



**HAL**  
open science

## Narrowing down the single homoeologous FaPFRU locus controlling flowering in cultivated octoploid strawberry using a selective mapping strategy.

Justine Perrotte, Amélia Gaston, Aline Potier, Aurélie Petit, Christophe Rothan, Beatrice Denoyes

### ► To cite this version:

Justine Perrotte, Amélia Gaston, Aline Potier, Aurélie Petit, Christophe Rothan, et al.. Narrowing down the single homoeologous FaPFRU locus controlling flowering in cultivated octoploid strawberry using a selective mapping strategy.. *Plant Biotechnology Journal*, 2016, 14 (11), pp.2176-2189. 10.1111/pbi.12574 . hal-02637036

**HAL Id: hal-02637036**

**<https://hal.inrae.fr/hal-02637036v1>**

Submitted on 22 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Narrowing down the single homoeologous *FaPFRU* locus controlling flowering in cultivated octoploid strawberry using a selective mapping strategy

Justine Perrotte<sup>1,2,3,‡</sup>, Amèlia Gaston<sup>1,2,†,‡</sup>, Aline Potier<sup>1,2</sup>, Aurélie Petit<sup>3</sup>, Christophe Rothan<sup>1,2</sup> and Béatrice Denoyes<sup>1,2,\*</sup>

<sup>1</sup>INRA, UMR 1332 BFP, Villenave d'Ornon, France

<sup>2</sup>Université de Bordeaux, UMR 1332 BFP, Villenave d'Ornon, France

<sup>3</sup>Cref Création Variétale Fraises Fruits Rouges, Douville, France

Received 21 February 2016;

revised 23 April 2016;

accepted 2 May 2016.

\*Correspondence (Tel +33 557122460;

fax +33 557122786; email

beatrice.denoyes@bordeaux.inra.fr)

†Present address: CNRS, UMR 5200,

Laboratoire de Biogenèse Membranaire,

Université de Bordeaux, Bordeaux, France.

‡These authors contributed equally to this work.

## Summary

Extending the period of fruit production is a way to substantially increase crop yield in many fruit or ornamental species. In the cultivated octoploid strawberry (*Fragaria* × *ananassa*), the most consumed small fruit worldwide, fruit production season can be extended by selecting the perpetual flowering (PF) cultivars. This trait is of considerable interest to growers and to the food industry. Four homoeologous loci controlling a single trait can be expected in such a complex octoploid species. However, we recently showed that the PF trait is under the control of the single dominant *FaPFRU* locus (*J. Exp. Bot.*, 2013, **64**, 1837), making it potentially amenable to marker-assisted selection (MAS). Here, we report the successful use of a strategy, based on a selective mapping using a reduced sample of individuals, to identify nine markers in close linkage to the *FaPFRU* allelic variant. Thus, this strategy can be used to fine map the target homoeologous loci in other complex polyploid crop species. Recombinant analysis further enabled us to reduce the locus to a region flanked by two markers, Bx083\_206 and Bx215\_131, corresponding to a 1.1 Mb region in the diploid *F. vesca* reference genome. This region comprises 234 genes, including 15 flowering associated genes. Among these, the *FLOWERING LOCUS T (FT)* is known to be a key activator of flowering. The close association between the PF trait and the *FaPFRU* flanking markers was validated using an additional segregating population and genetic resources. This study lays the foundation for effective and rapid breeding of PF strawberry cultivars by MAS.

**Keywords:** *Fragaria*, polyploidy, perpetual flowering, bin mapping, marker-assisted selection.

## Introduction

The length of the flowering period is a key trait in the development of high-yielding fruit or ornamental crops, because it determines flower number and, consequently, reproductive success. This trait depends on floral initiation, which is controlled by both environmental and genetic factors. Commonly for polycarpic perennial plants, floral initiation happens once a year ('seasonal flowering', SF). However, some perennials have the ability to flower more than once, thereby prolonging a favourable season, through a process called 'remontancy' or 'perpetual flowering' (PF) (Albani and Coupland, 2010; de Camacaro *et al.*, 2002). The promise of enhancing fruit production by increasing the flowering period makes this trait a key objective of numerous breeding programmes.

Perpetual flowering control has recently been described for rose and woodland diploid strawberry (*Fragaria vesca*) where, respectively, transposon insertion and mutation lead to the inactivation of a homologue of the floral repressor *TFL1* (Iwata *et al.*, 2012; Koskela *et al.*, 2012). For the cultivated octoploid strawberry, the PF trait has been shown to be controlled by the major QTL *FaPFRU* locus, which has been mapped to the female linkage group IVb (LGIVb-f) (Castro *et al.*, 2015; Gaston *et al.*, 2013). The allelic

variant is dominant and confers the PF behaviour to the plant with a negative effect on the production of stolons, which are specialized branch crowns producing new plants (the so-called runner process) (Savini *et al.*, 2008). A fine mapping of this region has been impeded by the octoploid status of the cultivated strawberry, and thus, it remains a challenging task.

Polyploidy resulting from whole-genome duplications has occurred in many flowering plants (Hegarty and Hiscock, 2008; Jiao *et al.*, 2011), including major crops, such as wheat, rape, sugarcane and cultivated strawberry. This process has long been recognized as a major force in evolution (Ohno, 1970). Polyploidization is usually followed by a process of diploidization (Bowers *et al.*, 2003; Paterson *et al.*, 2004), whereby gene redundancy is reduced via gene silencing, sequence elimination or rearrangement, demethylation of retroelements and relaxation of imprinting (see review in Soltis *et al.*, 2015). Although our understanding of the consequences of polyploidization has considerably increased in the past decades (Conant *et al.*, 2014), high ploidy level has undoubtedly delayed the development of genomic tools for certain crop species. Consequently, few genomes of polyploid species have been sequenced (allotetraploid rape: Chalhoub *et al.*, 2014; hexaploid wheat: Mayer

Please cite this article as: Perrotte, J., Gaston, A., Potier, A., Petit, A., Rothan, C., Denoyes, B. (2016) Narrowing down the single homoeologous *FaPFRU* locus controlling flowering in cultivated octoploid strawberry using a selective mapping strategy. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12574

*et al.*, 2014; allotetraploid cotton: Zhang *et al.*, 2015). The virtual 'reference genome' composed of 211 588 sequences was the first attempt to provide the sequence of the cultivated octoploid strawberry (Hirakawa *et al.*, 2014). These sequences, which cover 70.3% of the total length of the *F. vesca* genome (v1.1), highlight the high level of macrosynteny within *Fragaria* genomes. This high level of macrosynteny and of colinearity between the *Fragaria* genomes was previously shown through the comparative mapping approaches (Isobe *et al.*, 2013; Rousseau-Gueutin *et al.*, 2008; Sargent *et al.*, 2009). However, rearrangements and duplications can be found within specific regions (van Dijk *et al.*, 2014; Goldberg *et al.*, 2010; Sargent *et al.*, 2009; Spigler *et al.*, 2010). In addition, the cultivated strawberry displays a major disomic behaviour at meiosis (Lerceteau-Köhler *et al.*, 2003; Tennessen *et al.*, 2014). The diploid reference genome provides a reliable tool for fine mapping and positional cloning in the cultivated octoploid strawberry (Lerceteau-Köhler *et al.*, 2012; Rousseau-Gueutin *et al.*, 2008; Weebadde *et al.*, 2008).

Here, we present a strategy based on selective mapping and further refined with QTL identification and fine mapping that has successfully pinpointed the single *FaPFRU* locus controlling the PF trait in a genetically complex octoploid species of strawberry. Consequently, markers in close linkage to the *FaPFRU* locus have been identified. Moreover, our ability to transfer markers to a new segregating population and to 17 PF accessions indicates a single origin of the PF trait in modern strawberry cultivars. These findings will lead to the efficient and rapid breeding of strawberry PF cultivars by the marker-assisted selection (MAS). In addition, we have shown that 15 flowering associated genes are present within the refined locus. Among these genes is the *FLOWERING LOCUS T (FT)* gene that is known to be a key activator of flowering.

## Results

### The strategy to fine map a specific locus in the genetically complex polyploid species *Fragaria × ananassa*

In the cultivated octoploid strawberry, each locus is represented eight times in the genome that is eight homoeologous alleles can be found. When a major locus shows a 1:1 segregation in test cross, only one out of the eight homoeologues carries the allele of interest. Efforts to fine map a single specific allele at the exclusion of the remaining seven is very challenging. To overcome these limitations, we have developed a strategy that combines several complementary approaches (Figure 1). This strategy uses a previously described strawberry framework map (Lerceteau-Köhler *et al.*, 2012) that facilitates the identification of a reduced number of individuals for selective mapping (Vision *et al.*, 2000). New markers are then developed in the genomic region of interest, exploiting the colinearity between the genomes of the octoploid and diploid strawberries (Shulaev *et al.*, 2011). The markers are subsequently used to redefine the boundaries for the locus of interest and to identify recombinants in larger populations that segregate for the trait. When applied to a single homoeologous locus, this strategy leads to reduction in and refinement of the locus.

### Reducing the sample size of individuals optimizes the mapping of simple sequence repeat (SSR) markers at the *FaPFRU* locus

The parents and the 213 progeny comprising the reference segregating population resulting from the cross between

'Capitola' and 'CF1116' (Table 1) (Lerceteau-Köhler *et al.*, 2003; Rousseau-Gueutin *et al.*, 2008) were the source of data, underlying the mapping framework (Figure 1, Step 1). The reduced number of individuals was chosen according to the distribution of their crossover sites of recombination, also referred to as breakpoints (Vision *et al.*, 2000). The breakpoints define bins, which are intervals along a linkage group (LG) lacking a recombinational crossover site for a given set of individuals and that are bounded by breakpoints in at least one individual or by the end of a LG (Vision *et al.*, 2000) (Figure 1, Step2). Because the *FaPFRU* locus is located at the bottom of LGIVb-f (Gaston *et al.*, 2013), we optimized the representation of the breakpoints to the bottom of the homoeologous group (HG) IV in the female map HGIV-f, which includes the four LGs, LGIVa-f, LGIVb-f, LGIVc-f and LGIVd-f. The final sample of individuals included the parents of the segregating population and six individuals, 'Capitola' × 'CF1116' 35, 62, 70, 75, 118 and 128. Selection was performed using the MapPop program (Vision *et al.*, 2000), which was further refined by visual inspection. The objective was to introduce recombination between the two markers located at the LGIVb-f ends, *gatt284* and *gata295*, to link the two sub-linkage groups (Figure 2).

