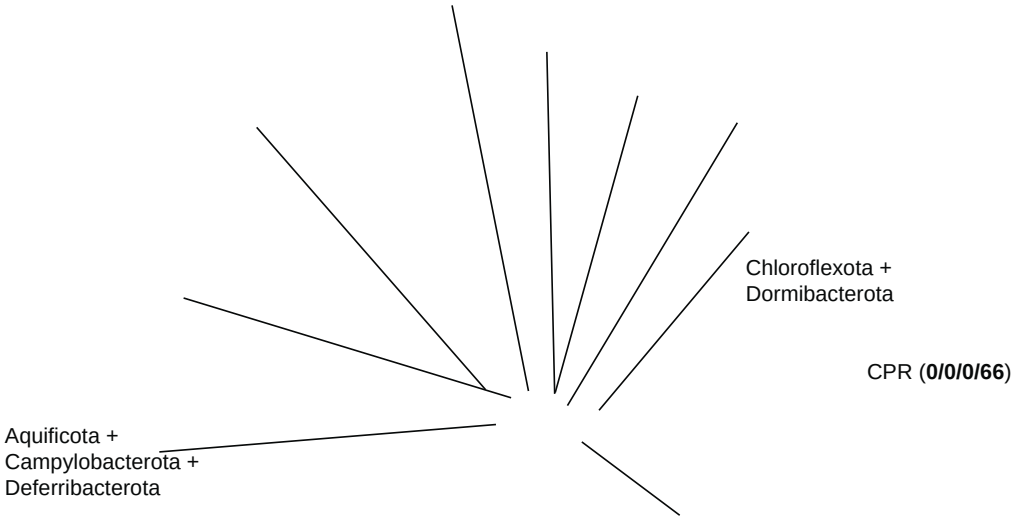
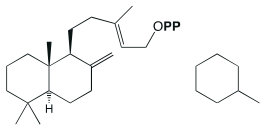
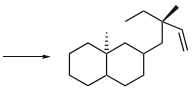
Received: 11 Jul 2024; Accepted: 30 Jul 2024; Published: 3 August 2024; Corrected and Typeset: 1 October 2024


CseDCS
ChjDCS
StrDCS
SpDCS
AcrDCS
DxDD
DDxxD / DDxxxD / DDxxxE
NDxxSxxxE / NGxxSxxxE







ent-

bacterial *DTC* and *TS* to undergo fusion and generate bifunctional DCSs is consistent with an origin for the ancestral *TPS* via acquisition of a *DCS* encoding a CPS–KS from bacteria by the last common ancestor of land plants. Notably, there are direct parallels to the cyanobacterial CPS–KS encoded by *CseDCS*, which was further shown here to extend to certain enzymatic structure–function relationships in common with the plant *TPSs* involved in phytohormone biosynthesis. While especially true for the CPS active site, the KS active site of CseDCS exhibits intriguing similarities to not only the KSs from plants but also those from bacterial GA biosynthesis and, hence, might be viewed as being intermediate between these. Although CseDCS provides only a single example of a bacterial bifunctional CPS–KS, the unknown function of such production of *ent*-kaurene in cyanobacteria, or even the selective advantage of fusing DTCs and subsequently acting TSs, argues against the expectation of more widespread conservation of such activity in bacteria.

Beyond supporting a bacterial origin for the initiating gene fusion event, the inclusion of the bacterial DCSs in phylogenetic analysis enabled the refinement of the evolutionary model for the plant *TPS* gene family. In particular, it has been hypothesized that the ancestral plant *TPS* underwent two gene duplication events, leading to the three extant groups of plant *TPSs*: *TPS*-c, *TPS*-e/f, and *TPS*-a/b/d/g/h subfamilies [11], here referred to as Clans I, II, and III (respectively). Clans I and II stem from subfunctionalization of the ancestral CPS–KS into monofunctional KS (Clan II) and, eventually, CPS (Clan I, which contains extant CPS–KSs), while Clan III stems from neo-functionalization into *TPSs* dedicated to more specialized metabolism, initially bifunctional, with the latter development of monofunctional (Class I) *TPSs* [11]. However, the order of these two gene duplication events could not be determined due to the lack of a proper outgroup, which is now provided with the discovery of bacterial DCSs. Notably, the bacterial

and the resulting fragments were then cloned back into pET28a. Mutants were constructed via whole-plasmid PCR. In all cases, the constructs were verified by whole-genome sequencing. Enzymatic activity was analyzed via a previously described modular metabolic engineering system [22]. Briefly, this enables the expression of the putative DCSs, as well as DTC-only derivatives (mutants and N-terminal γβ-didomains), with a GGPP synthase (GGPS), as well as subsequently acting Class I *TPS* (where indicated) or the TS-only derivatives (mutants and C-terminal α-domains) with the GGPS and stereospecific *TPS* that exhibit just DTC activity. The resulting recombinant *E. coli* were then grown, induced, and extracted for GC–MS analysis of product outcome, as recently described [11]. Although most of the observed products could be identified by comparison to authentic standards (see Supplementary Data), the StrDCS main product could not. To provide enough for structural analysis the relevant StrDCS product was extracted from 3 l of culture, isolated by flash chromatography, and purified by high-performance liquid chromatography (HPLC). This provided ∼2 mg, which was dissolved in 0.3 ml deuterated chloroform ($CDCl_3$) and transferred to an NMR microtube (Shigemi). Structural analysis was carried out by NMR using a Bruker AVANCE 700 MHz instrument equipped with a 5-mm HCN cryogenic probe, as previously described [36].

A

25. Aaron JA, Christianson DW. Trinuclear metal clusters in catalysis by terpenoid synthases. *Pure Appl Chem*. 2010;82: 1585–97
26. Jia M, Zhang Y, Siegel JB.