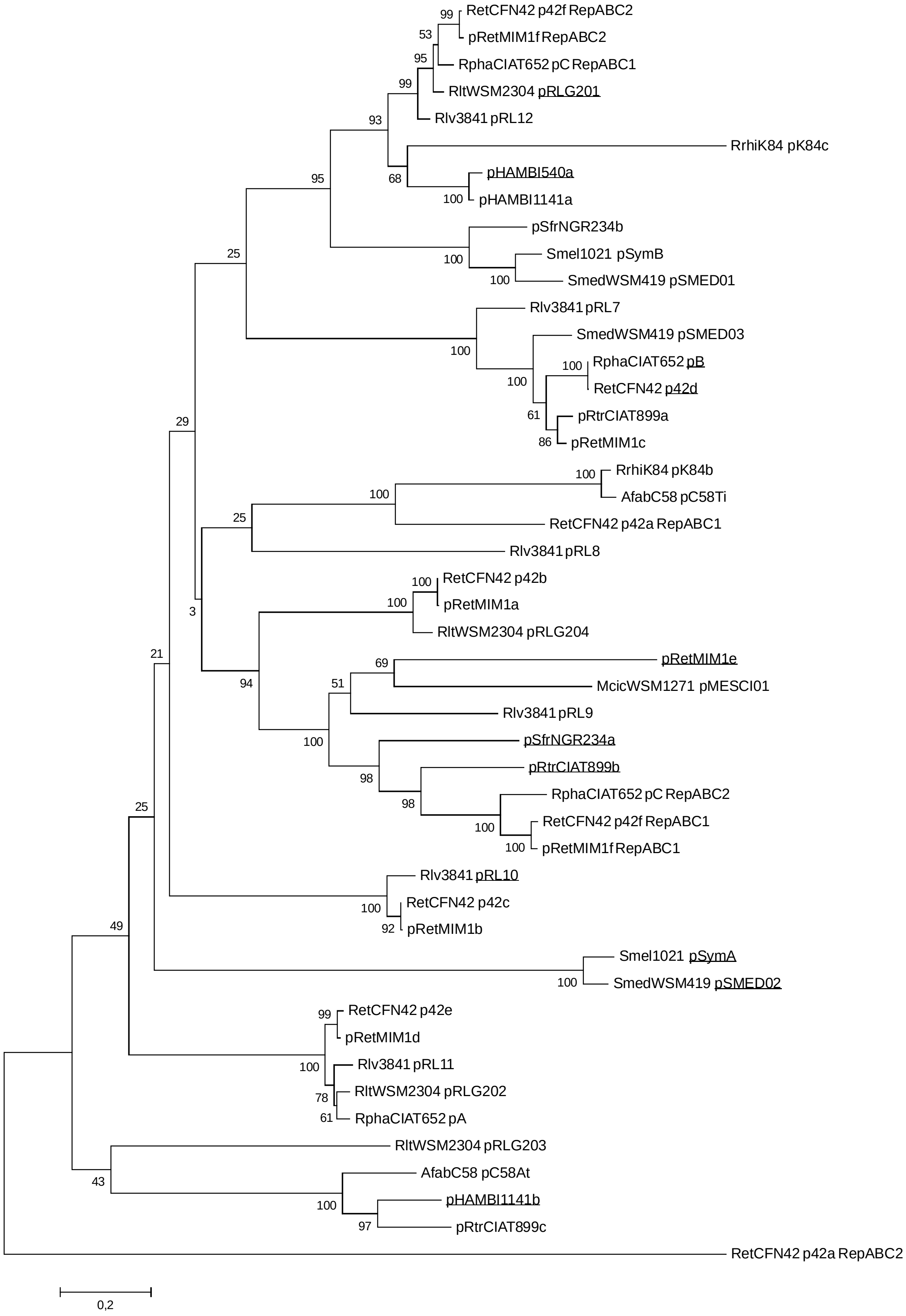

Figure S1. Phylogeny of concatenated RepABC proteins of replicons with this replication system in *N. galegae* strains HAMBI 540^{T} and HAMBI 1141, together with 12 strains representing *Rhizobium*, *Sinorhizobium*, *Mesorhizobium* and *Agrobacterium*. For those replicons which have multiple *repABC* operons, the operon is indicated. Replicons containing symbiosis genes are underlined. Rlv3841; *R. leguminosarum* sv. viciae 3841, RltWSM2304; *R. leguminosarum* sv. trifolii WSM2304, RtrCIAT899;*R. tropici* CIAT 899, RphaCIAT652; *R. phaseoli* CIAT 652, RetMIM1; *R. etli* sv. mimosae str. Mim1, RetCFN42; *R. etli* CFN 42, SfrNGR; *S. fredii* NGR234, Smel1021; *S. meliloti* 1021, SmedWSM419; *S. medicae* WSM419, McicWSM1271; *M. ciceri* sv. biserrulae WSM1271, RrhiK84; *R. rhizogenes* K84, AfabC58; *A. fabrum* C58. The evolutionary history was inferred by using the Maximum Likelihood method. Bootstrap values are shown next to the branches.