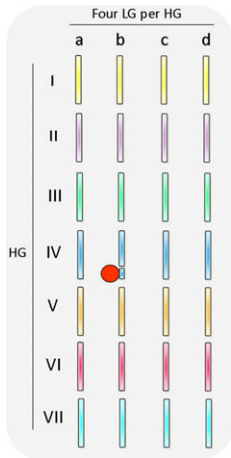
Within this reduced sample, the four female LGs belonging to HGIV, LGIVa-f, LGIVb-f, LGIVc-f and LGIVd-f were divided into three, four, five and three bins, respectively, yielding a total of 15 bins. Bin construction resulted in an average length of 11.2 cM per bin (Figure 2 and Table S1). The longest bin identified was BIN-FIVc-3 that spanned 35.1 cM. Furthermore, this reduced sample produced a male map divided into 11 bins (Figure 2 and Table S1) that was used to rapidly map a large number of SSR markers.

### Development of SSR primers for the fine mapping of the *FaPFRU* locus

For developing new markers, we focused on the genomic region of the diploid reference genome upstream of the UDF008 marker (Gaston *et al.*, 2013) (Figure 1, Step3). This region is orthologous to the *FaPFRU* locus (Figure 3a) and corresponds to the large scaffold scf0513158\_v1.0 (4 198 144 bp) (Shulaev *et al.*, 2011; GDR, <http://www.rosaceae.org>). Based on this scaffold sequence, a total of 275 novel, non-redundant SSR primer pairs were developed leading to an average spacing of one SSR per every 14 000 Kb of sequence. In addition, 16 SSR primer pair markers previously reported for the HGIV locus and mapped more precisely onto the scaffold scf0513158\_v1.0 were also included (Isobe *et al.*, 2013).

A total of 291 SSR primer pairs used in this study were tested on the reduced sample. Among them, 111 (38%) of the SSR primer pairs failed either to generate a PCR product or did not yield a clear pattern on polyacrylamide gel. The remaining 180 SSR primer pairs generated between one and 15 amplicons with an average of 4.6 products per primer pair. After excluding the monomorphic SSRs, an average of 2.4 alleles was obtained from the polymorphic primer pairs. Actually, if sequences of the primer-binding sites are conserved in all sub-genomes, eight homoeologous alleles are expected. The actual number (2.4) is lower, indicating either partial homozygosity or mismatching at primer-binding sites resulting from sequence divergence in several homoeologous regions.

Out of the 180 SSR primer pairs, 154 SSR primer pairs produced at least one polymorphic band, of which 227 bands were present exclusively in one parent. Each of these 227 bands,

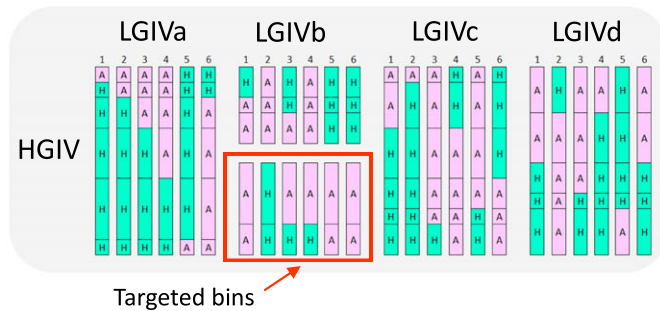


**1 A framework linkage map already available**

In the octoploid cultivated strawberry, a framework linkage map was achieved prior to the current work. This linkage map is organised in seven Homoeology Groups (HG) ranged from I to VII. Each HG includes four homoeologous Linkage Groups (LG) representing each the information in coupling and repulsion phase of the two homologous LG. The trait-of-interest (here on the LGIVb) is represented by a red dot.

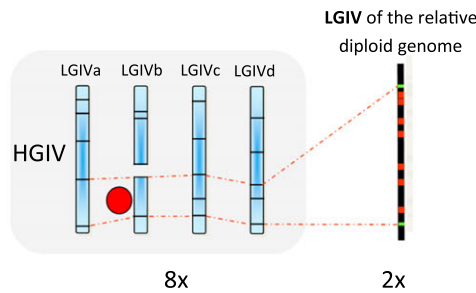
**2 Choice of a reduced sample of individuals for selective mapping on bins**

A reduced sample of individuals is chosen according to the distribution of their recombinational crossover sites in the mapping data framework. This choice leads to the identification of bins, which consist of intervals along LG where groups of markers display the same patterns.



**3 Designing new primers**

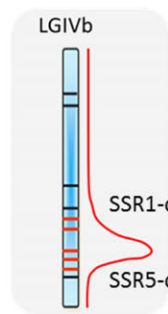
Owing to the high level of colinearity between octoploid and diploid linkage groups, a large number of SSR primers can be designed in the target region



New markers are mapped on the different bins using the reduced sample of individuals.

**4 New linkage map**

Markers presenting the pattern similar to the ones of the targeted bins are mapped on the entire octoploid segregating population.



**5 QTL identification**

After identification of QTL performed using this new linkage map, two new markers framing the QTL are chosen.

**6 Identification of recombinants**

Using the two markers that frame the QTL, recombinants identified among an enlarged population show one of the marker (AH or HA).

SSR1-qt1	A	H
SSR5-qt1	H	A

**7 Reduced QTL locus delimited by flanking markers**

Recombinants are phenotyped (X or Y) and genotyped (presence H or absence A of markers) using all new markers localised in the QTL. Combining phenotyping and genotyping information leads to reduce the locus and to identify two flanking markers useful for further studies, such as marker assisted selection and candidate gene identification.

		LGIVb											
Recombinant pattern		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12
SSR1-qt1		A	A	A	A	A	A	H	H	H	H	H	H
SSR2		H	A	A	A	A	A	H	H	H	H	H	A
SSR3		H	H	A	A	A	A	H	H	H	H	A	A
Flanking_mk1		H	H	H	A	A	A	H	H	H	A	A	A
<b>Phenotype</b>		<b>X</b>	<b>X</b>	<b>X</b>	<b>X/Y</b>	<b>Y</b>	<b>Y</b>	<b>X</b>	<b>X</b>	<b>X/Y</b>	<b>Y</b>	<b>Y</b>	<b>Y</b>
Flanking_mk2		H	H	H	H	A	A	H	H	A	A	A	A
SSR4		H	H	H	H	H	A	H	A	A	A	A	A
SSR5-qt1		H	H	H	H	H	H	A	A	A	A	A	A

Figure 1 The overall strategy in seven steps based on selective mapping using a reduced sample of individuals and combined with QTL identification and fine mapping.

**Table 1** Populations studied for their recombination in the *FaPFRU* locus. Recombinants, number and percentage, between the flanking markers Bx089\_196 and Bx064\_216 were identified for each population

	No. of seedlings	No. recombinants		Total recombinants	Recombinant (%)
		Bx089_196– Bx064_216			
		0–1	1–0		
Populations from cross					
'Capitola' × 'CF1116'	372	20	9	29	7.8
'Capitola' × 'Pajaro'	212	5	5	10	4.7
'Mara des Bois' × 'Pajaro'	85	6	2	8	9.4
Total in additive configuration	669	31	16	47	7.0
Populations from selfing					
S <sub>1</sub> populations					
'Capitola'-S <sub>1</sub>	432	18	17	35	8.1
'Mara des Bois'-S <sub>1</sub>	210	7	13	20	9.5
F <sub>2</sub> populations					
['Mara des Bois' × 'Pajaro']F <sub>1</sub> -no. 3	56	1	0	1	1.8
['Mara des Bois' × 'Pajaro']F <sub>1</sub> -no. 45	84	0	1	1	1.2
Total in co-dominant configuration	782	26	31	57	7.3
Total	1451	57	47	104	7.2

'Capitola' × 'CF1116' (210 individuals out of the 372) and 'Mara des Bois' × 'Pajaro' were the two mapping populations.

considered hereafter as SSR markers, was mapped to one unique bin in either the female (80 SSR markers) or the male (103 SSR markers) parent, whereas the location of the remaining 44 SSR markers could not be determined (Table 2).

All markers showing a similar pattern for the reduced sample were grouped, and their bin assignment was determined. Among the 80 new markers placed into the female bins, 16, 16, 19 and 29 were respectively attributed to LGIVa-f, LGIVb-f, LGIVc-f and LGIVd-f (Table 2). Because two of the bins BIN-FIVb-3 and BIN-FIVb-4 overlapped with the *FaPFRU* locus, the eight SSRs assigned to these two bins (Bx052, Bx056, Bx059, Bx064, Bx083, Bx089, Bx215, Bx250) were further evaluated by mapping them onto the octoploid reference mapping population 'Capitola' × 'CF1116' (Figure 3; Figure 1, Step4).

The result of mapping these nine markers was in concordance with the bin results, because the six markers of BIN-FIVb-3 were located between the markers *catc139* and *gatt284*, and the three markers of BIN-FIVb-4 were located between markers *gatt284* and *gata295* (Figure 3b). Upon the addition of the markers of BIN-FIVb-3, LGIVb-f formed a unique linkage group instead of two distinct ones as previously reported by Gaston *et al.* (2013) (Figure 3a). LGIVb-f was 125.1 cM long, and the nine markers added to the bottom of this linkage group added an additional 18.7 cM (Figure 3b). Therefore, the use of the bins allowed the drastic reduction in the number of markers to be mapped onto our octoploid reference segregating population and achieved our objective of saturating the *FaPFRU* locus-containing region with genetic markers.

#### Conservation of the *FaPFRU* locus in an independent octoploid population segregating for the PF trait

To determine whether the nine markers are also linked to the PF trait in a genetic background different from 'Capitola', we subsequently developed, genotyped and phenotyped a new segregating population from a cross between the PF 'Mara des

Bois' and SF 'Pajaro'. Primer pairs Bx089, Bx083, Bx056, Bx059, Bx064 and Bx215 resulted in the amplification of nine polymorphic markers that were mapped onto LGIVb-f (Figure 3c). Bx052 was monomorphic and Bx250 did not produce an amplicon and so were not mapped. Colinearity of the mapped markers indicated the conservation of this region in both PF cultivars 'Capitola' and 'Mara des Bois' and suggests that the same introgression event occurred at the *FaPFRU* locus (Figure 3).

