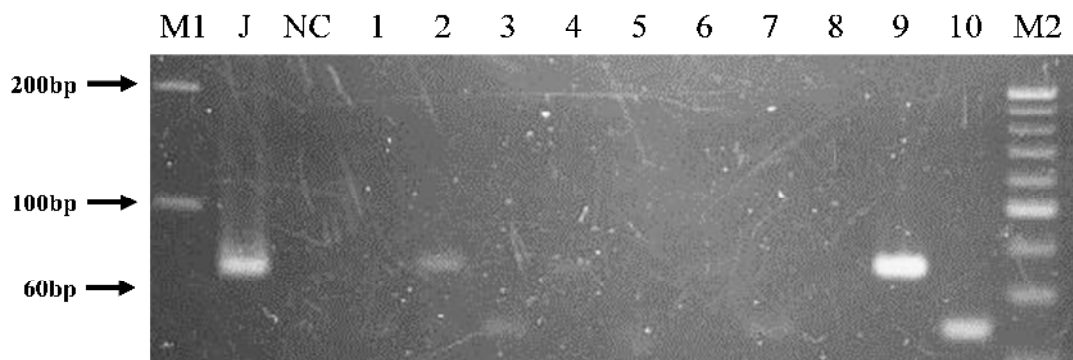


Figure S1: PCR amplification for the nuclear genome region *Sh4*; and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with specific primer sets, F2 and R2 primers. M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), J: Modern *japonica* cv. 'Nipponbare', NC: negative control (dH₂O), 1-10: represent remains from Ban Non Wat V200 7:Σ3 Δ2 Samples 11-20. **(b)** Sequence detected in wild rice, *japonica* and archaeological remains from Ban Non Wat V200 7:Σ3 Δ2. *Wild rice*: Modern *O. rufipogon* (EU999926); *Japonica*: Modern *japonica* cv. 'Nipponbare' (NC_008397); Remains: from Ban Non Wat V200 7:Σ3 Δ2 represent 14, 15 and 19. "*" shows identical nucleotide sequencing to *Wild rice*.

a)



b)

<i>Wild rice</i>	ATGCTATTGATGTATACTGGACATTGCATGATACACGTACCACCAAATGCTAAGGATCTT
<i>Japonica</i>	*****T*****
Remains	*****G*****
	
	F primer R primer
<i>Wild rice</i>	TGGACGAGATG
<i>Japonica</i>	*****
Remains	*****
	
	R primer

Figure S2: PCR amplification for the nuclear genome region *qSh1*; and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with specific primer sets, F and R primers. M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), J: Modern *japonica* cv. 'Nipponbare', NC: negative control (dH₂O), 1-10: represent remains from Noen U-Loke #105 Samples 1-10. **(b)** Sequence detected in wild rice, *japonica* and archaeological remains from Noen U-Loke #105. *Wild rice*: Modern *O. rufipogon* (EU999846); *Japonica*: Modern *japonica* cv. 'Nipponbare' (NC_008394); Remains: from Noen U-Loke #105 represent 2 and 9. "*" shows identical nucleotide sequencing to *Wild rice*.

