Supplementary Material for

Smash++: finding rearrangements

Morteza Hosseini[1], Diogo Pratas[1,2], Armando J. Pinho[1]

[1]IEETA/DETI, University of Aveiro, Portugal
[2]Department of Virology, University of Helsinki, Finland

{seyedmorteza,pratas,ap}@ua.pt

Contents

Note S1 GGA 18 compared to MGA 20

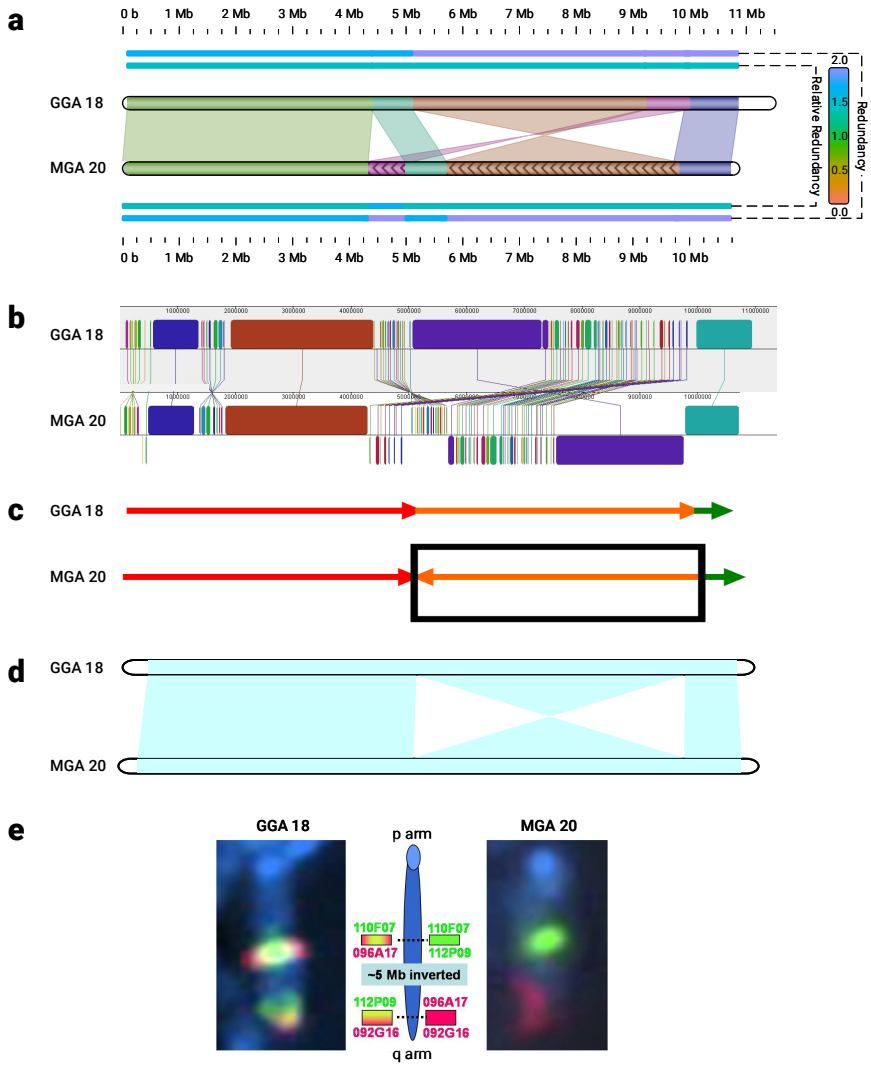

Fig. S1. (a) ; (b) adopted from []. FISH confirmation of the turkey-chicken inversion rearrangement due to apparent unequal recombination between NME1 and NME2 orthologs on GGA18/MGA20. CHORI-260 BACs 110F07 (GGA18 end coordinates: 4,850,650–5,056,016) and 112P09 (9,665,995–9,865,995) were labeled in green (FITC), while 96A17 (5,087,535–5,266,203) and 92G16 (9,980,713–10,142,396) were labeled with red (Enzo Red) and used for FISH analysis of chicken and turkey pachytene chromosomes, which are 14–20× more extended than mitotic metaphase chromosomes allowing for greater resolution. A view of GGA18 (left frame) affirms the arrangement predicted by the BES alignments noted above, 110F07 and 96A17 signals co-localize to generate a yellow signal