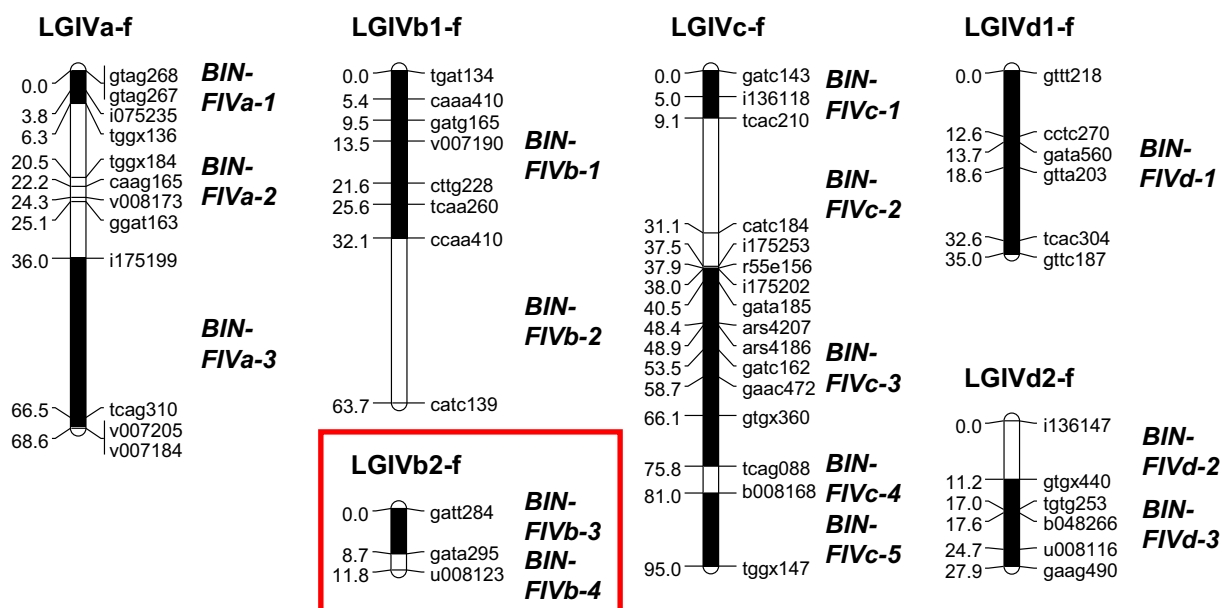
#### Identification of two markers flanking the *FaPFRU* locus

From the two new maps of 'Capitola' and 'Mara des Bois', we identified which markers flanked the *FaPFRU* locus using a QTL approach (Figure 1, Step5). Analyses were performed using the total number of inflorescences taken just after emergence that had appeared within the period from the end of May to the end of September, thereby representing the PF trait. This was carried out over 6 and 3 years of observation for 'Capitola' × 'CF1116' (Gaston *et al.*, 2013) and 'Mara des Bois' × 'Pajaro', respectively. A highly significant QTL (LOD superior to 15 and 10 for the female linkage maps of 'Capitola' and 'Mara des Bois', respectively) linked to the PF trait was detected on the bottom of LGIVb-f. The results obtained from 'Capitola' allowed us to determine that the maximum values for the QTL were located between the two markers Bx089\_196 and Bx064\_216 (Figure 3). These two markers were further used for screening larger populations to identify recombinants at the *FaPFRU* locus (Figure 1, Step6).

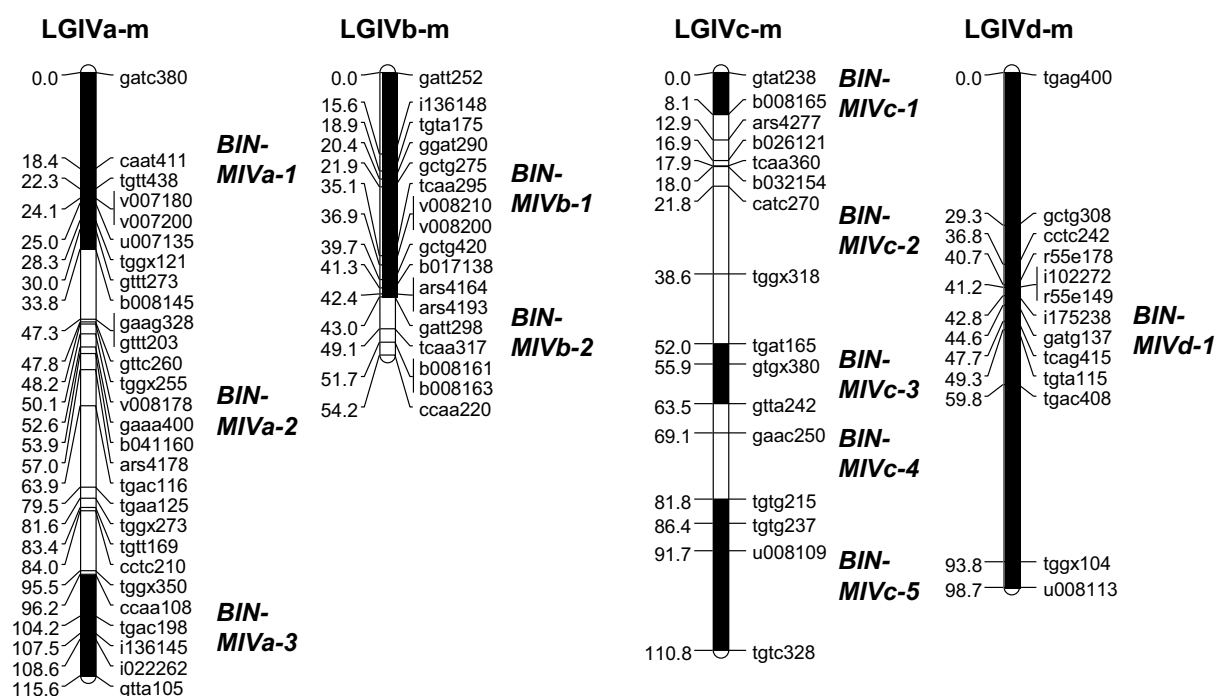
#### Characterization of recombinants reduced the *FaPFRU* locus interval

We screened 1451 individuals issued from different populations (Table 1) for the presence/absence of the two markers, Bx089\_196 and Bx064\_216, which delimit the *FaPFRU* locus. Out of the 1451 individuals tested (including the individuals of the two mapping populations), 104 recombinants were identified.

## (a) Female bins



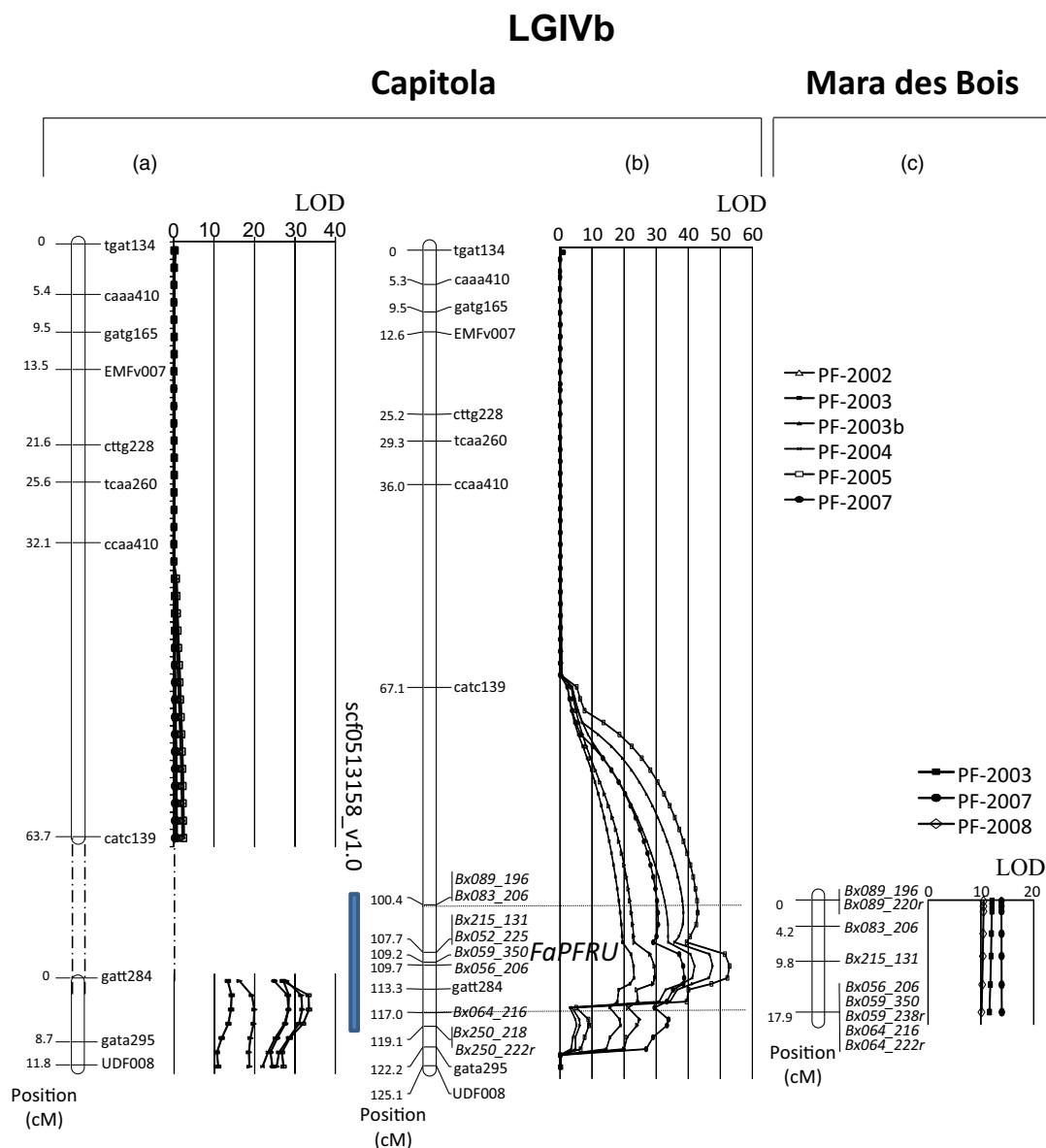
## (b) Male bins



**Figure 2** Representation of the bins on (a) the four female and (b) male linkage groups (LG) belonging to the homoeologous group IV. Bins consist of groups of markers separated by recombinational crossover sites. The position of each bin is shown by a black or white rectangle on the LG, and the name of each bin is indicated in *italics*. The targeted bins overlapping the *FaPFRU* locus are framed in red.

