A functionally conserved STORR gene fusion in *Papaver* species that diverged 16.8 million years ago

Theresa Catania1,7, Yi Li1,7, Thilo Winzer1, David Harvey1, Fergus Meade1, Anna Caridi1, Andrew Leech2, Tony R. Larson2, Zemin Ning3, Jiyang Chang4, Yves Van de Peer4,5,6 & Ian A. Graham1

The STORR gene fusion event is considered essential for the evolution of the promorphinan/morphinan subclass of benzylisoquinoline alkaloids (BIAs) in opium poppy as the resulting bi-modular protein performs the isomerization of (S)- to (R)-reticuline essential for their biosynthesis. Here, we show that of the 12 *Papaver* species analysed those containing the STORR gene fusion also contain promorphinans/morphinans with one important exception. *P. californicum* encodes a functionally conserved STORR but does not produce promorphinans/morphinans. We also show that the gene fusion event occurred only once, between 16.8-24.1 million years ago before the separation of *P. californicum* from other Clade 2 *Papaver* species. The most abundant BIA in *P. californicum* is (R)-glaucine, a member of the aporphine subclass of BIAs, raising the possibility that STORR, once evolved, contributes to the biosynthesis of more than just the promorphinan/morphinan subclass of BIAs in the Papaveraceae.
The benzylisoquinoline alkaloids or BIAs represent a structurally diverse group predominantly identified in the order Ranunculales. The naturally synthesised morphinans thebaine, oripavine, codeine and morphine are part of the BIA class of alkaloids, with morphine renowned for its powerful analgesic properties. They are naturally synthesised in the genus *Papaver* and are currently commercially produced in opium poppy, *Papaver somniferum*, from the Papaveraceae family. The commercial importance of opium poppy has led to its use as a model species for research into the biosynthetic pathway for morphinan production. The common precursor and central branch point in the pathway for production of the many structurally distinct subclasses of BIAs in the Ranunculales including morphinan, protoberberine, phthalideisoquinoline and benzophenanthridine is the 1-benzylisoquinoline alkaloid, (S)-reticuline (Fig. 1a). The gateway reaction leading to morphinan biosynthesis is catalysed by the STORR protein. Composed of P450 and oxidoreductase modules, this fused protein completes the epimerization of (S)- to (R)-reticuline, the first step in the morphinan pathway. STORR is a member of the first four genes involved in synthesis of the first morphinan, thebaine and ten genes involved in synthesis of the phthalideisoquinoline, noscapine, which together make up the BIA gene cluster in opium poppy.

Advances in genome sequencing technology and assembly offers the opportunity to compare the genome organisation of related species and provide insight into the role of events such as gene fusion and gene clustering in the evolution of specialized metabolites. Here, we use such an approach together with transcriptomics, metabolomics and gene function analysis to determine the evolutionary sequence of events leading to the clustering of the genes encoding the STORR modules, the fusion of these genes to form a functional STORR and the clustering of the other four genes involved in morphinan production across a number of *Papaver* species.

**Results**

**Metabolite and transcriptomic analysis of *Papaver* species.** The epimerization of (S)- to (R)-reticuline is followed by a sequence of conversions from (R)-reticuline to salutaridine, salutaridinol, salutaridinol-7-O-acetate and thebaine catalysed by the products of SALSYN, SALR, SALAT and THS, respectively (Fig. 1a). To investigate the presence of the STORR gene fusion across the *Papaver* genus we selected nine other *Papaver* species that provided good taxonomic coverage (Table 1; Fig. 1b; Supplementary Table 1), and analysed these alongside the reported metabolite and whole genome assembly data of opium poppy, *P. rhoes* and *P. setigerum*.

To determine the presence of promorphinan and morphinan compounds and related genes in the 11 species metabolite profiling of latex from juvenile plants and capsule material post-

---

*Fig. 1 Metabolite and transcriptomic analysis to investigate STORR and its distribution in *Papaver* species.*

a. Schematic of the benzylisoquinoline pathway and enzymes in Opium Poppy with a focus on morphinan production. (S)-reticuline is the central branch point of BIA metabolism in opium poppy for the production of structurally distinct compounds including the morphinans noscapine, berberine and sanguinarine. Conversion of (S)- to (R)-reticuline by STORR represents the first committed step in biosynthesis of the promorphinans (salutaridine, salutaridinol and salutaridinol-7-O-acetate) and morphinans (thebaine, oripavine, codeine and morphine). Compound names are in black and enzymes specific to the promorphinan and morphinan pathway are in red.

b. Species phylogeny inferred by Bayesian inference species tree using eight single copy conserved ortholog sequences. Phylogeny and clustering of the genes encoding the STORR gene fusion across the *Papaver* species. Clades are indicated and the twelve *Papaver* species are placed into two different clades (Clade 1 and Clade 2) as described by Carolan et al. Species highlighted with an asterisk indicate all gene sequences were identified and retrieved from the corresponding annotated genomes whereas for the remaining nine *Papaver* species, transcriptomic datasets were used. Presence of the STORR and the four promorphinan genes are shown under species names (Supplementary Data 2). Source data are provided as a Source Data file.
Table 1 Metabolite analysis for promorphinan and morphinan compounds in 12 Papaver species*.

| Species          | Reticuline | Promorphinan | Morphinan |
|------------------|------------|--------------|-----------|
|                  | Salutaridine | salutaridinol | Salutaridinol-7-O-acetate | Thebaine | Oripavine | Codeine | Morphine |
| Papaver setigerum | ND         | ND           | ND         | +       | ND       | +       | +        |
| Papaver somniferum | ND         | ND           | ND         | +       | ND       | +       | +        |
| Papaver rhoes     | ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver dubium    | ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver armeniacum| ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver trinilfolium | ND     | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver bracteatum| ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver orientale | ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver atlanticum| ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver californicum| ND     | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver nudicaule | ND         | ND           | ND         | ND      | ND       | ND      | ND       |
| Papaver pavonium  | ND         | ND           | ND         | ND      | ND       | ND      | ND       |