halfway along the chromosome q arm, as do 112P09 and 92G16 near the q terminus. Whereas for MGA20 (right frame), the 110F07 and 112P09 BAC probes co-localize (green) as do the two red probes, indicative of the 5 Mb inversion. (Prior FISH experiments utilized the BAC probes singly or in pairs of two to ensure all probes hybridized equally well.) This inversion was previously indicated by inconsistent BAC mate pairs: CHORI-260 111D05 (5,106,305–10,099,832), 95I22 (5,109,664–10,107,855), 89F20 (5,134,762–10,035,123), 94C02 (5,157,115–10,042,702), and 95H13 (5,268,786–9,982,916) and 78TKNMI 18A07 (5,109,437–10,066,115), all of which had BES that aligned with the same strand in the chicken sequence, as expected for BACs that cross inversion breakpoints. Additional FISH, overgo mapping, and fingerprint analyses confirm the inversion and narrow the breakpoint regions to sites near the NME1 and NME2 orthologs (unpublished data) ; (c) SynBrowser ; (d)

Note S2 GGA 14 compared to MGA 16

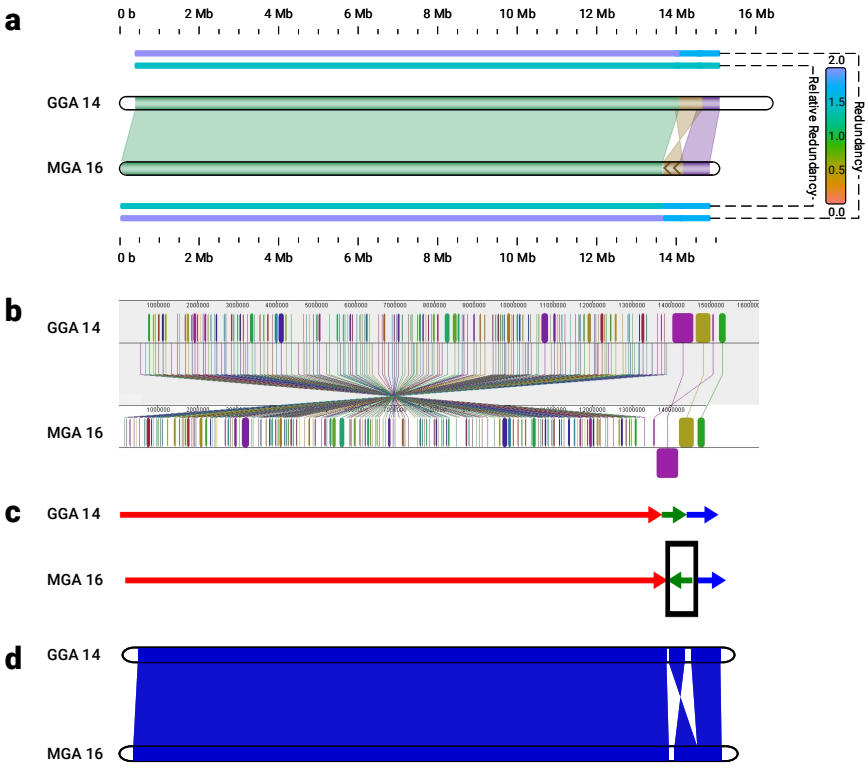

Fig. S2. (a) ; (b) adopted from []. ; (c) SynBrowser ; (d)