They displayed only one of the two markers, exhibiting a 0:1 or 1:0 profile (Bx089\_196:Bx064\_216) (Table 1). Similar percentages of recombinants were observed for each of the segregating population types, derived either from biparental crossing (7%) or by selfing (7%), which were independent of the PF parent, 'Capitola' (6%) or 'Mara des Bois' (10%). These results suggest that the recombination in this region was stable across different environments and genetic backgrounds.

To reduce the size of the locus, we genotyped the 104 recombinants with the five SSR markers Bx083\_206, Bx215\_131, Bx052\_225, Bx059\_350 and Bx056\_206 that are located within the interval framed by the Bx089\_196 and Bx064\_216 markers (Figure 4; Figure 1, Step7). Combining the five markers, the two frame markers and present (H) or absent (A) controls for each marker, but excluding the two profiles for the absence of recombination, AAAAAAA or HHHHHHH, 12



**Figure 3** QTLs linked to perpetual flowering (PF) trait identified in ‘Capitola’: (a) in Gaston *et al.* (2013) and (b) in this study and (c) in ‘Mara des Bois’ on the Linkage Group LGIVb. The blue bar represents the position of the scaffold scf0513158\_v1.0. Phenotypic data analysed for QTL identification are the number of newly emerged inflorescences from the end of May to the beginning of August, representing the PF trait, for 6 and 3 years, respectively ‘Capitola’ × ‘CF1116’ –(a) and (b)– and ‘Mara des Bois’ × ‘Pajaro’ –(c)– segregating populations. Markers added in this study are in italic.

profiles of recombination could be obtained. Of the 104 recombinants, only the AAAAAHH profile was not observed. Observing 11 out of 12 possible profiles indicated that recombination events at the *FaPFRU* locus are well represented (Figure 4 and Table S2). By adding the phenotypic value of the trait to the locus, PF or SF, we were able to assign the *FaPFRU* locus to the interval between the Bx083\_206 and Bx215\_131 markers, which were 7.3 cM apart in the ‘Capitola’ × ‘CF1116’ reference map.

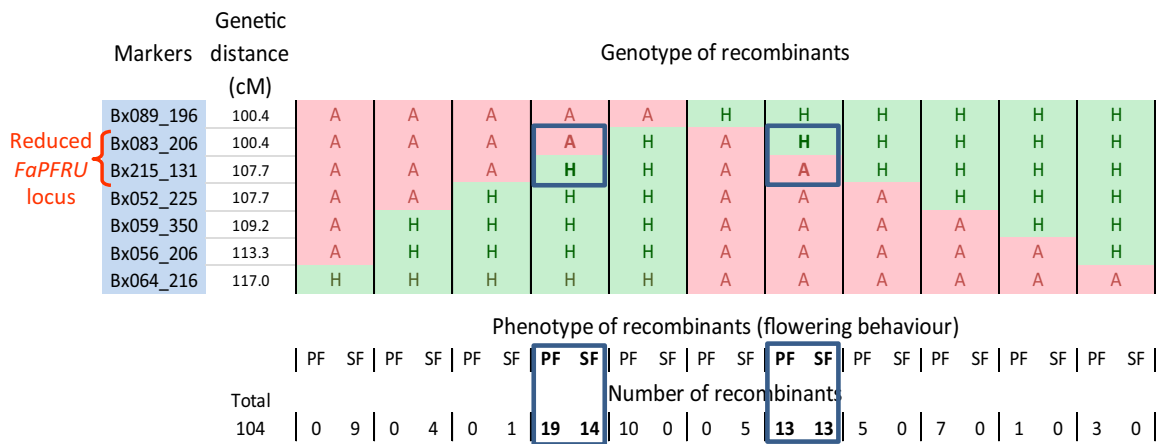
#### Identification of the *FLOWERING LOCUS T (FaFT)* gene as a promising candidate controlling the PF trait

As a prelude to a detailed investigation into the region corresponding to the reduced *FaPFRU* locus interval for the

diploid reference genome, we compared for each HGIV of the cultivated strawberry its genetic distance in cM with the physical distance of the corresponding LGIV in the diploid reference genome v2.0.a1 (available on [www.rosaceae.org](http://www.rosaceae.org)) (Jung *et al.*, 2014; Tennesen *et al.*, 2014). For each linkage group of HGIV, the markers are evenly distributed along a line (Figure 5) revealing a well conserved genetic organisation between the diploid reference genome and the octoploid genetic map for each of the HGs and the presence of a region without SSR markers between 17.3 Mb and 25.3 Mb. This gap may have been due to the presence of exclusively homozygous sequences or to substantial rearrangements between the diploid and the octoploid genomes in this region. However, previously mapped AFLP markers (Lerceteau-Köhler *et al.*, 2003), which were based on

**Table 2** Assignment of the 227 SSR markers to the female or male bins based according to the genetic profile they displayed on the six individuals of the reduced sample and the parents ‘Capitola’ and ‘CF1116’

	Marker number assigned to each bin according to its pattern									
	Female bins					Male bins				
	LGIVa-f	LGIVb-f	LGIVc-f	LGIVd-f	Not assigned	LGIVa-m	LGIVb-m	LGIVc-m	LGIVd-m	Not assigned
BIN-1	13	3	4	2		–	7	–	18	
BIN-2	–	4	2	14		6	3	3	No bin	
BIN-3	3	6	2	13		50	No bin	8	No bin	
BIN-4	No bin	3	6	No bin		No bin	No bin	7	No bin	
BIN-5	No bin	No bin	5	No bin		No bin	No bin	1	No bin	
Total	16	16	19	29	17	56	10	19	18	27
					97					130



**Figure 4** Genotype (H and A for the presence and absence of the markers), phenotype (PF and SF for perpetual and seasonal flowering) and the number of recombinants according to genotype and phenotype. The individuals presenting a recombination (framed in blue) between the markers Bx083\_206 and Bx215\_131 present both PF and SF behaviours (in bold), indicating that the *FaPFRU* locus controlling the trait of interest is localized between these two markers, which represent the two flanking markers of the reduced *FaPFRU* locus (bracket in red).

arbitrary sequences and could not be placed in the diploid reference genome, successfully revealed polymorphism in this region. These results support the hypothesis of a major DNA rearrangement rather than the consequence of homozygosity.

Based on the diploid reference genome v2.0.a1, the region between markers Bx083\_206 and Bx215\_131 flanking the *FaPFRU* locus is 1108 Kb and contains 234 predicted genes. An analysis of gene ontology (GO) was performed to assign possible functions to each of the 234 genes. Out of the 234 genes, 72 yielded no hit in the SwissProt database and were not annotated, but 15 were classified as flowering associated (Table 3). Among these 15 genes, the most likely candidate for the control of PF was a gene homologous to *FT*, the *gene04680* (Gaston *et al.*, 2013).

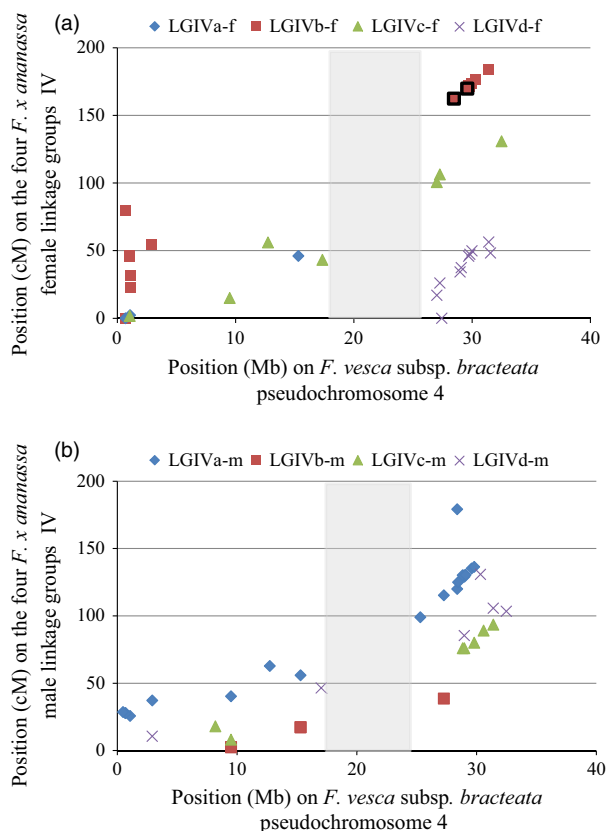
**Effect of the *FaPFRU* locus on the plant development**

An outstanding question regarding the *FaPFRU* locus is its influence on other plant developmental traits. One trait of interest is the relationship between the number of inflorescences and stolons produced by a plant, which was addressed by scoring the number of inflorescences and stolons during the period from the end of May until the end of September in a subset of four

populations segregating for this trait, ‘Capitola’-S<sub>1</sub>, ‘Mara des Bois’-S<sub>1</sub>, ‘Capitola’ × ‘CF1116’ and ‘Capitola’ × ‘Pajaro’. In agreement with previous findings in Gaston *et al.* (2013), we determined that PF individuals generated 4–28 times more inflorescences than stolons, whereas SF individuals produced a large number of stolons, but almost no inflorescences (Figure 6). These results revealed the opposite effects of *FaPFRU* on flowering (positive) and runnering (negative) independent of genetic background.

The *FaPFRU* locus also negatively affected plant size. Using co-dominant markers, in particular Bx089\_215c/Bx089\_220r, we investigated the dose effect of these markers on plant height and diameter in the ‘Mara des Bois’-S<sub>1</sub> population. In SF individuals showing homozygosity of the marker Bx089\_220r (or Null Dose (ND) of Bx089\_215c), the plants displayed the largest height and diameter. Among PF individuals, heterozygous plants (Bx089\_215c/Bx089\_220r) were taller and larger than homozygous plants (Bx089\_215c in Double Dose, DD) (Figure 7). However, no difference between homozygous and heterozygous PF individuals was observed in the number of either inflorescences or stolons. The negative effect on plant height was confirmed by the identification of a QTL located at the *FaPFRU*





**Figure 5** Comparison between physical and genetic distances for all informative markers (SSRs) localized on the four homoeologous (a) female and (b) male linkage groups (LG) IV. Physical distance was identified thanks to the pseudochromosome-4 sequence of the *F. vesca* reference genome v2.0.a1. Genetic distances are identified thanks to linkage maps of the octoploid population ‘Capitola’ × ‘CF1116’. The two flanking markers of the reduced *FaPFRU* locus on the LGIVb-f are framed in black. In grey, genomic region without SSR markers in our study.

locus, which was linked to the height measurements recorded in 2003 and 2004 (maximum LOD values of 7.1 and 5.9, respectively). For both years, the decrease in plant height was around 3 cm for PF individuals compared to that for SF individuals.

