1 Figures

1.1 Full tree

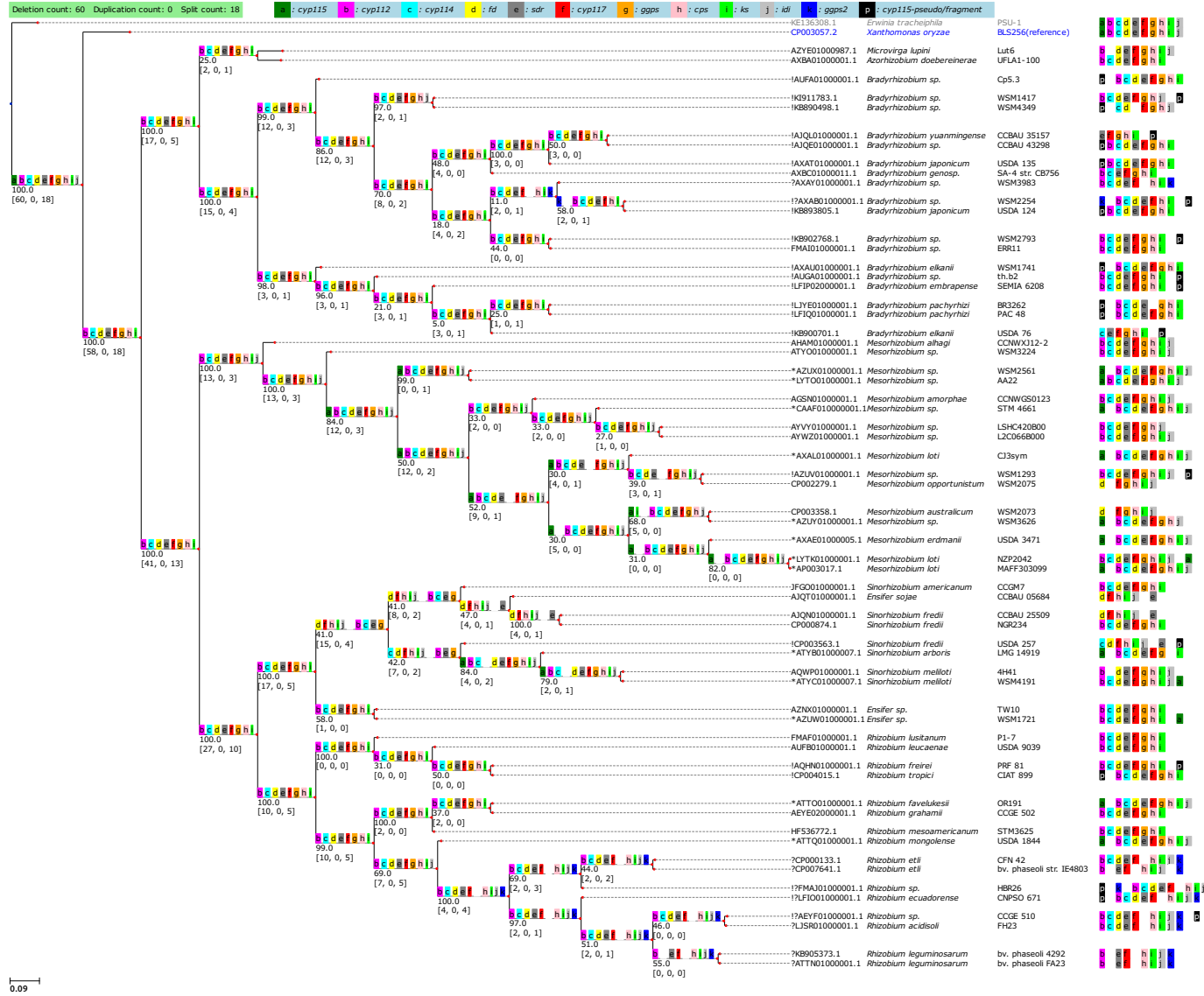

Figure 1: Ancestral reconstruction of operon Gibberellin using *rpoB* species tree (contains 120 species).The lower-case letters in each tree node represent the genes in the orthoblock (e.g. "a" represents "*cyp115*", we also color code genes, see legend at top). A ' ' (blank space) designates a split (i.e. a distance ≥500bp between the genes to either side of the ' '). The green bar on top left shows the total number of events that took place in this reconstruction. For each inner node *u*, the floating number (e.g. 98.0) represents the bootstrap value of the tree, the numbers in the brackets are a 3-tuple showing the cumulative count of events going from the leafnodes to node *u* in the following order: [deletions, duplications, splits]. Each leaf node is accompanied with some symbols(*, ?, !), the accession number, the species name, and the gene block. An asterisk (*) indicates the gene block contains gene *cyp*115, an exclamation ! indicates the gene block contains a pseudo or fragment of gene *cyp*115, and ? indicates the gene block contains gene *ggps*2. The reference genome *Xanthomonas oryzae pv. oryzicola* is in blue, and the outgroup genome *Erwinia tracheiphila* is in gray. These naming and color conventions persist through this study.

2

Figure 2: Ancestral reconstruction of operon Gibberellin using operon tree (contains 120 species).The lower-case letters in each tree node represent the genes in the orthoblock (e.g. "a" represents "*cyp115*", we also color code genes, see legend at top). A ' ' (blank space) designates a split (i.e. a distance ≥500bp between the genes to either side of the ' '). The green bar on top left shows the total number of events that took place in this reconstruction. For each inner node *u*, the floating number (e.g. 98.0) represents the bootstrap value of the tree, the numbers in the brackets are a 3-tuple showing the cumulative count of events going from the leafnodes to node *u* in the following order: [deletions, duplications, splits]. Each leaf node is accompanied with some symbols(*, ?, !), the accession number, the species name, and the gene block. An asterisk (*) indicates the gene block contains gene *cyp*115, an exclamation ! indicates the gene block contains a pseudo or fragment of gene *cyp*115, and ? indicates the gene block contains gene *ggps*2. The reference genome *Xanthomonas oryzae pv. oryzicola* is in blue, and the outgroup genome *Erwinia tracheiphila* is in gray. These naming and color conventions persist through this study.