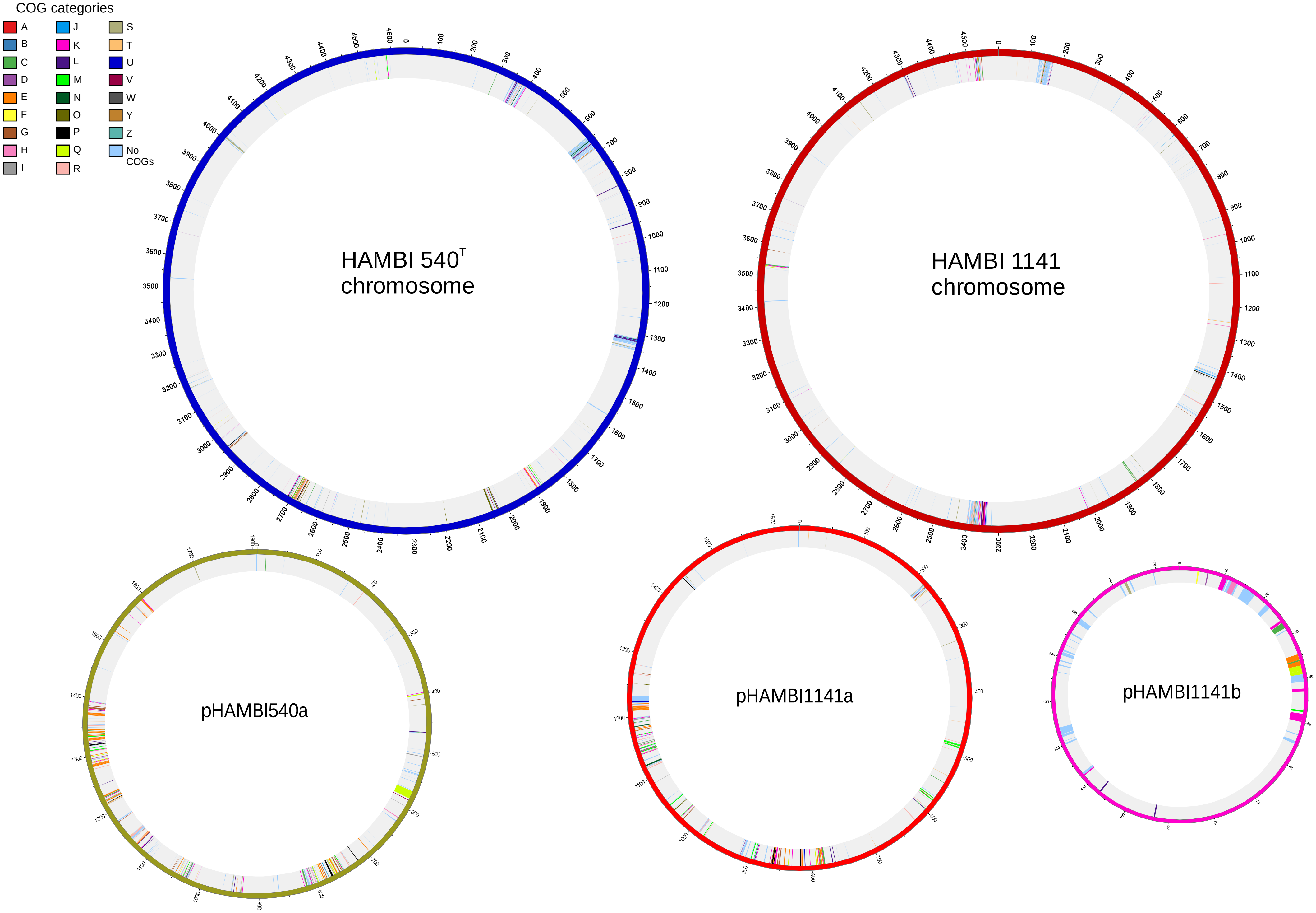

Figure S2. Illustration of the location singletons identified in the genomes of *N. galegae* HAMBI 540^T and HAMBI 1141 in the OrthoMCL analysis.