#### Additional validation using genetic resources

The presence or absence of the two markers Bx083\_206 and Bx215\_131 flanking the *FaPFRU* locus was investigated in 79 cultivars or selections of *F. × ananassa* (Table 4 and Table S3). The absence of both markers observed for 47 SF cultivars or selections was strictly associated with SF genotypes. The presence of both markers flanking the *FaPFRU* locus was associated with the PF trait for 14 (82%) out of the 17 PF cultivars/selections. However, nine out of the 62 SF cultivars/selections displayed both markers and six cultivars showed recombination in this region by displaying one of the two flanking markers. This study demonstrated that these markers will be useful in breeding programmes, independently of parental display of either or both flanking markers. It further reinforces the hypothesis that a single allelic variant controlling PF has been maintained through breeding of modern strawberry cultivars.

**Table 3** Gene Ontology (GO) for the 15 candidate genes linked to flowering and present in the reduced interval of the *FaPFRU* locus delimited by Bx083 and Bx215 SSRs (E values superior to  $10^{-5}$ )

Name	Annotation	Gene function
<i>gene04680</i>	Flowering locus T (FT)	Florigen
<i>gene04683</i>	Protein Argonaute 7 (AGO7)	Gene silencing by miRNA
<i>gene04726</i>	Agamous-like Mads-box protein (AGL30)	Member of the MIKC family of transcriptional regulators
<i>gene04596</i>	Leucine-Rich Repeat Receptor like kinase (RPK2)	Regulator of meristem maintenance
<i>gene03989</i>	Sucrose transport protein (SUC3)	Sucrose transport
<i>gene04730</i>	Cyclic dof factor 2 – zinc finger protein (CDF2)	DNA binding transcription factor activity
<i>gene04745</i>	Zinc finger homeodomain protein 9 (ZHD9)	DNA binding transcription factor activity
<i>gene03991</i>	Homeobox leucine zipper protein (ATHB-13)	DNA binding transcription factor activity
<i>gene04031</i>	Homeobox-leucine zipper protein (HAT5)	DNA binding transcription factor activity
<i>gene04729</i>	Probable calcium-binding protein 45 (CLP45)	Involved in calcium transport
<i>gene04028</i>	Cullin-1 (CUL1)	Involved in ubiquitination process
<i>gene04757</i>	Cullin-3a (CUL3a)	Involved in ubiquitination process
<i>gene04002</i>	Ubiquitin-conjugating enzyme E2 (UBC2)	Involved in ubiquitination process
<i>gene04611</i>	Ureide Permease 2 (UPS2)	Uracyl transport
<i>gene04612</i>	Ureide Permease 2 (UPS2)	Uracyl transport

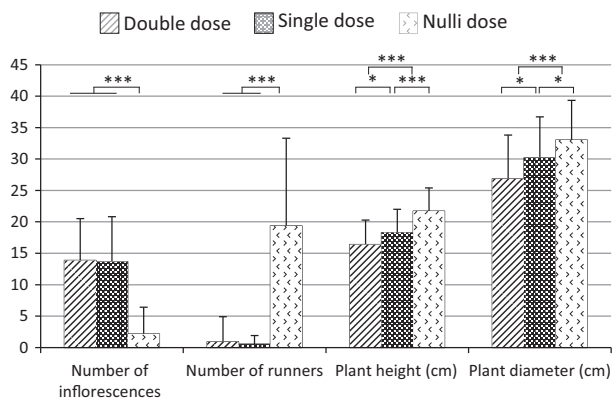
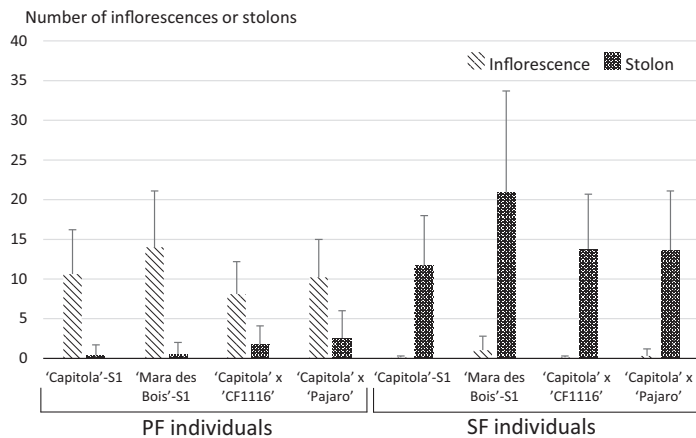
## Discussion

The perpetual flowering (PF) trait of strawberry extends the length of the flowering period and so can prolong the fruit production season. Thus, it contributes not only to improve fruit yield but also to open new markets for strawberries, for example by offering fruit to consumers outside the production season. For these reasons, most strawberry breeding programmes have aimed at introgressing this trait into elite cultivars. However, this task remains challenging and requires considerable effort, due to the interaction of the PF trait with its genetic and environmental context (Maltoni *et al.*, 1996). The identification of genetic markers in close linkage to the *FaPFRU* locus will permit the discrimination of PF cultivars/selections at very early stages of plant development, thereby circumventing the need for phenotyping the large numbers of mature plants under favourable environment conditions. Identification of the *FaPFRU* allelic variant will further provide insights into the mechanisms governing the balance between flowering and runnering in strawberry.

#### A strategy to narrow down the *FaPFRU* locus that circumvents polyploid complexity

The *FaPFRU* locus found in the cultivated strawberry results from the introgression of a PF locus from a *F. virginiana* ssp. *glauca* Wasatch accession collected from the Wasatch Mountains in Utah (Bringhurst and Voth, 1984; Powers, 1954). From then on, this original introgression has been largely used in breeding until

**Figure 6** The number of inflorescences and stolons of individuals belonging to four populations scored in 2012 according to their flowering behaviour, perpetual flowering (PF) or seasonal flowering (SF).



**Figure 7** Dose effect of the co-dominant markers Bx089\_215c/ Bx089\_220r localized in the *FaPFRU* locus. Phenotyping was performed on 'Mara des Bois'-S<sub>1</sub> population. Double Dose indicates homozygous individuals for Bx089\_215c (linked in coupling to the *FaPFRU* allelic variant), Single Dose indicates heterozygous individuals for Bx089\_215c/ Bx089\_220r and Nulli Dose indicates homozygous individuals for Bx089\_220r. Asterisks indicate significant differences by Student's t-test: \*, \*\*, \*\*\* for statistical significance at  $P \leq 0.05$ ,  $0.01$ ,  $0.001$ , respectively.

**Table 4** Distribution of the 79 cultivars/selections belonging to genetic resources according to their status on the two flanking markers

	Marker status				Total
	Bx083_206/Bx215_131				
Flowering behaviour of genotypes	+/+	+/- or -/+	-/-		
Seasonal flowering	9	6	47		62
Perpetual flowering	14	3	0		17

today (Hancock, 1999). However, developing markers to follow this trait has been largely impeded by the polyploid status of the cultivated octoploid strawberry.

With recent and ancestral events of whole-genome duplication (WGD) events, nearly all seed plants have experienced episodes of polyploidy (Jiao *et al.*, 2011; Renny-Byfield and Wendel, 2014). In diploid species, such as *Arabidopsis thaliana*, successive events of

duplications have led to the colinearity of sections of the genome (Blanc *et al.*, 2000). Duplication of genomes complicates mapping of markers for several reasons. It requires substantial effort to find polymorphisms between homoeologous regions in closely related sub-genomes (Njuguna *et al.*, 2013) and to decipher the effect of multi-dose markers (van Dijk *et al.*, 2014). Furthermore, the divergence among homoeologs could be influenced by dynamic genomic changes revealing asymmetrical evolution of sub-genomes (Liu *et al.*, 2014) or affecting many sites at once by introgression and chromosomal rearrangements (Tennessen *et al.*, 2014). Most mapping strategies require considerable work to circumvent these problems in polyploid crop species, and so map-based breeding remains very costly. In recent years, highthroughput genotyping of complex polyploid species has benefited considerably from improved genetic linkage maps possessing greater marker density. Nevertheless, gaps are still present as shown in the work using SNP arrays in strawberry (Basil *et al.*, 2015) and cotton (Hulse-Kemp *et al.*, 2015) or through DArT markers in strawberry (Sánchez-Sevilla *et al.*, 2015) or *Brassica* (Raman *et al.*, 2013). Within these linkage maps, marker coverage may be uneven and/or sparse in some genomic regions due to the lower polymorphism level in some sub-genomes, as shown for wheat (Maccaferri *et al.*, 2015), cotton (Blenda *et al.*, 2012) and strawberry (Tennessen *et al.*, 2014). Consequently, a unique LG may become split into different sub-LGs, as observed for the strawberry LGIVb-f that has been separated into two sub-LGs (Gaston *et al.*, 2013).

In strawberry, one genome of the diploid species corresponds to four homoeologous genomes in the cultivated octoploid species (Rousseau-Gueutin *et al.*, 2009; Tennessen *et al.*, 2014). In this context, the fine mapping of a specific homoeologous region by marker enrichment may prove unrevealing or, at best, very time consuming. For example, markers suffer from random mapping to one of four sub-genomes instead of being mapped to the specific sub-genome carrying the homoeologous region of interest. We reasoned that the selective mapping strategy first described by Vision *et al.* (2000), which was successfully used in diploid crop species (e.g. Cabrera *et al.*, 2009; Howad *et al.*, 2005; Sargent *et al.*, 2008), could be applied to complex polyploid species to overcome these limitations. When applied to polyploids, this strategy employs a reduced sample of highly informative individuals chosen according to the position of their recombinational crossover sites in the homoeologous region of interest. In strawberry, it facilitated the joining of the two sub-

linkage groups of LGIVb-f into a unique LG and helped to dramatically narrow down the single homoeologous *FaPFRU* locus carrying the allelic variant controlling the PF trait. As a consequence, flanking markers tightly linked to *FaPFRU* are available for MAS, and the candidate genes comprising the locus have been identified. This locus is likely the same as the one described in previous studies (Castro *et al.*, 2015; Honjo *et al.*, 2016) because all markers developed in our work and in previous studies are located on the same chromosomal region of the LG4 in the diploid reference genome. However, the flanking markers developed in our study are closer to the locus as shown in Figure S1.