*The presence of the promorphinan and morphinan compounds in latex and capsule samples was determined by high resolution accurate mass spectrometry (HRAM) (Supplementary Data 1). + sign represents species where the compounds were quantitatively measured at a level 10x above the calculated limits of detection (LOD) and the ND (not detectable) represents species where the compounds were below our LOD. LOD values were calculated using authentic standards. Promorphinan compounds were identified in four species indicative of a functional STORR, with morphinan production identified in two species plus P. somniferum and P. setigerum. Source data are provided as a Source Data file.

harvest was carried out (Table 1 and Supplementary Data 1). As previously reported, we found thebaine to be the most prominent metabolite in P. bracteatum and oripavine the most prominent in P. orientale19–21. These two species along with P. somniferum and P. setigerum14 were the only species identified as producing morphinans. A recent report of trace amounts of morphinans in P. rhoeas appears to have been conducted without the use of known morphinan standards13. Our analysis of P. rhoeas capsule material conducted using high resolution accurate mass LC-MS does not detect morphinans or promorphinans above defined limits of detection (Supplementary Data 1), which for morphine is 226-fold lower than morphine levels in P. setigerum. Our results of zero peak area for morphinans in P. rhoeas are in agreement with other published results17,22,23 and consistent with the absence of morphinan-related genes in this species.

From the other species in our analysis, promorphinans were identified as minor compounds in P. armeniacum and P. trinilfolium (Table 1 and Supplementary Data 1). Identification of these morphinan and promorphinan compounds is suggestive of the presence of a STORR ortholog which was selectively lost in some Papaver species.

Transcriptomic analysis revealed expression of STORR and either none, some or all of the promorphinan genes in a subset of the nine other Papaver species, with the full complement in opium poppy as reported previously5. These are shown on the branches of a species tree generated using conserved orthologs of 8 low-copy nuclear genes identified from the same transcriptomic datasets or the annotation of whole genome assemblies (Fig. 1b; Supplementary Tables 2 and 3, Supplementary Data 2 and 3).

The topology of the tree generated based on the species included is largely congruent with other phylogenetic trees constructed based on taxonomic sequence datasets of chloroplast, ribosomal, and plastid markers from the genus Papaver and the wider family and order24–28. The differences observed in the ordering of the Papaver species can be attributed to the marker sets and methods used for the assembly of the species tree and range of plant samples sequenced. The divergence times estimated between the species of the tree are in the regions of those previously estimated, for example the divergence of P. californicum at around 16.8 million years ago (MYA) compares to previously estimated timings27–30.

Papaver STORRs convert (S)- to (R)-reticuline in vitro. Similar to the opium poppy STORR, the homologues present in the other five Papaver species (Fig. 1b) encode full length P450-oxidoreductase fusion proteins (Supplementary Fig. 1). The direction of both modules and the 9-13 amino acid linker sequences are also conserved (Fig. 2a; Supplementary Fig. 1) with overall sequence identity ranging from 89.4-99.9% at the nucleotide level and 86.0-99.9% with the predicted amino acid sequences (Supplementary Data 4).

The STORR gene fusion from P. somniferum and P. bracteatum have previously been shown to catalyse the gateway reaction to promorphinan biosynthesis3,4,7. The identification of STORR in conjunction with promorphinans in P. armeniacum and P. orientale suggest a similar role in these species. In P. californicum, STORR was expressed but promorphinans and genes related to promorphinans were not detected. However, the main BIA identified from our extraction of P. californicum was a member of the aporphine subclass, glaucine in the (R) configuration (Supplementary Table 1 and Supplementary Fig. 2).

In order to investigate the function of the P. californicum STORR, we expressed this gene in Saccharomyces cerevisiae, performed epitmerization assays on microsomal fractions and found the same activity as the P. somniferum, P. bracteatum and P. armeniacum STORR proteins (Fig. 2b). Activity of P. californicum STORR microsomal fractions on alternative
Monophyletic origin of **STORR** in the genus *Papaver* 16.8 MYA. To understand the evolutionary relationship, we performed gene tree analysis of the **STORR**s and the closest paralogues to the two **STORR** modules within the genomes of a representative subset of *Papaver* species. The presence of a closely linked gene pair of closest paralogs to the opium poppy **STORR** cytochrome P450 CYP82Y2 and oxidoreductase modules was previously discovered through analysis of a whole-genome assembly5. The segmental duplication resulting in these paralogs was suggested to have occurred 20.0–27.8 MYA by the Ks estimation of the paralogous pairs 6. In order to establish if equivalent paralogous pairs are present in related *Papaver* species we used whole-genome sequencing approaches to assemble draft genomes for *P. nudicaule* from Clade 1 and *P. californicum, P. bracteatum, Por P. orientale* and *P. armeniacum* from Clade 2 (Supplementary Table 4). We compiled and annotated all homologous sequences that contain full-length genes corresponding to either of the **STORR** modules in these draft assemblies. We combined all CYP82Y2 and oxidoreductase sequences from the five draft genomes presented here with those retrieved from searches of the annotated opium poppy5,14, *P. rhoeas* and *P. setigerum*14 genomes, as well as the transcriptomic data of other species in the present study. We then constructed gene trees for the two gene subfamilies containing the coding sequences closely related to the **STORR** modules CYP82Y2 (Fig. 3a) and oxidoreductase (Fig. 3b), respectively. Consistent with previous work6, both trees revealed robustly supported cytochrome P450 CYP82Y2-like (CYP82Y2-L) and oxidoreductase-like (COR-L) groups. All P450 modules and oxidoreductase modules of **STORR**s are from Clade 2 *Papaver* species and they form monophyletic orthologous subgroups highlighted in orange (Fig. 3a, b). The **STORR** subgroups contain no other complete coding sequences for P450 or oxidoreductase apart from *Pca_CYP82Y2, L*storr, which actually contains the conserved linker sequence and the first 13 codons of an oxidoreductase before a stop codon, indicating deletion after the fusion. Therefore these two monophyletic **STORR** subgroups and the presence of functional **STORR** in *P. californicum* (Fig. 2b) lead us to conclude that all **STORR** genes from these *Papaver* species have derived from a single fusion event in the common ancestor of Clade 2 species after their divergence from Clade 1 between 16.8 to 24.1