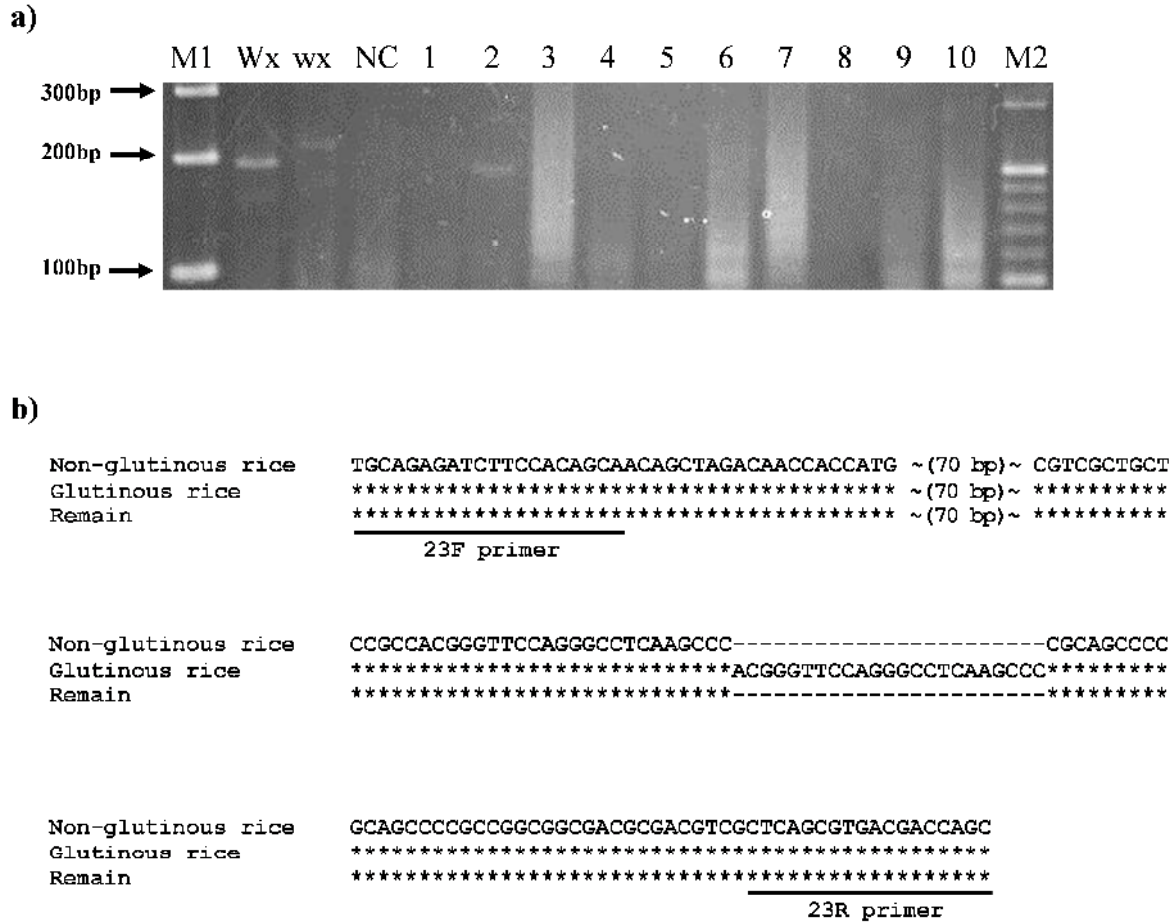
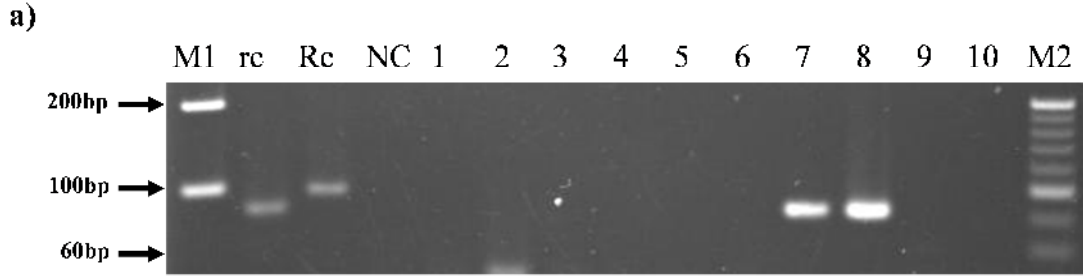


Figure S3: PCR amplification for the nuclear genome region *Waxy*; and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with specific primer sets, F and R primers. M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), *Wx*: Modern *japonica* cv. 'Nipponbare', *wx*: Modern *japonica* glutinous cv. accession no. DQ280635, NC: negative control (dH₂O), 1-10: represent remains from Ban Non Wat K500 4:2 GEN Δ Samples 1-10. **(b)** Sequence detected in non-glutinous rice, glutinous rice and archaeological remains from Ban Non Wat K500 4:2 GEN Δ. *Non-glutinous rice*: Modern *japonica* cv. 'Nipponbare' (NC_008399); *Glutinous rice*: accession no. DQ280635; Remains: from Ban Non Wat K500 4:2 GEN Δ represent 12. "-" in the sequence symbolises a gap introduced to optimise alignment, whereas "*" shows identical nucleotide sequencing to *non-glutinous rice*.



