafey MK, Rafalski JA and Tingey SV 1993 Genetic analysis using random amplified polymorphic DNA markers. *Methods Enzymol.* **218** 705–740

Wolf M, Achtziger M, Schultz J, Dandekar T and Müller T 2005 Homology modeling revealed more than 20,000 rRNA internal transcribed spacer 2 (ITS2) secondary structures. *RNA* **11** 1616–1623

Xiao LQ, Möller M and Zhu H 2010 High nrDNA ITS polymorphism in the ancient extant seed plant Cycas: Incomplete concerted evolution and the origin of pseudogenes. *Mol. Phylogenet. Evol.* **55** 168–177

Xu J 2006 Fundamentals of Fungal Molecular Population Genetic Analyses. *Curr. Issues Mol. Biol.* **8** 75–90

Yokota Y, Kawata T, Iida Y, Kato A and Tanifuji S 1989 Nucleotide sequences of the 5.8S rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. *J. Mol. Evol.* **29** 294–301

Zewdie Y and Bosland PW 2000 Evaluation of genotype, environment, and genotype-by-environment interaction for capsaicinoids in *Capsicum annum* L. *Euphytica* **111** 185–190

Zhang LL, Yan Liu D, Ma LQ, Luo ZD, Cao TB, Zhong J, Yan ZC, Wang LJ, *et al.* 2007 Activation of transient receptor potential vanilloid type-1 channel prevents adipogenesis and obesity. *Circ. Res.* **100** 1063–1070

MS received 26 June 2012; accepted 16 July 2012

Corresponding editor: MAN MOHAN JOHRI