Supplemental figure S2: Maximum likelihood tree of the CHASE domains from bacteria and eukaryotes. The sequence set was obtained by a PSI-BLAST search against Genbank's (http://www.ncbi.nlm.nih.gov/Genbank) non-redundant protein database using the CHASE domain of AHK4 as a query sequence. The search was iterated 14 times, until no further sequences were found. In the resulting dataset of 640 sequences, the presence of the CHASE domain was confirmed by an HMMsearch with the CHASE HMM from Pfam (PF03924.3). Sequences showing more than 95% sequence identity were also removed from the dataset. From the remaining 325 sequences, a multiple sequence alignment was built with Muscle and a phylogenetic tree constructed with PHYML, as described in the Materials and Methods section. The sequence identifiers are composed of the NCBI's gi number and the species name. The numbers next to branches represent support values from 100 bootstrap replicates. All plant sequences are highlighted in a green box.