---

**Fig. 2 Functional analysis of **STORR** proteins in *Papaver* species.** a Alignment of the amino acid linker connecting the **STORR** P450 and oxidoreductase modules. *P. setigerum* Pso, *P. somniferum* Par, *P. armeniacum* Pam, *P. bracteatum* Pbr, *P. orientale* Por and *P. californicum* Pca. b Functional characterisation of the **STORR** proteins by heterologous expression in *S. cerevisiae*. **STORR** protein from *P. californicum*, *P. armeniacum*, *P. bracteatum* and *P. somniferum* were separately produced with a *P. somniferum* cytochrome P450 reductase (CPR) redox partner in *S. cerevisiae*. Microsomal proteins were assayed with (S)-reticuline as substrate. Relative abundance is used to show the reticuline epimers present in the sample. Reticuline standards (black), pESC-trp empty vector (red), pESC-trp vector + cytochrome P450 reductase (CPR) (green), pESC-TRP + CPR + Pso **STORR** (blue), pESC-TRP + CPR + Pbr **STORR** (gold), pESC-TRP + CPR + Pam **STORR** (pink) and pESC-TRP + CPR + Pca **STORR** (cyan). Source data are provided as a Source Data file.
MYA (Fig. 3c). Our analysis of STORR gene evolution also reveals the importance of lineage-specific deletion (P. atlanticum and P. rhoesas), duplication (P. californicum, P. armeniacum, P. sommiferum and P. setigerum) and rearrangement after duplication (P. californicum) (Fig. 3c). In addition, we note that in P. sommiferum one copy of STORR has been lost after a whole-genome duplication event5.

We also observed Clade 2 orthologous subgroups (CYP82Y2-La and COR-La, highlighted in red), which contains the closest parologue pair (Pso_CYP82Y2_La4-Pso216860 and Pso_-COR_La4-Pso216870) from the opium poppy genome in the respective gene trees (Fig. 3a, b). This and the presence of the corresponding copies in P. californicum indicates all sequences within the subgroups are derived from a single copy Clade 2 ancestor as found to be the case for STORR. We found the same pairing arrangement of closest parologue pairs in all our assembled genomes except P. bracteatum which appears to have lost this paralogous gene pair (Fig. 3c). In P. californicum and P. atlanticum we found evidence for the ancestral arrangement of the paralogous gene pairs but in both cases one member of the pair now is present as a pseudogene (Fig. 3c). The order of the P450 and oxidoreductase genes for both paralogous pairs in P. armeniacum and P. rhoesas, as well as two of the three gene pairs in P. setigerum, have switched compared to the other species (Fig. 3c).

Taken into consideration gene deletion, duplication, rearrangement and erosion it is still apparent from these findings that paralogous pairing of a CYP82Y2-La and COR-La exist in the Clade 2 common ancestor (Fig. 3c), supporting the hypothesis that a segmental duplication would have occurred prior to STORR fusion/neofunctionalization.

Fig. 3 Evolutionary history of the formation of STORR fusion gene inferred from gene tree analyses of its P450 CYP82Y2 and oxidoreductase modules. a Phylogenetic gene tree of CYP82Y2-L P450 subfamily constructed using Bayesian Inference. b Phylogenetic gene tree of COR-L oxidoreductase subfamily constructed using Bayesian Inference. The posterior probabilities are shown on the branches and scale bar represents substitutions per nucleotide site. Both trees were rooted based on branch position of orthologous outgroup containing P. sommiferum sequences CYP82Y-PS0215450 for CYP82Y2-L tree and COR-L-PS0211450/PS01020430/PS01020440/PS01020450/PS01021890 for the COR-L tree as described in Li et al.6. STORR, La, Lb, Lc, Ld and Ln are defined as orthologous subgroups and their members are highlighted by coloured branches. A three letter prefix followed by an underscore is used as a species identifier for each gene; including Par P. armeniacum, Pat P. atlanticum, Pbr P. bracteatum, Pca P. californicum, Pnu P. nudicaule, Por P. orientale, Prh P. rhoesas, Ptr P. trinifolium, Pse P. setigerum and Pso P. sommiferum. c Schematic representation of the evolutionary history of STORR reconciling the species tree with gene trees. The fusion event was preceded by the clustering of a gene pair of CYP82Y2-La and COR-La genes and subsequent segmental duplications. The grey background branches denote species divergence with speciation time points indicated by red arrows. The orange lines denote STORR and the red lines a gene pair of CYP82Y2-La and COR-La. The grey line indicates an ancestral CYP82Y2-L and COR-La gene pair prior to the divergence of P. nudicaule from the Clade 2 Papaver species. The exclusive presence of Clade 2 species in La and STORR subgroups and a single P. nudicaule Ln subgroup in both CYP82Y2-L and COR-L trees is consistent with the segmental duplication occurring after the divergence of P. nudicaule from Clade 2 species but before STORR formation between 24.1 and 16.8 MYA. Source data are provided as a Source Data file.
Within both CYP82Y2-L and COR-L groups, the gene trees have identified single orthologous groups (CYP82Y2-Ln and COR-Ln, highlighted in grey) containing exclusive Clade 1 P. nudicaule sequences, which are absent from all other subgroups (Fig. 3a, b). Among the members, there is at least one paralogue pair (Pnu_CYP82Y2_Ln3/Pnu_COR_Ln3) just over 1 kb apart (Fig. 3c), which was confirmed by direct sequencing of a PCR amplified genome fragment. This finding implies that the origin of the CYP82Y2-L and COR-L pairing event occurred before the separation of P. nudicaule from Clade 2 Papaver species 24.1 MYA. This is consistent with the presence of additional paralogous pairs identified in the Clade 2 Lb and Lc subgroups which will have arisen by multiple segmental duplication events (Fig. 3a, b). One such duplication event will have led to the formation of STORR 16.8-24.1 MYA.