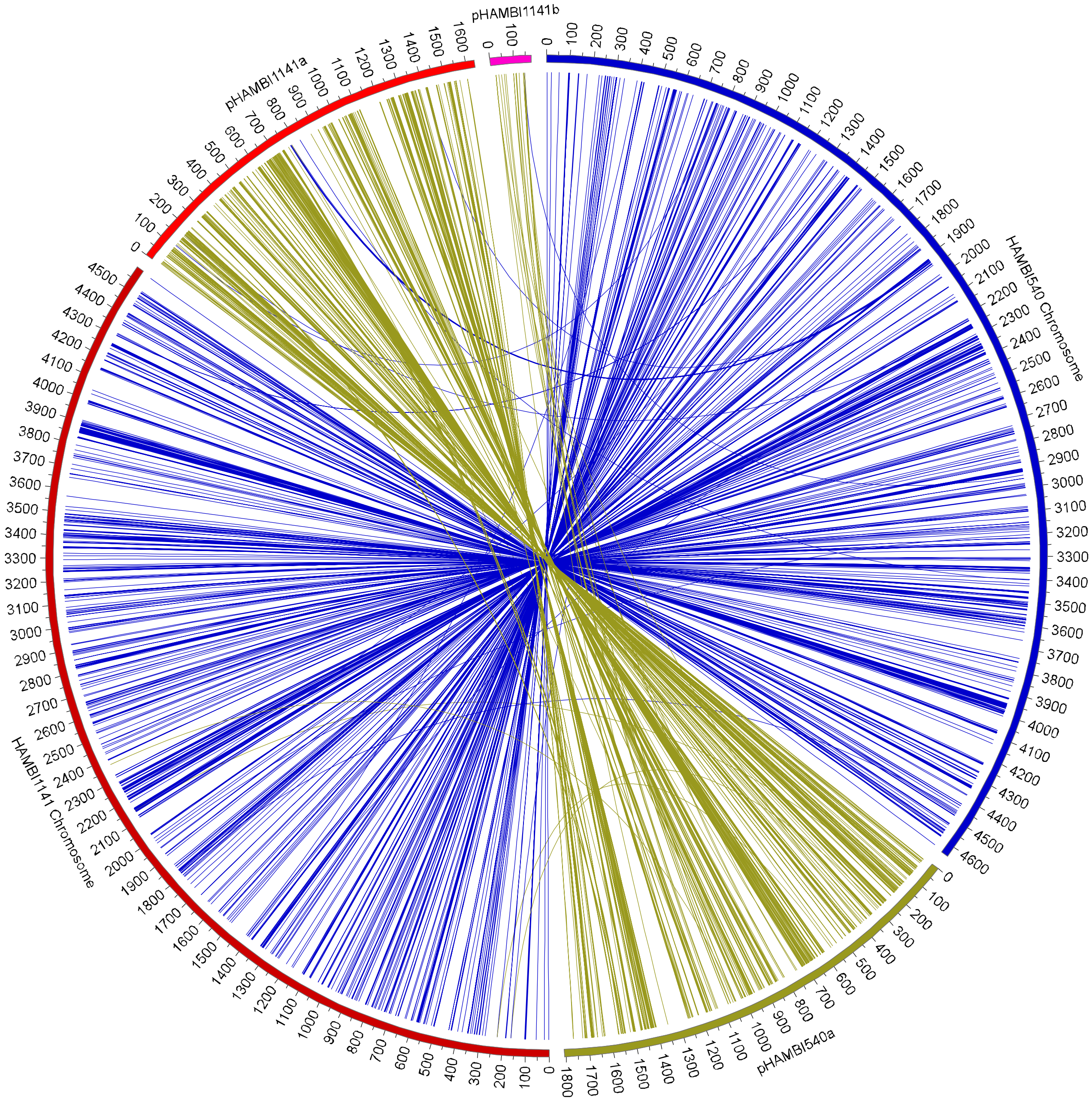

Figure S3. Orthologous genes of *N. galegae* strains HAMBI 540^T