Note S3 HS 12 compared to PT 12

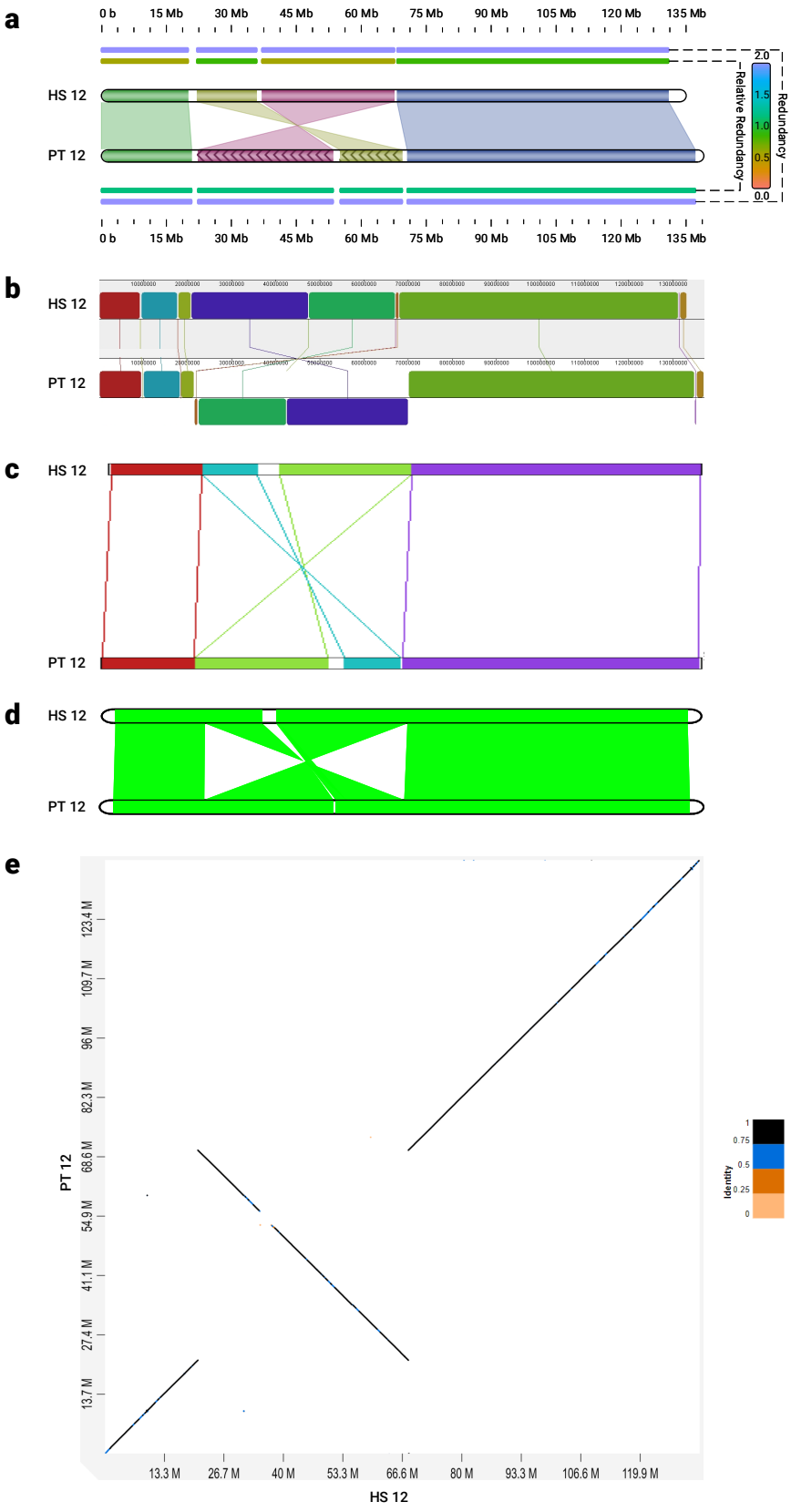

Fig. S3. (a) ; (b) adopted from []. ; (c) SynBrowser ; (d)