**STORR clustered with four promorphinan genes 9.7 MYA.** As part of the 800 kb BIA cluster, STORR is clustered with four promorphinan genes in the opium poppy genome, which also contains a second syntenic region containing paralogs of the promorphinan genes but not STORR (Fig. 4a). Both these syntenic regions are well conserved in P. setigerum, the closest related sister species of opium poppy as inferred from the species phylogeny in the genomes. The phylogeny for the selected subset of Papaver species and divergent times at the branching points were extracted from the species tree in Fig. 1b. The emergence of STORR and its clustering with all four promorphinan genes are indicated on the ancestral nodes. These are derived from the presence and organisation of these genes in the genomes of these extant species, which are also shown. Source data are provided as a Source Data file.
the opium poppy regions containing the promorphinan genes with a notable difference being an extra copy of SALR in the former (Fig. 4a; Supplementary Data 5). *P. armeniacum* shows synteny with STORR and SALSYN with gene order conserved in two smaller contigs, one of which extends to reveal five additional genes in the flanking region. Only STORR was found in the draft assembly of *P. californicum*, the *P. atlanticum* draft assembly lacks STORR and the promorphinan genes (Supplementary Data 2). These findings are consistent with our metabolite profiling and transcriptomic analyses across these five species (Table 1; Fig. 1b), highlighting the role played by gene deletion events in formation of the diversity of BIA composition within the Papaver lineage.

Together these findings lead us to conclude that the STORR plus promorphinan component of the BIA gene cluster formed prior to the divergence of *P. bracteatum* and *P. somniferum* 9.7 MYA (Fig. 4b), but after the formation of STORR at least 16.8 MYA. Orthologues of SALSYN, SALAT, SALR and THIS have only been identified in the genomes of a subset of the Clade 2 Papaver species that produce promorphnin/morphinan alkaloids and contain the STORR gene fusion (Supplementary Data 2). Percentage nucleotide and amino acid identity matrix of these orthologues range from 94.1-99.8% and 90.5-100% respectively (Supplementary Data 4), supporting the case for evolution of the promorphinan/morphinan biosynthetic pathway being triggered after the STORR gene fusion event.

**Discussion**

We demonstrate a functional STORR protein in *P. californicum*, which has persisted since its divergence from the common ancestor 16.8 MYA. *P. californicum* represents the earliest branching species in the Papaver lineage that we have found to contain the STORR gene fusion event occurring between 24.1 – 16.8 MYA (Fig. 1b). While we are confident that the STORR gene fusion occurred before the neofunctionalization of promorphinan genes we cannot conclude whether the latter was before or after the branching of *P. californicum* from the other Clade 2 species as shown in Fig. 1. Therefore it is possible that either neofunctionalization did not happen prior to the branching of *P. californicum* or it did happen and the promorphinan genes were lost.

Genome rearrangements result in considerable structural variation even within a single species as evidenced by the loss of the noscapine component of the BIA gene cluster9 and significant variation in copy number of genes associated with morphinan biosynthesis in *P. somniferum*8. That STORR persists in *P. californicum* suggests it is providing some selective advantage.

*P. californicum* is a *Papaver* species indigenous to California (‘New World’) thought to have evolved in parallel to the Eurasian ‘Old World’ members of the family24,27,29,30. The specific distribution of *P. californicum* to North Western America is an example of an ‘Old World/New World’ disjunction at 28 – 10 MYA in the Papaveraceae distribution30. The occurrence of such North American/Eurasian disjunctions is recognised in a number of species, and attributed to historic changes in climate and existence of previous land-bridge connections30-33. The differences observed in the metabolite profile of *P. californicum* compared to the morphinan-producing species could be attributed to its parallel evolution in a different environment with different selective pressures.

We found the most abundant BIA in *P. californicum* to be (R)-glaucine, which to our knowledge has not previously been reported in nature. Glaucine in the (S) configuration isolated from *Glaucium* species in the Papaveraceae is associated with bronchodilator, anti-inflammatory and neuroleptic effects34,35. A chemically synthesised form of (R)-glaucine has been shown to increase the efficacy of serotonin36. Evidence for (S)-glaucine biosynthesis via (S)-reticuline has previously been proposed37. It is therefore interesting to speculate, based on the emergence of STORR at 16.8 MYA in the *Papaver* lineage, that the *P. californicum* STORR could be involved in (R)-glaucine formation. Epimerisation assays using *P. californicum* STORR with (S)-glaucine and (S)-laudanosine as substrate showed no activity implying that neither of these are intermediates in the biosynthesis of (R)-glaucine but rather the pathway is dependent on epimerization of reticuline by STORR followed by formation of the tetracyclic aporphine structure.

A recent report based on genome comparison of opium poppy with two very closely related *Papaver* species, *P. setigerum* and *P. rhoes* has proposed the fusion event that resulted in STORR occurred following the whole-genome duplication event in opium poppy after its divergence from *P. rhoes*14. However, that proposal does not consider the possibility of STORR gene loss within *P. rhoes* nor does it take into account the previously reported presence of STORR in *P. bracteatum* which diverged earlier from all three species (Fig. 1b)24. Furthermore, that proposal associates the segmental duplication giving rise to the CYP82Y2-La/COR-La gene pair and STORR with the opium poppy WGD event but no supporting evidence was presented14. Figure 3 demonstrates that the required gene duplication event must have happened before this WGD event.

Our findings with wider taxonomic sampling clearly show that the STORR gene fusion was a single event that occurred between 16.8-24.1 MYA in the *Papaver* lineage and was preceded by clustering and segmental duplication of the P450 oxidase and oxidoreductase genes that fused to form it. While the STORR gene fusion is regarded as a key event enabling the evolution of the morphinan subclass of BIA’s our findings show that it may also have enabled the production of compounds in other BIA subclasses such as the aporphine (R)-glaucine in *P. californicum*, which 16.8 MYA branched from the common ancestor of those other *Papaver* species that now produce morphinans.