3

1.2 Partial tree

Figure 3: Ancestral reconstruction of operon Gibberellin using reduced *rpoB* species tree (contains 64 species).The lower-case letters in each tree node represent the genes in the orthoblock (e.g. "a" represents "*cyp115*", we also color code genes, see legend at top). A ' ' (blank space) designates a split (i.e. a distance ≥500bp between the genes to either side of the ' '). The green bar on top left shows the total number of events that took place in this reconstruction. For each inner node *u*, the floating number (e.g. 98.0) represents the bootstrap value of the tree, the numbers in the brackets are a 3-tuple showing the cumulative count of events going from the leafnodes to node *u* in the following order: [deletions, duplications, splits]. Each leaf node is accompanied with some symbols(∗, ?, !), the accession number, the species name, and the gene block. An asterisk (∗) indicates the gene block contains gene *cyp*115, an exclamation ! indicates the gene block contains a pseudo or fragment of gene *cyp*115, and ? indicates the gene block contains gene *ggps*2. The reference genome *Xanthomonas oryzae pv. oryzicola* is in blue, and the outgroup genome *Erwinia tracheiphila* is in gray. These naming and color conventions persist through this study.

5

Figure 4: Ancestral reconstruction of operon Gibberellin using reduced operon species tree (contains 64 species).The lower-case letters in each tree node represent the genes in the orthoblock (e.g. "a" represents "*cyp115*", we also color code genes, see legend at top). A ' ' (blank space) designates a split (i.e. a distance ≥500bp between the genes to either side of the ' '). The green bar on top left shows the total number of events that took place in this reconstruction. For each inner node *u*, the floating number (e.g. 98.0) represents the bootstrap value of the tree, the numbers in the brackets are a 3-tuple showing the cumulative count of events going from the leafnodes to node *u* in the following order: [deletions, duplications, splits]. Each leaf node is accompanied with some symbols(*, ?, !), the accession number, the species name, and the gene block. An asterisk (*) indicates the gene block contains gene *cyp*115, an exclamation ! indicates the gene block contains a pseudo or fragment of gene *cyp*115, and ? indicates the gene block contains gene *ggps*2. The reference genome *Xanthomonas oryzae pv. oryzicola* is in blue, and the outgroup genome *Erwinia tracheiphila* is in gray. These naming and color conventions persist through this study.