Note S4 PXO99^{A} compared to MAFF 311018

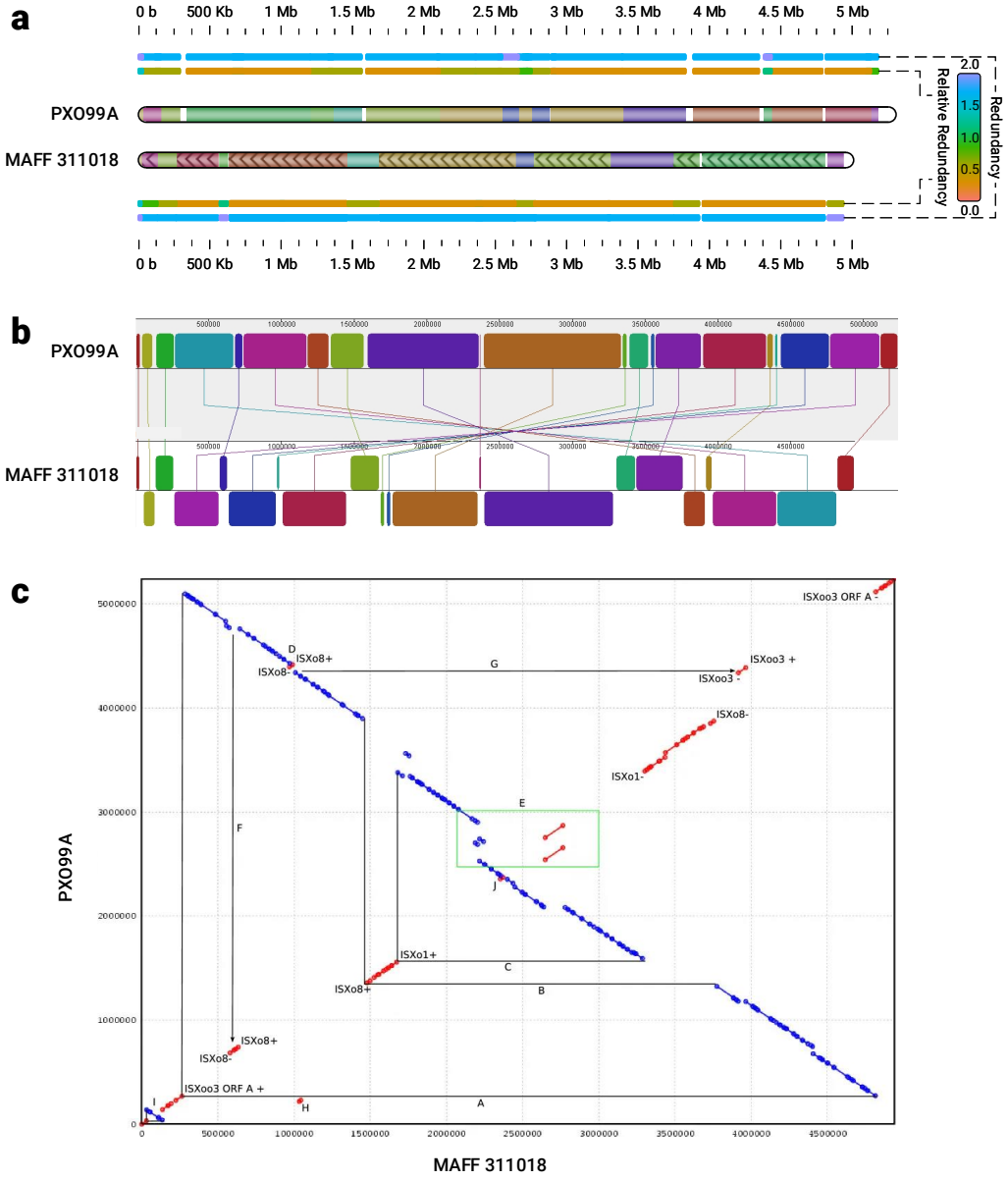

Fig. S4. adopted from [1]. Inversions and rearrangements in PXO99A compared to MAFF. The alignment shows regions of PXO99A that align to the same (red) or opposite (blue) strand of MAFF. Transposase genes and their orientation (+ or -) are shown at the sites of each rearrangement. Letters A-J indicate specific rearrangement events. A: the IS element ISXoo3 is composed of two distinct and independently conserved ORFs and is responsible for an inversion spanning coordinates 267869–5114959 (all coordinates refer to the PXO99A genome). B: ISXo8 occurs in opposite orientation at each end of a 2.6 Mbp inversion spanning positions 1356757–3898472. C: ISXo1 occurs in inverted copies at the endpoints of a 1.8 Mbp inversion spanning 1558996–3391786. D: a 33270 bp inverted region spanning 4394742–4428012 is flanked by oppositely-oriented copies of ISXo8. E: Each copy of the 212-kb duplication is flanked by ISXo5, which also occurs adjacent to two other translocations in this region. The duplication appears as two parallel diagonal lines in this box. F: ISXo8 also occurs in inverted copies at the boundaries of a 47540 bp segment that is translocated from approximately 4800000 to 685272. G: ISXoo3 flanks both ends of a 47540 bp translocation from approximately 1117000 to 4339239. H: A 9,862 bp region occurs in inverted copies at 217,455 and 4,305,307. MAFF311018 contains only one copy of this region. I,J: Segments spanning 96,753 bp (I) and 17,021 bp (J) are inverted with respect to MAFF311018 but not associated with transposases.

References

[1] S. L. Salzberg, D. D. Sommer, M. C. Schatz, A. M. Phillippy, P. D. Rabinowicz, S. Tsuge, A. Furutani, H. Ochiai, A. L. Delcher, D. Kelley *et al.*, "Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* pxo99a," *BMC genomics*, vol. 9, no. 1, p. 204, 2008.