b)

```

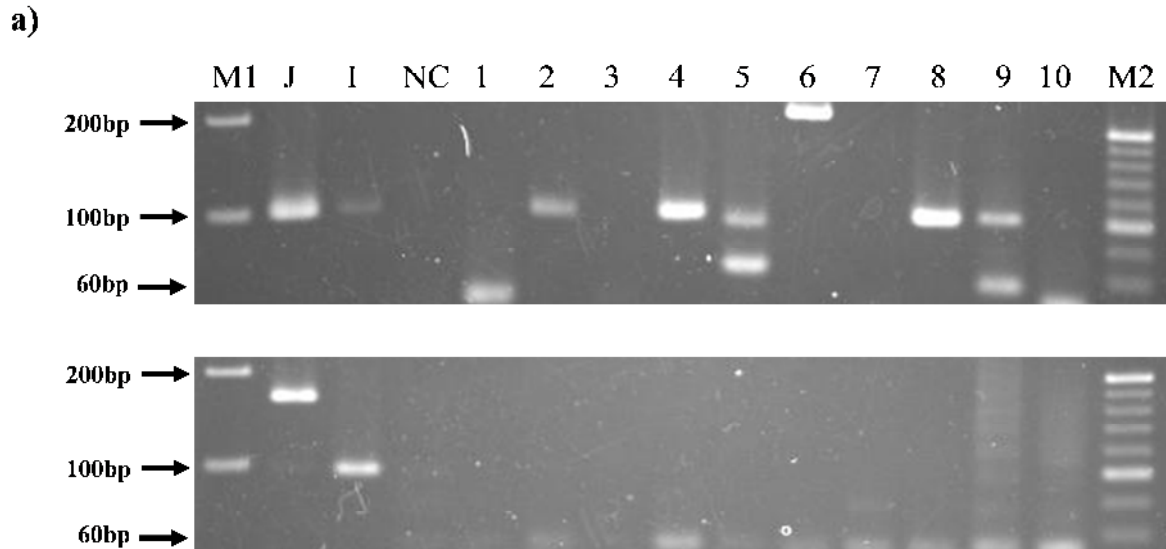
White pericarp    CACCTGAATCAAGGGCGGGAAAGGCCCAAGTGGG-----TGCCATCCAAG
Red pericarp     *****ACGCGAAAAGTCGG*****
Remain           *****-----*****

                F primer

White pericarp    GTGATTTTCAGTGCCAACCATGTGCTGAAAGAGA
Red pericarp     *****
Remain           *****

                R primer
  
```

Figure S4: PCR amplification for exon 7 in the *Rc* gene; and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with specific primer sets, F and R primers. M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), *rc*: Modern *japonica* cv. 'Nipponbare', *Rc*: Modern *indica* cv. 'AC417', NC: negative control (dH₂O), 1-10: represent remains from Ban Non Wat K500 4:2 GEN Δ Samples 1-10. **(b)** Sequence detected in rice with a white and red pericarp, and archaeological remains from Ban Non Wat K500 4:2 GEN Δ. *White pericarp*: Modern *japonica* cv. 'Nipponbare' (NC_008400); *Red pericarp*: Modern *indica* cv. 'Kasalath' (AB247503); Remains: from Ban Non Wat K500 4:2 GEN Δ represent 7 and 8. "-" in the sequence symbolises a gap introduced to optimise alignment, whereas "*" shows identical nucleotide sequencing to *White pericarp*.