This strategy remains relevant even with more dense linkage maps, which still show gaps.

If a more saturated map is used instead of a framework linkage map, the overall strategy will be the same, but smaller bins will be obtained. In addition, when taking into account multi-dose markers useful for the identification of homoeoQTLs (Lerceteau-Köhler *et al.*, 2012), the choice of an appropriate reduced sample of individuals should allow the fine mapping of different homoeologous loci.

### PF has a single origin in cultivated strawberry and has a different genetic control than in diploid strawberry

In addition to the *FaPFRU* locus that originated from the *F. virginiana* species (Bringhurst and Voth, 1984; Powers, 1954), two other origins of PF trait in the cultivated strawberry have been reported (Darrow, 1966). The cultivar 'Pan American' released in the USA in 1902 is believed to be a mutant of the Bismarck cultivar, whereas Gloede's seedling was first identified in France (Richardson, 1914). Thus, there is still a doubt regarding the origin of the PF trait in modern cultivated cultivars. In this study, we tested 79 different cultivars and selections from our strawberry germplasm, including 17 PF cultivars from various breeding programmes from around Europe, Japan and USA. Results showed the presence of at least one of the flanking markers of the *FaPFRU* locus within the various PF cultivars, confirming that the modern PF cultivars have been derived from the same *FaPFRU* introgression event. Because of time inconsistency between the release in Europe of the PF parent of 'Mara des Bois' ('Hummi Gento' 1967) and the first release of modern PF cultivars related to 'Capitola' in USA (1971–1972) (Bringhurst and Voth, 1980), it had been previously assumed that the origin of the PF behaviour in these cultivars was different. However, we identified in this study six SSR markers in linkage disequilibrium with the PF allele that were identical in 'Mara des Bois' and 'Capitola'. It is likely that lines selected from offsprings from the Wasatch genotype were sent from USA to Europe before the release of the first PF cultivars in USA. We therefore believe that the origin of the PF allele is the same in 'Mara des Bois' and 'Capitola'.

In addition, we found that PF cultivars in the tested germplasm are heterozygous at this locus (data not shown), suggesting that the presence of a dominant homozygous allelic variant may be a disadvantage. In fact, a significant negative co-dominant effect of an *FaPFRU* allelic variant on plant size was observed, indicating that the *FaPFRU* locus plays a major role in strawberry plant fitness by controlling the balance between sexual reproduction and vegetative development independent of dose. Although this effect was not linked to the number of inflorescences produced by the plant, a trade-off between fruit production and growth of vegetative organs is possible, such as through competition for available carbon.

The diploid and octoploid PF strawberries share similar physiological control of the perpetual flowering process (Sønsteby and Heide, 2008). PF occurs during long days and is antagonistic to the stolon production (Tenreira T and Denoyes B, unpublished data). Deciphering the mechanisms involved in the regulation of PF and running in strawberry through the identification of allelic variants underlying the variations in both traits will be of considerable help. Genetic controls of known PF traits in diploid and cultivated octoploid strawberry are likely to be different. PF was mapped to LGVI in diploid strawberry *F. vesca* (Albani *et al.*, 2004) and to LGIVb in octoploid cultivated strawberry *F. × ananassa* (Gaston *et al.*, 2013). We and others previously showed that a two-bp mutation in *FvKSN*, a strawberry ortholog of the *TERMINAL FLOWER 1 (TFL1)* floral repressor, was responsible for the PF trait in diploid strawberry (Iwata *et al.*, 2012; Koskela *et al.*, 2012). The role of *TFL1* as a floral repressor in the cultivated octoploid strawberry was also recently confirmed by Koskela *et al.* (2016). In this study, silencing the floral repressor *FaTFL1* in a SF cultivar through a transgenic approach was sufficient to induce PF under long days. In the cultivated octoploid strawberry, Nakano *et al.* (2015) proposed that *FaFT3*, one of the three *FT-like* genes described in *Fragaria* (Iwata *et al.*, 2012), is a key gene in flower induction. In a recent study using *F. vesca*, Mouhu *et al.* (2013) further explored the possible implication of orthologs of *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (FvSOC1)* and *FLOWERING LOCUS T (FvFT1)*, which are two genes upstream of *KSN/TFL1* in the flowering signalling pathway. Inactivation of *FvSOC1* through RNAi-silencing led to PF. However, evidence is lacking that *FvFT1* inactivation would lead to PF in a SF genetic background. The patterns of *FvFT1* expression suggest that it may influence the photoperiodic regulation of *FvSOC1* (Koskela *et al.*, 2012; Mouhu *et al.*, 2013). Although *FaTFL1*, *FaFT3*, *FaFT1* and *FaSOC1* appear as potential candidates to regulate PF in cultivated octoploid strawberry, they can be excluded because *FvSOC1* is located on LGVII, *FvFT3* is located on LGIII and *FvTFL1* and *FvFT1* are located on LGVI in diploid strawberry. In contrast, the *FaFT-like* gene (named *FaFT2* in Nakano *et al.*, 2015) that is orthologous to *gene04680-v1.0-hybrid*, and which is located within the reduced *FaPFRU* locus, appears to be an excellent candidate. Like *TFL1*, *FT* belongs to the phosphatidylethanolamin binding protein (PEBP)-encoding gene family and plays a crucial role in flowering. Whereas *TFL1* protein represses the transition from vegetative to reproductive phase, *FT* protein promotes these phases and is considered to be the florigen (Corbesier *et al.*, 2007). *FT* and *TFL1* proteins may compete for binding to the transcription factor *FD* to either activate or repress the meristem identity genes (Hanano and Goto, 2011; Pin and Nilsson, 2012; review in Wickland and Hanzawa, 2015). In this hypothesis, an allelic variant of *FaFT2* would act as a positive regulator of flowering. *FaFT2* would therefore overcome the floral repression in long days carried by *FaTFL1* in PF-cultivated octoploid strawberry. Although *FaFT-like* is the most likely candidate, the other known flowering associated genes should not be neglected in future studies aimed at deciphering the PF-regulating mechanisms in cultivated strawberry.

### Conclusion

Perpetual flowering is as a major agronomical trait, because its characteristics permit the extension of the fruit production season. The development of an original approach, based on

selective mapping to enrich a single region of one specific homoeologous linkage group, allowed us to fine map the *FaPFRU* locus controlling PF in the highly polyploid cultivated strawberry. We described the rapid development of two flanking markers, which were subsequently validated in a second segregating population and in genetic resources. These findings lay the foundation for more rapid and efficient creation of new PF strawberry cultivars through the marker-assisted selection. In addition, the identification of likely candidate genes will advance our understanding of flowering in strawberry.

## Experimental procedures

### Plant materials and PF phenotyping

Plant material of the cultivated octoploid strawberry, *F. × ananassa*, which was used for fine mapping the region controlling the PF trait and enhancing the prevalence of crossover events in this region, was obtained from various crosses or selfings. These populations were issued from the heterozygous PF parents 'Capitola' or 'Mara des Bois' and the homozygous SF parents 'CF1116' or 'Pajaro' (Table 1). For the two pseudo, full-sibling  $F_1$  populations 'Capitola' × 'CF1116' and 'Mara des Bois' × 'Pajaro' linkage maps were created. The first one representing our reference segregating population to identify the *FaPFRU* locus (Gaston *et al.*, 2013) was increased to 375 individuals for this study to facilitate identifying recombinants. For the second population, a framework linkage map was obtained from 85 individuals. Recombinants were screened in these two populations described above and 212 individuals of a third  $F_1$  population 'Capitola' × 'Pajaro', and selfings of four genotypes 'Capitola'- $S_1$ , 'Mara des Bois'- $S_1$ , ['Mara des Bois' × 'Pajaro']- $F_2$ -3 and ['Mara des Bois' × 'Pajaro']- $F_2$ -45 that were already known to be heterozygous for the PF trait. All together, these various populations provided 1451  $F_1$ ,  $S_1$  or  $F_2$  individuals.

Seeds were germinated in the greenhouse with a 16-h photoperiod using supplementary incandescent light and a day and night temperatures of  $24 \pm 6$  °C and  $18 \pm 4$  °C, respectively. Four to six weeks after sowing, seedlings were transplanted to 72-cell plug trays containing sand/vermiculite substrate (1:1 by volume) at a density of one plant per  $5 \times 5$  cm well.

For studying the presence of markers linked to the PF behaviour in the genetic resources, we analysed 79 cultivars or selections with various pedigrees and origins (Table S3).

Flowering behaviour of individuals of the two mapping populations and recombinants was identified as described in the study by Gaston *et al.* (2013). In addition, selfings of 'Mara des Bois'- $S_1$  were characterized for their plant size, height and diameter.

For QTL analyses, phenotypic data were represented by the number of newly-emerged inflorescences from the end of May to the beginning of August and were described in the study by Gaston *et al.* (2013) or obtained after 3 years of observation for 'Capitola' × 'CF1116' and 'Mara des Bois' × 'Pajaro' populations.

### DNA extraction

Genomic DNA was extracted using the DNeasy™ Plant Mini Kit or DNeasy™ 96 Plant Kit (Qiagen Inc., Hilden, Germany) with 90 mg or 45 mg, respectively, of young, developed, leaf tissue according to the manufacturer's protocol and modified as previously described (Lerceteau-Köhler *et al.*, 2003). gDNA was quantified visually using the DNA-binding dye GelGreen™ (Biotium, Hayward, CA) and standards of known quantity.

### Reduced sample choice for optimizing the identification of markers on the LGIVb-f

A goal was to reduce the effort required to obtain new markers in coupling-phase linkage with the PF-specific allele at *FaPFRU*, rather than the corresponding, SF-specific allele. Therefore, we developed a selective mapping procedure based on the selection of a specific set of highly informative individuals chosen for the position of their recombinational crossover sites (Vision *et al.*, 2000) in the homoeologous group IV, in which the LGIVb-f carries the allelic variant *FaPFRU*. The reduced set was a subset of plants from the 'Capitola' × 'CF1116' pseudo full-sibling  $F_1$  population, for which framework female and male (f and m) maps were previously obtained (Lerceteau-Köhler *et al.*, 2003; Rousseau-Gueutin *et al.*, 2008). BIN names consisted of BIN, the female or male map, the LG, and a number indicating the relative position of the bin on the LG. For example, BIN-MIVb-2 indicates BIN 2 on the male LGIVb.