and HAMBI 1141.

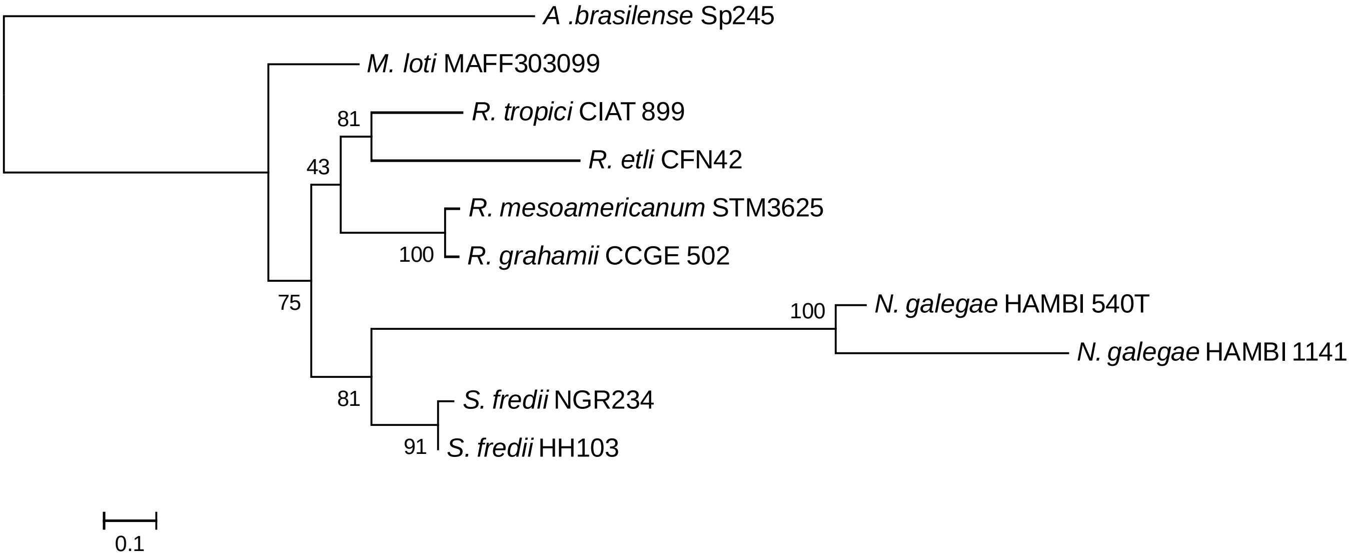

Figure S4. NifQ ML tree used for PAML analyses.