b)

```

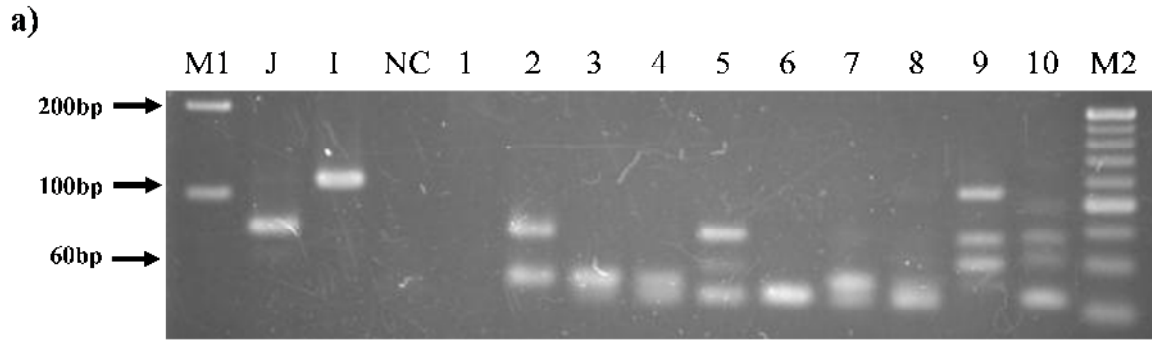
Indica          TGGATTTTCGAAAGTCAATTTTTCTTTTCAATATCTTTACTTTTTTCA-----
Japonica      *****GAATCCTATTTT
Remains         *****GAATCCTATTTT
                _____
                F1 primer

Indica          -----TTT
Japonica      TGTTCCTTATACCCATGCAATAGAGAGCGAGTGGGAAAAGGGAGGTTACTTTTTTCA***
Remains         TGTTCCTTATACCCATGCAATAGAGAGCGAGTGGGAAAAGG
                _____
                R1 primer

Indica          TTCCCTTAAAAAATAGGCTTTCTTGAAAATAGCAATCATGGA
Japonica      *****
                _____
                R2 primer

```

Figure S5: Examples of PCR amplification for the chloroplast genome region *Orf100* and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with F and R1 primer set (upper image) and F and R2 primer set (lower). M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), J: Modern *japonica* cv. ‘Nipponbare’, I: Modern *indica* cv. ‘IR36’, NC: negative control (dH₂O), 1-10: represent remains from Noen U-Loke #105 Samples 1-10. **(b)** Sequence detected in *indica*, *japonica* and archaeological remains from Noen U-Loke. *Indica*: Modern *indica* (NC_008155); *Japonica*: Modern *japonica* cv. ‘Nipponbare’ (NC_001320); Remains: from Noen U-Loke represent samples 2,4,5,8 and 9. “-” in the sequence symbolises a gap introduced to optimise alignment, whereas “*” shows identical sequencing to *indica*.



b)

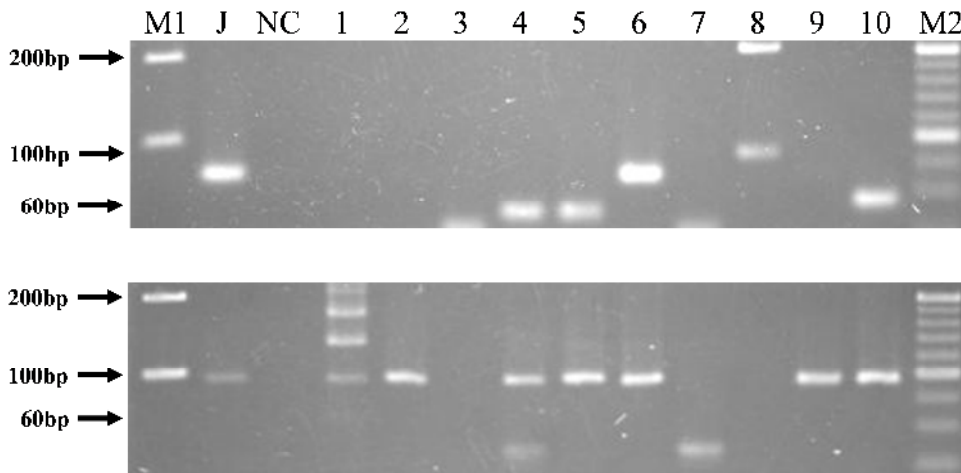
```

Indica      ATCAGTTCAAACAATTTACTCTTAACAAATTTCTAGAGTATTTCTGGTAGAATTTAACAA
Japonica    *****-----
Remains     *****-----
            F2 primer

Indica      ATTCTTAGAGTATTTCTGGTAGAATTGGGGAGCATTAAGTATAAATA
Japonica    -----*****
Remains     -----*****
            R2 primer
  