**Methods**

**Plant material.** The plant material used in the current study were voucher specimens sourced primarily from botanic institutions as detailed in Supplementary Table 1. Plants were grown under controlled long day conditions in the glasshouse facilities and in the experimental gardens located in the University of York. Samples for DNA cDNA and RNA were collected from young leaves of juvenile plants and flash frozen in liquid nitrogen. Metabolite analysis was conducted on latex harvested from three individual glasshouse grown plants at 70 days post germination and from capsules pre-dehiscence.

**Genomic DNA and cDNA isolation for PCR.** Genomic DNA was extracted from frozen ground young leaf using the BioSprint 15 Plant Kit on the BioSprint 15 Workstation (Qiagen, Crawley, UK). The DNA was quantified on the nanodrop 1000 (Thermo Fisher). cDNA was synthesised from RNA isolated from young leaf tissue using superscript V (Thermo Fisher) and used for gene-specific amplification.

**Metabolite profiling.** Metabolite extraction and UPLC analysis were carried out on both latex and dried capsule material. Fresh latex samples were collected from the stems of juvenile plants into 10% acetic acid for extraction. Following flowering capsule material was also collected, dried and ground. A 10 mg sample of material from the dried capsules was then extracted in 10% acetic acid. The extracts were analysed using an Acquity UPLC system (Waters Ltd., Elstree, UK) linked to a Thermo LTQ Orbitrap (Thermo Fisher, Hemel Hempstead, UK). Authentic standards for salutaridine and salutaridinol (Toronto research chemicals, Canada), and also a MCONT standard mix (morphine, codeine, oripavine noscapine and thebaine) were included to confirm the presence or absence of promorphinan and morphinan compounds in the 10 *Papaver* species. Limit of detection for all standards were calculated by preparing each individual standard to a top concentration of 0.05 mg/ml in 10% Acetic acid equating to 100 ng on column in a 2 pl injection. From each of the top standards a 20 point serial dilution was prepared. The limit of detection (LOD) calculation was based on the standard deviation of the response (S) of the curve and the slope of the calibration curve (S) at a level approximating the LOD (according to the formula: LOD = 3.3(Sy/S) Supplementary Data 1)38.
Isolation of glaucine by preparative HPLC. A total of 20 g of dried capsule material from *P. californicum* or *G. flavum* plants was ground to a fine powder and extracted with 250 ml of 80% acetic acid in 70% methanol, followed by centrifugation at 4000 g for 10 minutes and the extract dried down using an EZ-2 Elite genevac. The dried residue was taken up in 2 ml of ethyl acetate and further spun to remove debris prior to purification. Isolation and purification of glaucine was performed using the interchum pulirrh 4500 prep HPLC system with Advion Expression, compact mass spec (CMS). The ethyl acetate extracts were applied to a 12 g BUCHI FlashPure Econflex silica 50 µm irregular column and all fractions collected using a 0-100% ethyl acetate in hexane gradient, followed by isocratic 100% ethyl acetate and 100% methanol. Samples from fractions thought to contain the glaucine peak (356 ion) were confirmed by running on an Acquity UPLC system (Waters, Elstree, UK) linked to a Thermo LTQ Orbitrap (Thermo Fisher, Hemel Hempstead, UK) on an Acquity BEH C18 1.7 µm 2.1 x 100mm column with the mass spec using APCI ionization in positive polarity. The glaucine containing fractions were dried down and resuspended in 10% acetic acid for further purification on a 4 g TELOS Flash C18 column with fractions collected using a 2-80% gradient of Solvent B in Solvent A (where Solvent A is 10 mM Ammonium bicarbonate pH10.2; Solvent B is 100% methanol). This method yielded 3 mg of glaucine from *P. californicum* and 2 mgs from *G. flavum*. These were resuspended in 100% ethanol to a 2mg/ml final concentration for circular dichroism (CD) analysis.

Analysis of glaucine extracts by circular dichroism. CD spectra were collected on a Jasco J-1500 Circular Dichroism Spectrometer. The system was purged with oxygen-free nitrogen for 10 minutes before lamp ignition and the lamp allowed to stabilize for at least 10 minutes before data acquisition. A nitrogen purge flow of at least 5 l/min was employed during data collection. Temperature was regulated by a Peltier thermostat monitored at the cell holder at 20 °C.

CD spectra were collected from 400 nm down to 190 nm which was the limit of solvent absorption allowing a HIT maximum of 600 V; scan speed was 50 nm/min; bandwidth was averaged and absorbance was measured in a path length quartz cuvettes at the specified concentration. Blank spectra were collected in the same cuvette as the sample using the same solvent, and subtracted from the sample spectra. Data analysis was performed with Jasco Spectro Manager v2 software. UV absorption spectra of the samples were collected on a Jasco V560 spectrophotometer in the same cuvettes over the range 400 to 210 nm.

**RNA-Seq sequencing and transcriptomic data analyses.** Transcriptomic analysis was performed on RNA extracted using the Direct-zol RNA Miniprep Kit (Zymo Research, USA) according to the manufacturer’s instructions from young frozen leaf material. RNA was quantified on the Qubit 3.0 Fluorometer (Thermo Fisher Scientific) according to the manufacturer’s protocol and quality assessed by running 1 µl on the agilent tapestain.

RNA sequencing was performed on the nine *Papaver* species (Supplementary Table 2). RNA samples were prepared from 1 µg high-quality RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module and NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs), according to the manufacturer’s guidelines. Libraries were subject to 150 base pair end sequencing on one lane of a HiSeq 3000 system at the University of Leeds Next Generation Sequencing Facility (Leeds, UK) except for *P. californicum* which was sequenced on a NovaSeq 6000 at Novogene Co, Cambridge, UK. An average of 7.5 Gb PE reads sequencing data were generated for each species (Supplementary Table 2).

Transcriptomes were assembled with the Trinity (v2.2.0) software pipelines after filtering out any of the remaining 1-5% ribosomal RNA in the raw reads for each species with mapping to rRNA_115_tax_silva_v1.0 downloaded from SILVA database (https://www.arb-silva.de/). These transcriptome datasets were used for identification of orthologous genes of selected BIA biosynthetic genes (Supplementary Table 3) and the eight conserved orthology sets (COS) genes (Supplementary Data 3) by local BLAST searches. Sequences of the top matches were retrieved. Orthologous gene sequences belonging to gene families were identified and verified through subsequent gene tree analysis along with the related gene family’s datasets that were previously reported (Supplementary Data 2).