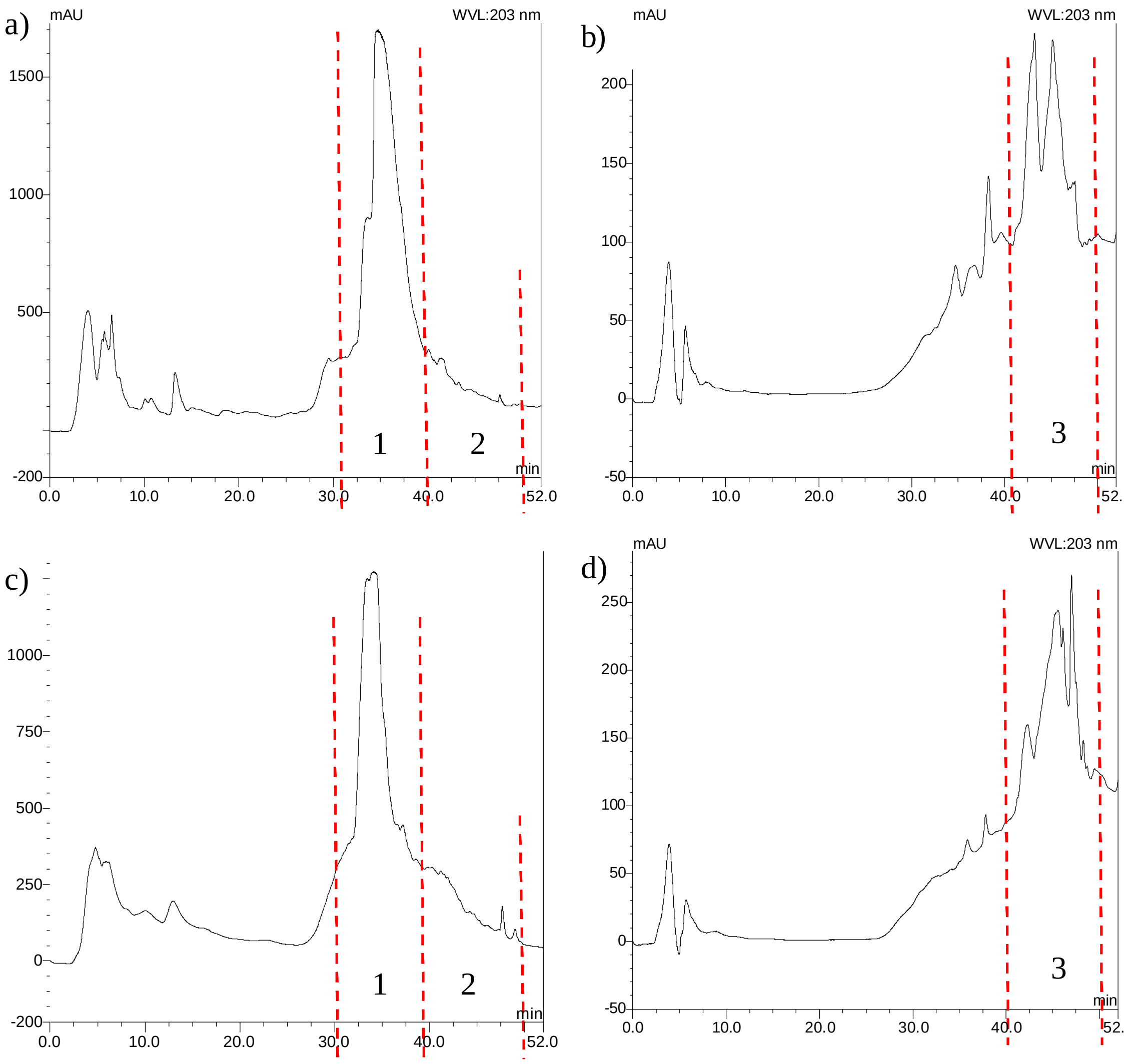

Figure S5. RP-HPLC fractionation of *N. galegae* HAMBI 1174 and *noeT* mutant strain LCO-containing SPE fractions. a) HAMBI 1174 (45 % SPE fraction) b) HAMBI 1174 (60 % SPE fraction), c) *noeT* mutant (45 % SPE fraction) d) *noeT* mutant (60 % SPE fraction). Regions 2 and 3 correspond to known elution positions for LCOs.