```

Figure S6: Examples of PCR amplification for the chloroplast genome region *PetN-TrnC* and the sequence analysis, indicating presence of deletion characteristic of *japonica* rice in archaeological specimens. **(a)** Electrophoresis of 10 PCR products amplified with specific primer set, F and R primers. M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), J: Modern *japonica* cv. 'Nipponbare', I: Modern *indica* cv. 'IR36', NC: negative control (dH₂O), 1-10: represent remains from Noen U-Loke #105 Samples 1-10. **(b)** Sequence detected in *indica*, *japonica* and archaeological remains from Noen U-Loke. *Indica*: Modern *indica* (NC_008155); *Japonica*: Modern *japonica* cv. 'Nipponbare' (NC_001320); Remains: from Noen U-Loke represent 2,5,9 and 10. "-" in the sequence symbolises a gap introduced to optimise alignment, whereas "*" shows identical sequencing to *indica*.

a)



b)

```

Indica      GCGCGAGTATCCGAAACTGTAGCTAGAGTAGCTATTTCCATAGCTGCCAGCAAAATGCCCC
Japonica   *****CG*****
Remains    *****CG*****
           _____
           F primer                                     R primer

Indica      ATACGAAGTCAATTTCTTCGGTTAGAAATATAACCCCCCAAAAAAAGTAGTATTGAA
Japonica   *****-----*****
Remains    *****-----*****
           _____
           2F primer
           _____
           R primer

Indica      AATAAAAACCAGGTTCTTCTTTCTTGAAAGACAATATTTCTTTC
Japonica   *****
Remains    *****
           _____
           2R primer

```

Figure S7: Examples of PCR amplification for the chloroplast genome region *Rp114-Rp116* and *Rp116* (the PS-ID region); and the sequence analysis. **(a)** Electrophoresis of 10 PCR products amplified with *Rp116* specific primer set (upper) and *Rp114-Rp116* specific primer sets (lower). M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), J: Modern *japonica* cv. 'Nipponbare', NC: negative control (dH₂O), 1-10: represent remains from Ban Non Wat K500 4:2 GEN Samples 1-10. **(b)** Sequence detected in *indica*, *japonica* and archaeological remains from Ban Non Wat K500 4:2 GEN. *Indica*: Modern *indica* (NC_008155); *Japonica*: Modern *japonica* cv. 'Nipponbare' (NC_001320); Remains: from Ban Non Wat K500 4:2 GEN represent 1,2,4,5,6,9 and 10. "-" in the sequence symbolises a gap introduced to optimise alignment, whereas "*" shows identical nucleotide sequence to *indica*.