### Development of SSR markers for fine mapping

Simple sequence repeat primer pairs for fine mapping were designed from the *F. vesca* v2.0.a1 Pseudomolecule Assembly (Tennesen *et al.*, 2014) using colinearity with the homoeologous genomes of the cultivated strawberry. We focused on the large scaffold scf0513158\_v1.0 (4 198 144 bp) located at LGIVb-f, which was upstream from the UDF008 marker at the end of the *FaPFRU* QTL (Gaston *et al.*, 2013) (Figure 3). SSR motifs were detected by the Phobos plug-in compatible with the Geneious software (version 6.1.4; Biomatters Ltd, Auckland, New Zealand). Primers were designed for repeats that contained at least 10 bp or for combinations of repeats. See Table S4 for the primers used in this study. Primers were designed in the SSR flanking regions based on the diploid reference genome using Primer3 plug-in of Geneious. In addition, SSR primer pairs previously published and located on the HGIV, and more precisely on the scaffold scf0513158\_v1.0, were also studied (Isobe *et al.*, 2013).

PCR was performed using the standard PCR protocols at an annealing temperature of 55 °C. Amplicons derived from the 291 SSRs obtained on the reduced sample were screened on 6% (w/v) polyacrylamide gels as previously described (Rousseau-Gueutin *et al.*, 2008). The seven markers tested on recombinants were formed on the basis of screening all 1451 individuals using a specific PCR protocol, and amplicons were visualized using an Applied Biosystems (ABI) 3730 DNA Analyzer (Foster City, CA, USA) to obtain allele sizes. Cycling conditions for all primers except Bx215 were as follows: an initial denaturation at 94 °C for 3 min followed by 30 cycles of 94 °C for 40 s, 55 °C for 45 s and 72 °C for 45 s. Primer pair Bx215 was annealed at 58 °C. These cycles were followed by eight cycles of PCR with a lower annealing temperature to incorporate the 6-FAM™ M13 or VIC® M13 label at 94 °C for 40 s, 53 °C for 45 s and 72 °C for 45 s, with a final extension at 72 °C for 30 min. PCR products were diluted from 1:20–1:50 according to the primer pairs for the fragment analysis on an ABI 3730 DNA Analyzer calibrated with an ABI Genescan™ 600 LIZ Size Standard (reference 4366589). Peaks data were scored using GeneMapper V4.0 software (Applied Biosystems, Foster City, CA, USA) and converted to a qualitative, binary code to facilitate scoring for either the presence or absence. For each SSR marker, the first part of the identifier indicates the name of the SSR followed by a number indicating the allele size in bp. For marker Bx064\_122 for example, Bx064 states the SSR primer pairs, and 122 represents the allele size of the amplicon.

### Development of markers that flanked the QTL for further recombinant identification

For QTL detection on the female linkage map of 'Capitola' × 'CF1116', we used the framework map previously developed (Lerceteau-Köhler *et al.*, 2012; Rousseau-Gueutin *et al.*, 2008), including the nine markers at LGIVb-f found in this study, and the phenotypic data were those described in the study by Gaston *et al.* (2013). QTL analyses were performed by the composite interval mapping (CIM) (Zeng, 1993, 1994) using QTL Cartographer software (Basten *et al.*, 1997) as described in the study by Gaston *et al.* (2013). For 'Mara des Bois' × 'Pajaro', similar QTL analyses were performed on the seven linkage groups constructed with 37 markers. This linkage map included the orthologous LGIVb.

The QTL analysis yielded the two flanking alleles delimiting the peak of the *FaPFRU* locus that were used for the subsequent identification of recombinants (Figure 3). The 1,451 S<sub>1</sub> or F<sub>2</sub> seedlings were amplified with primer pairs of Bx083 and Bx215 for genotyping the two flanking alleles Bx083\_206 and Bx215\_131. Only recombinants showing one of the two markers were kept for further phenotyping.

### Comparison between genetic and physical distances in the HGIV

To see whether the mapped markers were distributed regularly on the homoeologous chromosomes belonging to the HGIV, we compared genetic with physical distances. This comparison was performed on both identified female and male linkage groups with the pseudomolecule IV of the diploid reference genome (v2.0.a1). For identifying the physical map location of the microsatellite markers used in the genetic octoploid map, we used BLAST against the *F. vesca* v2.0.a1 Pseudomolecule Assembly: (Tenessen *et al.*, 2014). Genetic distances were obtained using the software MapMaker/Exp v3.0 (Lander *et al.*, 1987).

### Identifying candidate genes within the refined region

Within the interval framed by the SSRs Bx083\_206 and Bx215\_131, a gene search was made in GDR browser on the hybrid version of the genome annotation (<http://www.rosaceae.org>). The proteins of the identified genes were subjected to a Blast2Go analysis (Conesa *et al.*, 2005) to predict function through gene ontology (GO). A BLASTp was run against the SwissProt database with an E value at 10<sup>-5</sup>. The mapping and annotation steps were run using default parameters. Genes were filtered for those predicted to be involved in the biological processes related to flowering.

### Acknowledgements

This work was supported by Région Aquitaine (Régal project) and was carried out under the remit of the EU FP7-KBBE-2010-4 Project (Grant Agreement no. 265942). The authors gratefully acknowledge the participation of Laetitia Riva-Roveda, Sadia Schafleitner and Alain Bonnet in helping with phenotyping.

### References

Albani, M.C. and Coupland, G. (2010) Comparative analysis of flowering in annual and perennial plants. *Curr. Top. Dev. Biol.* **91**, 323–348.

- Albani, M.C., Battey, N.H. and Wilkinson, M.J. (2004) The development of ISSR-derived SCAR markers around the *SEASONAL FLOWERING LOCUS (SFL)* in *Fragaria vesca*. *Theor. Appl. Genet.* **109**, 571–579.
- Bassil, N.V., Davis, T.M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., Mahoney, L. *et al.* (2015) Development and preliminary evaluation of a 90 K Axiom<sup>®</sup> SNP array for the allo-octoploid cultivated strawberry *Fragaria × ananassa*. *BMC Genom.* **16**, 155.
- Basten, C.J., Weir, B.S. and Zeng, Z.-B. (1997) *QTL Cartographer: A Reference Manual and Tutorial for QTL Mapping*. Raleigh, NC: Dep. Stat. North Carolina State Univ.
- Blanc, G., Barakat, A., Guyot, R., Cooke, R. and Delseny, M. (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell*, **12**, 1093–1101.
- Blenda, A., Fang, D.D., Rami, J.F., Garsmeur, O., Luo, F. and Lacape, J.M. (2012) A high density consensus genetic map of tetraploid cotton that integrates multiple component maps through molecular marker redundancy check. *PLoS ONE*, **7**, e45739.
- Bowers, J.E., Chapman, B.A., Rong, J.K. and Paterson, A.H. (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, **422**, 433–438.
- Bringhurst, R.S. and Voth, V. (1980) 6 new strawberry varieties released. *Calif. Agric.* **34**, 12–15.
- Bringhurst, R.S. and Voth, V. (1984) Breeding octoploid strawberries. *Iowa State J. Res.* **58**, 371–381.
- Cabrera, A., Kozik, A., Howad, W., Arus, P., Iezzoni, A.F. and van der Knaap, E. (2009) Development and bin mapping of a *Rosaceae* Conserved Ortholog Set (COS) of markers. *BMC Genom.* **10**, 562.
- de Camacaro, M.E.P., Camacaro, G.J., Hadley, P., Battey, N.H. and Carew, J.G. (2002) Pattern of growth and development of the strawberry cultivars elsanta, bolero, and everest. *J. Am. Soc. Hortic. Sci.* **127**, 901–907.
- Castro, P., Bushakra, J.M., Stewart, P., Weebadde, C.K., Wang, D., Hancock, J.F., Finn, C.E. *et al.* (2015) Genetic mapping of day-neutrality in cultivated strawberry. *Mol. Breed.* **35**, 79.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X., Chiquet, J. *et al.* (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, **345**, 950–953.
- Conant, G.C., Birchler, J.A. and Pires, J.C. (2014) Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr. Opin. Plant Biol.* **19**, 91–98.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.
- Corbesier, L., Vincent, C., Jang, S.H., Fornara, F., Fan, Q.Z., Searle, I., Giakountis, A. *et al.* (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, **316**, 1030–1033.
- Darrow, G. (1966) *The Strawberry: History, Breeding and Physiology*. New York: Holt, Rinehart and Winston.
- van Dijk, T., Pagliarani, G., Pikunova, A., Noordijk, Y., Yilmaz-Temel, H., Meulenbroek, B., Visser, R.G. *et al.* (2014) Genomic rearrangements and signatures of breeding in the allo-octoploid strawberry as revealed through an allele dose based SSR linkage map. *BMC Plant Biol.* **14**, 55.
- Gaston, A., Perrotte, J., Lerceteau-Köhler, E., Rousseau-Gueutin, M., Petit, A., Hernould, M., Rothan, C. *et al.* (2013) *PFRU*, a single dominant locus regulates the balance between sexual and asexual plant reproduction in cultivated strawberry. *J. Exp. Bot.* **64**, 1837–1848.
- Goldberg, M.T., Spigler, R.B. and Ashman, T.L. (2010) Comparative genetic mapping points to different sex chromosomes in sibling species of wild strawberry (*Fragaria*). *Genetics*, **186**, 1425–1433.
- Hanano, S. and Goto, K. (2011) *Arabidopsis* *TERMINAL FLOWER1* is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell*, **23**, 3172–3184.
- Hancock, J.F. (1999) *Strawberries*. Wallingford, UK: CABI Publishing, pp. 237.
- Hegarty, M.J. and Hiscock, S.J. (2008) Genomic clues to the evolutionary success of polyploid plants. *Curr. Biol.* **18**, R435–R444.
- Hirakawa, H., Shirasawa, K., Kosugi, S., Tashiro, K., Nakayama, S., Yamada, M., Kohara, M. *et al.* (2014) Dissection of the Octoploid Strawberry Genome by Deep Sequencing of the Genomes of *Fragaria* Species. *DNA Res.* **21**, 169–181.