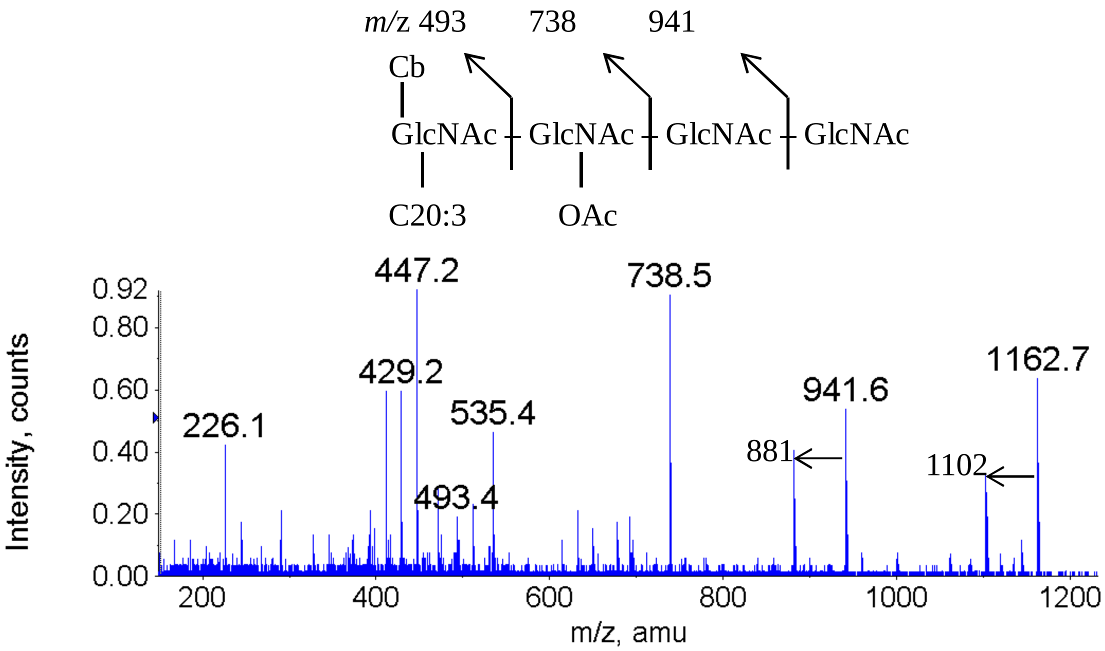

Figure S6. CID product ion spectrum of the LCO from wild type strain *N. galegae* HAMBI 1174

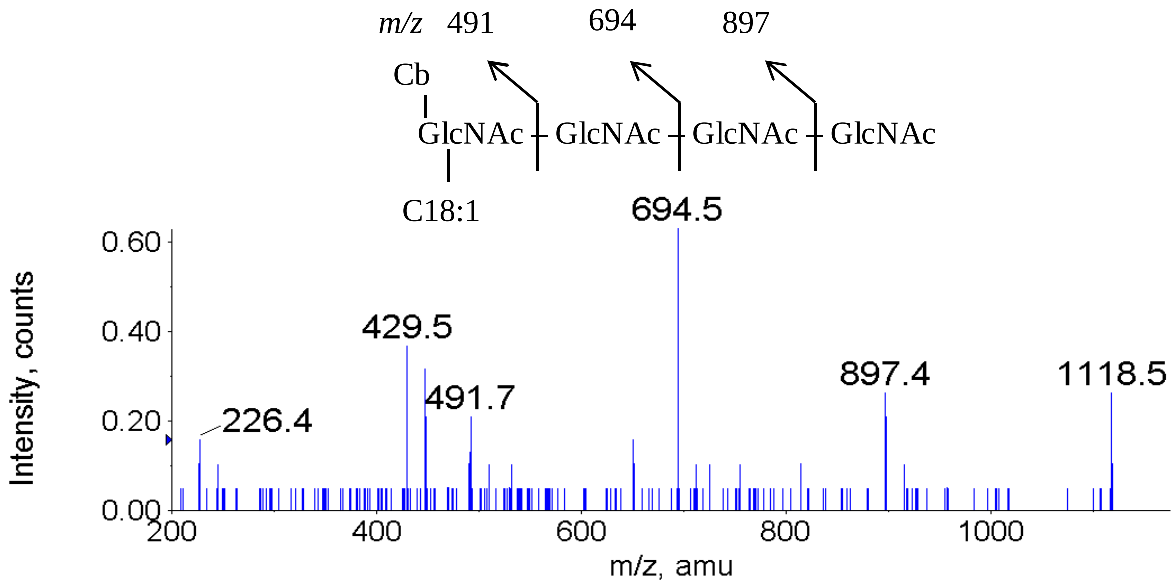

Figure S7. CID product ion spectrum of *N. galegae* HAMBI 1174 *noeT* mutant strain NF