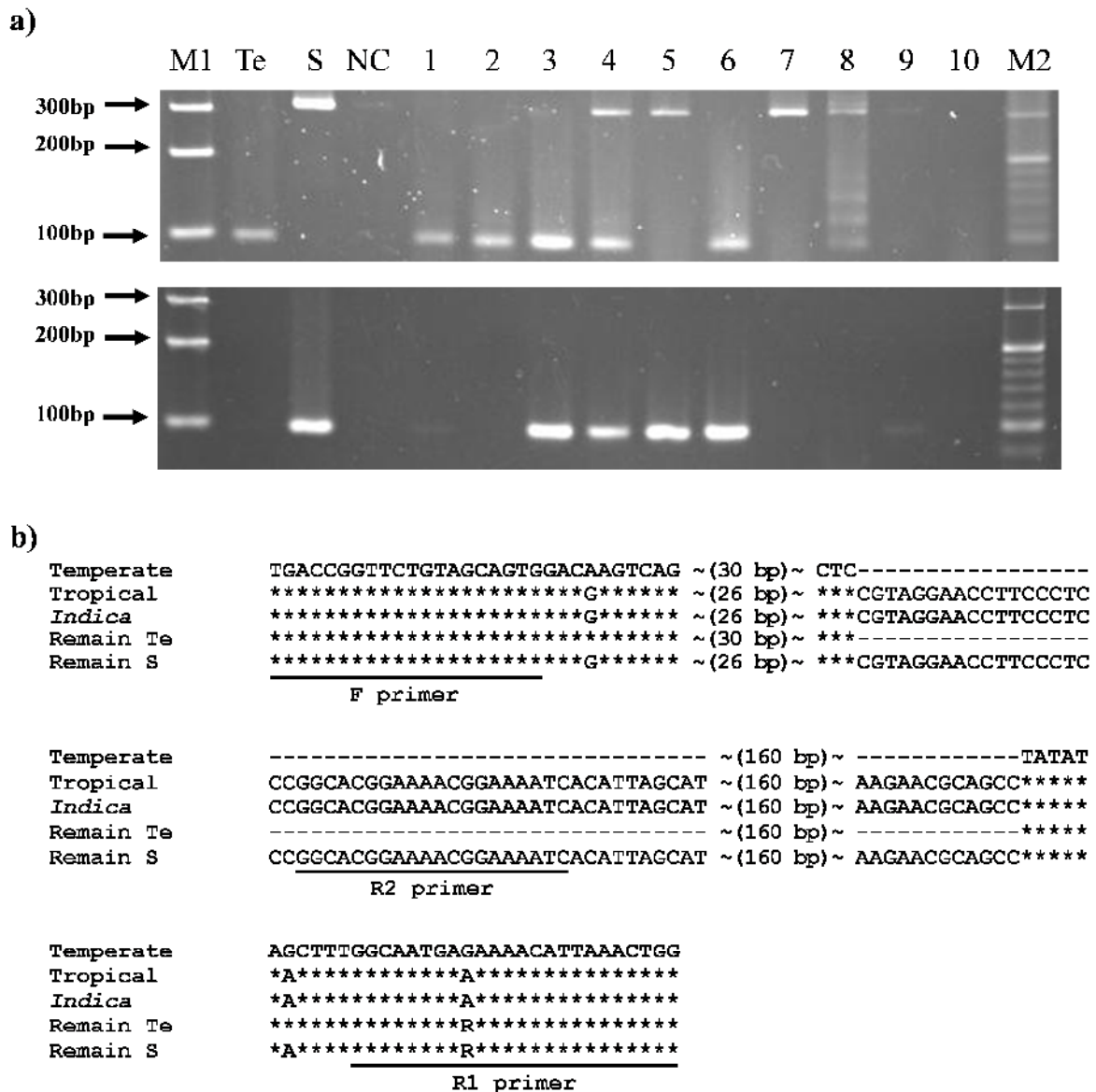


Figure S8: Polymorphism in a non-coding region in chromosome 6 (Ch6); and the sequence analysis. **(a)** Electrophoresis of PCR products amplified with specific primer sets, Ch6 F and R1 primers (upper) and Ch6 and R2 primers (lower). M1: 100bp DNA Ladder (TaKaRa, Japan), M2: 20bp DNA Ladder (TaKaRa, Japan), Te: Modern temperate *japonica* cv. 'Nipponbare', S: Modern tropical *japonica* cv. 'AC220', NC: negative control (dH₂O), 1-10: represent remains from Ban Non Wat K500 4:2 GEN Δ Sample 11-20. **(b)** Sequence detected in temperate *japonica*, tropical *japonica* and archaeological remains from Ban Non Wat K500 4:2 GEN Δ. *Temperate*: Modern *japonica* cv. 'Nipponbare' (NC_008399); *Tropical*: Modern *japonica* cv. 'AC220'; *Indica*: Modern *indica* cv. 'IR36'; *Remain Te*: from Ban Non Wat K500 4:2 GEN Δ represent samples 11, 12, 13, 14 and 16; *Remain S*: from Ban Non Wat K500 4:2 GEN Δ represent samples 14, 17, 18 and 19. "-" in the sequence symbolises a gap introduced to optimise alignment, whereas "*" shows identical nucleotide sequencing to *Temperate*.