- Honjo, M., Nunome, T., Kataoka, S., Yano, T., Hamano, M., Yamazaki, H., Yamamoto, T. *et al.* (2016) Simple sequence repeat markers linked to the everbearing flowering gene in long-day and day-neutral cultivars of the octoploid cultivated strawberry *Fragaria × ananassa*. *Euphytica*, **209**, 291–303.
- Howad, W., Yamamoto, T., Dirlewanger, E., Testolin, R., Cosson, P., Cipriani, G., Monforte, A.J. *et al.* (2005) Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics*, **171**, 1305–1309.
- Hulse-Kemp, A.M., Lemm, J., Plieske, J., Ashrafi, H., Buyyarapu, R., Fang, D.D., Frelichowski, J. *et al.* (2015) Development of a 63K SNP array for cotton and high-density mapping of intra- and inter-specific populations of *Gossypium* spp. *G3 (Bethesda)*, **5**, 1187–1209.
- Isobe, S.N., Hirakawa, H., Sato, S., Maeda, F., Ishikawa, M., Mori, T., Yamamoto, Y. *et al.* (2013) Construction of an integrated high density simple sequence repeat linkage map in cultivated strawberry (*Fragaria × ananassa*) and its applicability. *DNA Res.* **20**, 79–92.
- Iwata, H., Gaston, A., Remay, A., Thouroude, T., Jeauffre, J., Kawamura, K., Oyant, L.H. *et al.* (2012) The *TFL1* homologue *KSN* is a regulator of continuous flowering in rose and strawberry. *Plant J.* **69**, 116–125.
- Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P. *et al.* (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature*, **473**, 97–100.
- Jung, S., Ficklin, S.P., Lee, T., Cheng, C.H., Blenda, A., Zheng, P., Yu, J. *et al.* (2014) The genome database for *rosaceae* (GDR): year 10 update. *Nucleic Acids Res.* **42**, D1237–D1244.
- Koskela, E.A., Mouhu, K., Albani, M.C., Kurokura, T., Rantanen, M., Sargent, D.J., Batten, N.H. *et al.* (2012) Mutation in *TERMINAL FLOWER1* reverses the photoperiodic requirement for flowering in the wild strawberry *Fragaria vesca*. *Plant Physiol.* **159**, 1043–1054.
- Koskela, E.A., Sønsteby, A., Flachowsky, H., Heide, O.M., Hanke, M.V., Elomaa, P. and Hytönen, T. (2016) *TERMINAL FLOWER1* is a breeding target for a novel everbearing trait and tailored flowering responses in cultivated strawberry (*Fragaria × ananassa* Duch.). *Plant Biotechnol. J.* doi:10.1111/pbi.12545.
- Lander, E.S., Green, P., Abrahamson, J. and Barlow, A. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, **181**, 174–181.
- Lerceteau-Köhler, E., Guérin, G., Laigret, F. and Denoyes-Rothan, B. (2003) Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria × ananassa*) using AFLP mapping. *Theor. Appl. Genet.* **107**, 619–628.
- Lerceteau-Köhler, E., Moing, A., Guérin, G., Renaud, C., Petit, A., Rothan, C. and Denoyes, B. (2012) Genetic dissection of fruit quality traits in the octoploid cultivated strawberry highlights the role of homoeo-QTL in their control. *Theor. Appl. Genet.* **124**, 1059–1077.
- Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A., Zhao, M. *et al.* (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* **5**, 3930.
- Maccaferri, M., Ricci, A., Salvi, S., Milner, S.G., Noli, E., Martelli, P.L., Casadio, R. *et al.* (2015) A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* **13**, 648–663.
- Maltoni, M.L., Faedi, W., Simpson, D.W., Dradi, R. and Persano, S. (1996) Influenza dell'ambiente sulla espressione del carattere rifiorienza in Fragola. *Italus Hortus*, **3**, 30–36.
- Mayer, K.F., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B. *et al.* (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, **345**, 1251788.
- Mouhu, K., Kurokura, T., Koskela, E.A., Albert, V.A., Elomaa, P. and Hytönen, T. (2013) The *Fragaria vesca* homolog of *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* represses flowering and promotes vegetative growth. *Plant Cell*, **25**, 3296–3310.
- Nakano, Y., Higuchi, Y., Yoshida, Y. and Hisamatsu, T. (2015) Environmental responses of the *FT/FTL1* gene family and their involvement in flower induction in *Fragaria × ananassa*. *J. Plant Physiol.* **177**, 60–66.
- Njuguna, W., Liston, A., Cronn, R., Ashman, T.L. and Bassil, N. (2013) Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Mol. Phylogenet. Evol.* **66**, 17–29.
- Ohno, S. (1970) *Evolution by Gene Duplication*. Berlin-Heidelberg-New York: Springer-Verlag.
- Paterson, H., Bowers, J.E. and Chapman, B. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl Acad. Sci. USA*, **101**, 9903–9908.
- Pin, P. and Nilsson, O. (2012) The multifaceted roles of *FLOWERING LOCUS T* in plant development. *Plant, Cell Environ.* **35**, 1742–1755.
- Powers, L. (1954) Inheritance of period of blooming in progenies of strawberries. *Proc. Am. Soc. Hortic. Sci.* **64**, 293–298.
- Raman, H., Raman, R., Kilian, A., Detering, F., Long, Y., Edwards, D., Parkin, I.A. *et al.* (2013) A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: applications in genetic dissection of qualitative and quantitative traits. *BMC Genom.* **14**, 277.
- Renny-Byfield, S. and Wendel, J.F. (2014) Doubling down on genomes: polyploidy and crop plants. *Am. J. Bot.* **101**, 1711–1725.
- Richardson, C.W. (1914) A preliminary note on the genetics of *Fragaria*. *J. Genet.* **3**, 171–178.
- Rousseau-Gueutin, M., Lerceteau-Köhler, E., Barrot, L., Sargent, D.J., Monfort, A., Simpson, D., Arús, P. *et al.* (2008) Comparative genetic mapping between octoploid and diploid *Fragaria* species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. *Genetics*, **179**, 2045–2060.
- Rousseau-Gueutin, M., Gaston, A., Aïnouche, A., Aïnouche, M.L., Olbricht, K., Staudt, G., Richard, L. *et al.* (2009) Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. *Mol. Phylogenet. Evol.* **51**, 515–530.
- Sánchez-Sevilla, J.F., Horvath, A., Botella, M.A., Gaston, A., Folta, K., Kilian, A., Denoyes, B. *et al.* (2015) Diversity Arrays Technology (DArT) marker platforms for diversity analysis and linkage mapping in a complex crop, the octoploid cultivated strawberry (*Fragaria × ananassa*). *PLoS ONE*, **10**, e0144960.
- Sargent, D.J., Clarke, J., Simpson, D.W., Tobutt, K.R., Arús, P. and Monfort, A. (2008) An enhanced microsatellite map of diploid *Fragaria*. *Theor. Appl. Genet.* **112**, 1349–1359.
- Sargent, D.J., Marchese, A., Simpson, D.W., Howad, W., Fernández-Fernández, F., Monfort, A., Arús, P. *et al.* (2009) Development of “universal” gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within *Rosaceae*. *Tree Genet. Genomes*, **5**, 133–145.
- Savini, G., Giorgi, V., Scarano, E. and Neri, D. (2008) Strawberry plant relationship through the stolon. *Physiol. Plant.* **134**, 421–429.
- Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., Jaiswal, P. *et al.* (2011) The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* **43**, 109–116.
- Soltis, P.S., Marchant, D.B., Van de Peer, Y. and Soltis, D.E. (2015) Polyploidy and genome evolution in plants. *Curr. Opin. Genet. Dev.* **35**, 119–125.
- Sønsteby, A. and Heide, O.M. (2008) Long-day rather than autonomous control of flowering in the diploid everbearing strawberry *Fragaria vesca* ssp. *semperflorens*. *J. Hortic. Sci. Biotechnol.* **83**, 360–366.
- Spigler, R.B., Lewers, K.S., Johnson, A.L. and Ashman, T.L. (2010) Comparative mapping reveals autosomal origin of sex chromosome in octoploid *Fragaria virginiana*. *J. Hered.* **101**, S107–S117.
- Tennessen, J.A., Govindarajulu, R., Ashman, T.-L. and Liston, A. (2014) Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biol. Evol.* **6**, 3295–3313.
- Vision, T.J., Brown, D.G., Shmoys, D.B., Durrett, R.T. and Tanksley, S.D. (2000) Selective mapping: a strategy for optimizing the construction of high-density linkage maps. *Genetics*, **155**, 407–420.
- Weebadde, C.K., Wang, D., Finn, C.E., Lewers, K.S., Luby, J.J., Bushakra, J., Sjulín, T.M. *et al.* (2008) Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry. *Plant Breed.* **127**, 94–101.
- Wickland, D.P. and Hanzawa, Y. (2015) The *FLOWERING LOCUS T/TERMINAL FLOWER 1* gene family: functional evolution and molecular mechanisms. *Mol. Plant.* **8**, 983–997.
- Zeng, Z.B. (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc. Natl Acad. Sci. USA*, **90**, 10972–10976.

Zeng, Z.B. (1994) Precision mapping of quantitative trait loci. *Genetics*, **136**, 1457–1468.

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., Zhang, J. *et al.* (2015) Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* **33**, 531–537.

### Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Figure S1** Position on the *F. vesca* subsp. *bracteata* pseudochromosome 4 (Tennessee *et al.*, 2014) of the markers linked to perpetual flowering in the octoploid strawberry developed in this study (in bold), in Castro *et al.* (2015) and in Honjo *et al.* (2016) (in italic).

**Table S1** Description of the 15 and 11 female and male bins of the four linkage groups (LGs) IV in the octoploid *Fragaria* reference map.

**Table S2** Number of recombinants according to their population of origin, genotype and phenotype. Individuals were fine mapped between Bx089\_196 and Bx064\_216 SSRs. Phenotype was recorded as Perpetual Flowering (PF) or Seasonal Flowering (SF).

**Table S3** List of 79 cultivars or selections tested for presence (1) or absence (0) of two SSR markers that flanked the *FaPFRU* locus. Pedigree, year of cultivar release, country of origin, and year of sampling are included.

**Table S4** List of the SSR primer pairs used in this study.