**Species phylogeny and estimation of divergence times.** Estimations of divergence dates among the nine *Papaver* species together with *P. rhoes, P. setigerum* and opium poppy were based on the 8 COS gene sequences from the transcriptomic datasets (Supplementary Data 3) along with other Papaveraceae species *E. californica*, *M. cordata* and *M. armeniacum*. DNA samples were used in PCR amplification using the BIA primers listed in Table S3.

Orthologous sequences of the eight COS genes were identified and retrieved from the annotated gene datasets of *P. somniferum*, *E. californica*, *M. cordata* and *M. armeniacum* after conducting BLAST searches. Multiple sequence alignments of each gene set were obtained with MUSCLE v3.8.31,32. Conserved alignment blocks were generated with Gblocks v0.91b30 to remove highly polymorphic regions and the alignments of all eight genes were subsequently concatenated in MegaX48.

Species divergence times were estimated under the strict clock model implemented in BEAST v2.5.1 to generate a Bayesian posterior sample of time-calibrated phylogenies and the remaining species were included using a tree two using two prior calibration points (Ranunculales 110 ± 5 MYA and Papaveraceae 77 ± 4 MYA). Priors were treated as setting a Yule speciation process and lognormal distribution; the nucleotide substitution model used was GTR with 4 Gamma categories. The chain length of Markov chain Monte Carlo (MCMC) was set to 30,000,000 runs, per million generations and collected every 100th generation; with 10% of the total trees discarded as burn-in samples, the remaining trees were used for generating the consensus tree. Convergence of the runs performed by BEAST was assessed by effective sample sizes (ESS) using Tracer v1.7.4. ESS for the parameter statistics fell above the threshold of 200 that is considered to indicate good sample quality. The species tree and divergence times were visualised using Figtree v1.4.3.

**Whole genome sequencing and assemblies.** 10x Genomics whole genome sequencing and assemblies were performed on five *Papaver* species, including *P. nudicaule, P. californicum, P. atlanticum, P. bracteatum* and *P. armeniacum*. Further Oxford Nanopore Technology (ONT) sequencing and genome assemblies were carried out on two of these species *P. bracteatum* and *P. armeniacum*.

High molecular weight (HMW) genomic DNA was prepared for 10X Genomics sequencing for five species. Young seedling material was grown and sent to Amplicon Express (Pullman, WA, USA), where HMW genomic DNA was prepared by using their proprietary protocol for HMW grade (megabase size) DNA preparation. This protocol involves isolation of plant nuclei and yields pure HMW DNA with >100 kb fragment length. The DNA samples were then used to construct the 10X Chromium libraries, which were subsequently sequenced on Illumina NovaSeq platform to produce 2 × 150 bp reads at HudsonAlpha Institute for Biotechnology, Huntsville, Alabama. This produced a total of 2 billion reads of >100 kb fragment length for each species. Samples were prepared with the linked reads at approximately 50-60 times coverage of the estimated genome size, as required by the Supernova assembly software package (Supplementary Table 4) for optimal performance.

Further sequencing of *P. bracteatum* and *P. armeniacum* was carried out using ONT long read sequencing technology by the Platform Technology Facility at the University of York. HMW DNA samples were prepared from frozen young leaf tissue using a CTAB extraction method followed by a Qian genomic tip clean-up and removal of short DNA fragments using the Circulomix SRE kits (PacBio)32. Long read sequencing was performed by the University of York Bioscience Technology Facility Genomics lab, using the Oxford Nanopore Technologies MinION and ONT Oxford Nanopore Technologies MinION sequencer setup on the ONT long read sequencing kit SQK-LSK106, using a minimum of 5 µg high-quality DNA. The resulting library was split into 4 to allow loading onto MinIon (FLO-MINION) and
PromethION (FLO_PR00002) R9.4.1. flow cells, and sequencing for 48 hours (MiniION) or 72 hours (PromethION), with a flow cell wash (using ONT wash kit EXP-Wash kit) and re-loading of the flow cells for 24 hours into the run. Basecalling was performed using ONT’s Guppy toolkit, version 4.0.11. Just over 7 million reads were generated with average read length of 13,395 bp and 95 Gb in total for \textit{P. bracteatum}, whereas 4.3 million reads with average length of 22,216 bp covering 99 Gb (votide bases for \textit{P. armeniacum}. We ran the FLYE version 2.8-b16743 de novo assembly pipeline with the ONT datasets, producing initial assemblies. This was followed by one round base polishing with the starting ONT raw reads using RACON (https://github.com/isovic/racon). Purge_dups42 was then used to remove haplotigs and contig overlaps to produce a unique, non-redundant representation of the genome. Final draft genome assemblies (Supplementary Table 4) were achieved after two more rounds of base polishing using FREEBAYES software too.55 After the Illumina short reads from the 10X Chromium libraries were mapped to the working assembly with Longranger software package (https://support.10xgenomics.com/genome-exome/software/pipelines/latest/what-is-long-ranger). The protein-coding genes of selected BIA biosynthetic genes (Supplementary Table 3) were conducted firstly by local BLAST searches in the 10X assemblies of \textit{P. nudicaule}, \textit{P. californicum}, and \textit{P. atlanticum}, the final draft ONT assemblies of \textit{P. bracteatum} and \textit{P. armeniacum} and the annotated gene sets of \textit{P. rhoes} and \textit{P. setigerum}. Sequences of the top matches were retrieved and verified through subsequent gene tree analyses along with the related gene families. The region containing full length coding regions were previously reported.6 (Supplementary Data 2).

**Repeat annotation and gene prediction.** For each of the five assembled \textit{Papaver} genomes, a repeats library was constructed ab initio using RepeatModeler (v2.0.1, http://www.repeatmasker.org/RepeatModeler)(66). The consensus TE sequences generated by the RepeatModeler software were then used as repeats library in RepeatMasker (v4.1.1, http://www.repeatmasker.org/RepeatMasker) to identify repetitive elements in all five genomes. In addition, we identified intact LTR-RTs using LTR_retriever (v2.8)\textsuperscript{37}, which integrates results of LTR_FINDER (v1.1)\textsuperscript{58} and LTR_harvest (v1.5.9)\textsuperscript{59}. Insertion times of LTRs were also calculated by LTR_retriever with parameter -u set to 1.396-8. An integrative approach combining homology-based, RNA-seq-based, and ab initio gene prediction was used to identify duplicates and protein-coding genes in the \textit{Papaver} genomes. Genome sequences and gff files of \textit{Arabidopsis thaliana} and two previously published \textit{Papaver} genomes, \textit{P. somniferum} and \textit{P. rhoes}, were used for homology-based prediction using GeneMoMa\textsuperscript{v1.8}, http://www.jspacs.de/index.php/GeneMoMa) with default parameters. Illumina RNA-seq data from the white frost leaf and brown frost leaf were used for transscriptome assemblies were mapped to the genome using HISAT2\textsuperscript{25} (v2.1.0) and the alignments were assembled with StringTie\textsuperscript{262} (v2.1.5). Transdecoder (v5.0.2, https://github.com/TransDecoder) was then used to find potential open reading frames (ORFs). For ab initio predictions, BRAKER\textsuperscript{263} was used and model training was based on RNA-seq data after the predicted repeats were soft-masked within the genome assembly. Finally, all structural gene annotations were joined with InterProScan version 5.52-86 with default parameters. BUSCO\textsuperscript{6} analyses were run on the finalised annotations of the five \textit{Papaver} genomes using BUSCO version 4.1.4 and unмоubhypha_odb10 used to estimate genome completion\textsuperscript{65}. The transcriptome and genome assembly data generated in this study have been deposited in the NCBI databases under the BioProject PRJNA770869. Raw RNA seq and DNA reads are available under the accession numbers SRR16389878–SRR16389886, SRR16690173, SRR16690174 and SRR16591806–SRR16591806. Assembled RNA-seq datasets are available under the following accession numbers GJOO00000000–GJOY00000000. Amplified region to the contigs/scaffolds in the five \textit{Papaver} genomes. The contigs/scaffolds were identified after BLAST searches for containing \textit{STORR} or \textit{SALSYN}/\textit{SALAT}/\textit{SAL/SYS} in the draft genome assemblies and their sequences were retrieved (Supplementary Data 4). A single contig was identified in \textit{P. bracteatum} genome and its length is 1.2 Mb, whereas the single scaffold in \textit{P. californicum} is the shortest with 44 kb in length. Two contigs were constructed by full length coding regions for \textit{STORR}/\textit{SALSYN} genes, one is 294 kb long and the other 94 kb in length. Ab initio gene prediction of these contigs and scaffolds was carried out using FGENesh, a web-based gene annotation tool.\textsuperscript{57} (http://www.softberry.com/berry.html?topic=genes&group=programs&subgroup=fgnd) with Dicot plants option as training set. The predicted genes were functionally annotated by homologous BLAST searches in the swissprot database. Reciprocal BLASTN searches were performed between the coding sequence of predicted genes in these contigs/scaffolds and the annotated opium gene sets. The matching regions to the contigs/scaffolds in the five \textit{Papaver} genomes. The top match records in both sets with an expected value less than 1E-12 corresponding to the opium positive \textit{BIA} gene cluster region were summarised in Supplementary Data 5 and Fig. 4a.

**Data availability**

The transcript and genome assembly data generated in this study have been deposited in the NCBI databases under the BioProject PRJNA770869. Raw RNA seq and DNA reads are available under the accession numbers SRR16389878–SRR16389886, SRR16690173, SRR16690174 and SRR16591806–SRR16591806. Assembled RNA-seq datasets are available under the following accession numbers GJOO00000000–GJOY00000000. Amplified region to the contigs/scaffolds in the five \textit{Papaver} genomes. The contigs/scaffolds were identified after BLAST searches for containing \textit{STORR} or \textit{SALSYN}/\textit{SALAT}/\textit{SAL/SYS} in the draft genome assemblies and their sequences were retrieved (Supplementary Data 4). A single contig was identified in \textit{P. bracteatum} genome and its length is 1.2 Mb, whereas the single scaffold in \textit{P. californicum} is the shortest with 44 kb in length. Two contigs were constructed by full length coding regions for \textit{STORR}/\textit{SALSYN} genes, one is 294 kb long and the other 94 kb in length. Ab initio gene prediction of these contigs and scaffolds was carried out using FGENesh, a web-based gene annotation tool.\textsuperscript{57} (http://www.softberry.com/berry.html?topic=genes&group=programs&subgroup=fgnd) with Dicot plants option as training set. The predicted genes were functionally annotated by homologous BLAST searches in the swissprot database. Reciprocal BLASTN searches were performed between the coding sequence of predicted genes in these contigs/scaffolds and the annotated opium gene sets. The matching regions to the contigs/scaffolds in the five \textit{Papaver} genomes. The top match records in both sets with an expected value less than 1E-12 corresponding to the opium positive \textit{BIA} gene cluster region were summarised in Supplementary Data 5 and Fig. 4a.

**Repeating summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Received: 22 October 2021; Accepted: 23 May 2022; Published online: 07 June 2022

**References**

1. Ziegler, J. et al. Evolution of morphine biosynthesis in opium poppy. \textit{Phytochemistry} \textbf{70}, 1696–1707 (2009).

2. Hajel, J., M. & Facchini, P. J. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. \textit{Plant Cell Physiol.} \textbf{54}, 647–672 (2013).

3. Winzer, T. et al. Morphinan biosynthesis in opium poppy requires a P450- oxidoreductase fusion protein. \textit{Science} \textbf{349}, 309–312 (2015).

4. Farrow, S. C., Hajel, J. M., Beaudoin, G. A., Burns, D. C. & Facchini, P. J. Stereocchemical inversion of (S)-reticuline by a cytochrome P450 fusion in anopine poppy. \textit{Nat. Chem. Biol.} \textbf{11}, 728–732 (2015).

5. Guo, L. et al. The opium poppy genome and morphinan production. \textit{Science} \textbf{362}, 343–347 (2018).

6. Li, Y., Winzer, T., He, Z. & Graham, I. A. Over 100 million years of enzyme evolution underpinning the production of morphine in the Papaveraceae family of flowering plants. \textit{Plant Commun.} \textbf{1}, 100029 (2020).
10. Nützmann, H. W. & Osbourn, A. Gene clustering in plant specialized metabolism. *Plant Physiol.* 154, 531–535 (2010).

9. Winzer, T. et al. A Papaver somniferum 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* 336, 1704–1708 (2012).

8. Olszewski, A. Gene clustering and copy number variation in alkaloid metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535 (2010).

7. Galanis, S., Thodey, K., Trenchard, I. J., Interrante, M. F. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* 349, 1095–1100 (2015).

6. Oberti, R., A. Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535 (2010).

5. Zetler, G. Neuroleptic-like, anticonvulsant and antinociceptive effects of aporphine alkaloids: bufotenine, bufotenine, buprenaline, and bufotenine. *Arch. Int. Pharmacodyn. Ther.* 296, 255–281 (1988).

4. Cortijo, J. V. et al. Bronchodilator and anti-inflammatory activities of glaucine: In vitro studies in human airway smooth muscle and polymorphonuclear leukocytes. *Br. J. Pharmacol.* 127, 1641–1651 (1999).

3. Celander, T., Leitner, K. & Kutsch, T. M. Molecular characterization of the salutaridin 7-O-acetyltransferase involved in morphin biosynthesis in opium poppy. *Plant J.* 267, 3071–30723 (2001).

2. Chen, X. et al. A pathogenesis-related 10 protein catalyzes the final step in thebaine biosynthesis. *Nat. Chem. Biol.* 14, 738–743 (2018).

1. Suzuki, Y. & Brunn, J. G. Papaver bracteatum – a summary of current knowledge. *Plant Syst. Evol.* 357, 97–117 (1979).

A. Yang, X. et al. Three chromosome-scale *Papaver* genomes revealed punctuated patchwork evolution of the morphinoph and noscapine biosynthesis pathway. *Nat. Commun.* 12, 6030 (2021).

Gesell, A. et al. CYP719B1 is salutaridine synthase, the CC phenol-coupling enzyme of morphine biosynthesis in opium poppy. *J. Biol. Chem.* 284, 24432–24442 (2009).

Ziegler, J. et al. Comparative transcript and alkaloid profiling in *Papaver* species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. *Plant J.* 48, 177–192 (2006).

Gadde, T. E. & Kutsch, T. M. The alkaloid biosynthetic pathway of *Papaver somniferum* differs from *Papaver bracteatum*. *Nat. Commun.* 11, 1–13 (2020).

Olszewski, A. et al. Gene clustering and copy number variation in alkaloid metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535 (2010).

Winzer, T. et al. A Papaver somniferum 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* 336, 1704–1708 (2012).

Nützmann, H. W. & Osbourn, A. Gene clustering in plant specialized metabolism. *Plant Physiol.* 154, 531–535 (2010).

Hagel, J. M. & Facchin, P. J. Tying the knot: occurrence and possible significance of gene fusions in plant metabolism and beyond. *J. Exp. Bot.* 68, 4029–4043 (2017).

Liu, Z. et al. Formation and diversification of a paradigm biosynthetic gene cluster in plants. *Nat. Commun.* 11, 1–11 (2020).

Li, Q. et al. Gene clustering and copy number variation in alkaloid metabolic pathways of opium poppy. *Nat. Commun.* 11, 1–13 (2020).

Yang, X. et al. Three chromosome-scale *Papaver* genomes revealed punctuated patchwork evolution of the morphinoph and noscapine biosynthesis pathway. *Nat. Commun.* 12, 6030 (2021).

Gesell, A. et al. CYP719B1 is salutaridine synthase, the CC phenol-coupling enzyme of morphine biosynthesis in opium poppy. *J. Biol. Chem.* 284, 24432–24442 (2009).

Ziegler, J. et al. Comparative transcript and alkaloid profiling in *Papaver* species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. *Plant J.* 48, 177–192 (2006).

Gadde, T. E. & Kutsch, T. M. The alkaloid biosynthetic pathway of *Papaver somniferum* differs from *Papaver bracteatum*. *Nat. Commun.* 11, 1–13 (2020).

Olszewski, A. et al. Gene clustering and copy number variation in alkaloid metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535 (2010).
Acknowledgements
I.A.G. received support from the Biotechnology and Biological Sciences Research Council, United Kingdom (grant BB/K018809/1) and the Garfield Weston Foundation, United Kingdom. We are grateful to York Technology Facility, COEMS for support with metabolite profiling and RNA-seq library preparation, the University of York High Performance Computing service, Viking and the Research Computing team for support with bioinformatics and J. Mitchell for administrative support. Y.V.d.P. acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (No. 833522) and from Ghent University (Methusalem funding, BOF.MET.2021.0005.01).

Author contributions
T.C. and F.M. conducted all RNA and genomic DNA preparation, molecular biology, biochemical and enzyme analysis, T.C., F.M., D.H. and T.R.L. conducted all metabolite analysis, T.C. and Y.L. all transcriptomic analysis, Y.L. and A.C. species phylogeny analyses, Y.L. all gene tree analyses, Y.L. all draft genome assemblies and Y.L. all synteny analysis, Y.C. and Y.V.d.P. performed all genome annotations. A.L. conducted circular dichroism analysis of glaucine extracts. T.C., Y.L., T.W. and I.A.G. analysed and interpreted results and were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-30856-w.

Correspondence and requests for materials should be addressed to Ian A. Graham.

Peer review information Nature Communications thanks Amit Rai, Michael Wink and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022, corrected publication 2022