

DGAL/SDQSPV/BSPIC

N° du projet : C-2017-05

Titre du projet : « GENOLAVANDE » Développement d'outils génomiques permettant la caractérisation et la sélection des variétés de lavande



Organisme chef de file : iteipmai

Nom et organisme du chef de projet : iteipmai, Berline FOPA FOMEJU

Site internet sur lequel les résultats du projet peuvent être consultés : www.iteipmai.fr

A – Format du PEI-AGRI sous format électronique (*fiche modèle jointe*)

Document joint en annexe.

B – Compte rendu technique détaillé (*à transmettre par courrier accompagné des annexes*)

I – Les modalités d'organisation

Modalités de pilotage :

- **Conventions avec les partenaires recevant des financements du CASDAR**

La convention entre le ministère de l'Agriculture et de l'Alimentation et l'iteipmai a été signée le 06 décembre 2017.

Les conventions Interpartenariales ont été signées entre l'iteipmai et :

L'INRAE Versailles-Grignon, le 03 juillet 2018

Le CRIEPPAM le 18 décembre 2017

Le protocole d'accord a été quant à lui signé le 24 avril 2018

- **Accords de propriété intellectuelle**

Aucun accord de propriété intellectuelle n'a été signé dans le cadre de ce projet

- **Accords spécifiques au projet (ex : relations avec des entreprises, des organisations professionnelles)**
Pas d'accords spécifique
- **Comités de pilotage : dates, participants, teneur des discussions et accords obtenus sur les réorientations éventuelles du projet**
4 comités de pilotage ont été mis en place dans le cadre du suivi de ce programme. Y étaient conviés les représentants des structures partenaires impliquées dans le programme, des représentants professionnels ainsi que des experts extérieurs, représentants de la Recherche Publique. Ces comités de pilotage ont été organisés :
 - Le 5 mars 2018
 - Le 31 janvier 2019
 - Le 13 février 2020
 - Le 3 février 2021

Les comptes rendus, incluant la liste des participants de ces réunions, figurent en annexe de ce document. De même, les supports de présentation utilisés lors des réunions sont disponibles en annexes.

Outre ces Comités de Pilotage, le responsable du projet organisait régulièrement des points téléphoniques avec les partenaires afin de s'assurer du bon suivi des différentes actions.

- **avenant obtenu : motifs et modalités**

La situation pandémique exceptionnelle vécue en 2020, avec une période confinée au printemps, a obligé l'iteipmai à demander une prolongation du projet de 9 mois (jusqu'au 30 septembre 2021). L'avenant est en date du 04 août 2020.

- autres modalités de pilotage (groupe de travail, comité technique...)
- **Calendrier comparatif prévu/réalisé et analyse des écarts**

Les tâches liées à la production de matériel végétal ainsi qu'à leur phénotypage se sont globalement bien déroulées conformément aux prévisions. En revanche, les tâches liées au génotypage des plantes et aux analyses subséquentes ont pris du retard pour deux raisons principales :

- Le choix d'utiliser le même outil de génotypage pour les actions 1 et 2, validé lors du COPIL du 31/01/2019 a entraîné un décalage sur la réalisation des tâches 2.1.2 et 2.1.3 par rapport au prévisionnel
- Les confinements successifs et les conséquences de la pandémie de covid 19 (difficulté d'accès aux consommables de laboratoire) ont lourdement impacté et retardé la réalisation des tâches 1.5, 1.6, 2.1.2, 2.1.3 et 2.2.3

Tâche	Objectif	2018	2019	2020	2021
Axe 1	Caractérisation de la collection de lavande				
1.1	Création de la collection d'étude	■			
	<i>Tâche 1.1 : réalisé</i>	■			
1.2	Clonage et mise en place multi-site de la collection d'étude	■	■	■	
	<i>tâche 1.2 : réalisé</i>	■	■	■	
1.3	Phénotypage de la collection d'étude "adulte"	■	■	■	
	<i>tâche 1.3 : réalisé</i>	■	■	■	
1.4	Phénotypage des collections d'étude "jeunes"			■	
	<i>tâche 1.4 : réalisé</i>			■	
1.5	Génotypage de la collection d'étude		■	■	
	<i>tâche 1.5 : réalisé</i>		■	■	
1.6	Analyse d'association phénotype / génotype			■	■
	<i>tâche 1.6 : réalisé</i>			■	■
Axe 2	Développement d'outils méthodologiques				
2.1	Construction d'une carte génétique de référence				
2.1.1	Autofécondations	■	■	■	
	<i>Tâche 2.1.1 : réalisé</i>	■	■	■	
2.1.2	Génotypage des descendants	■	■	■	■
	<i>Tâche 2.1.2 : réalisé</i>	■	■	■	■
2.1.3	Carte génétique de Maillette		■		■
	<i>Tâche 2.1.3 : réalisé</i>		■		■
2.2	Mise au point d'une méthode de croisements contrôlés				
2.2.1	Etude de la pollinisation de la lavande fine	■	■		
	<i>Tâche 2.2.1 : réalisé</i>	■	■		
2.2.2	Développement de la méthode de croisement contrôlés	■	■	■	
	<i>Tâche 2.2.2 : réalisé</i>	■	■	■	
2.2.3	Validation des croisements contrôlés par génotypage		■	■	■
	<i>Tâche 2.2.3 : réalisé</i>		■	■	■
2.2.4	Rédaction de la procédure de croisement retenue			■	■
				■	■
		réalisé	iteipmai	EPGV	iteipmai et CRIEPPAM

▪ **Tableau de bord de suivi des moyens mis en œuvre**

Des comptes-rendus intermédiaires ont été établis à la fin de chaque année de travaux. Ils étaient constitués d'un document scientifique, mais aussi d'un tableau financier et d'un tableau de suivi des temps passés, permettant d'estimer l'avancée des travaux. Ces dossiers ont été transmis au ministère en temps voulu.

▪ **Les modalités d'évaluation**

L'évaluation s'est opérée au sein des comités de pilotage où en fonction des résultats présentés par les opérateurs, les travaux ont été discutés voire réorientés si besoin.

II - Les partenariats :

▪ **le rôle et l'apport de chaque partenaire**

Le rôle de chaque partenaire est précisé dans le tableau « Calendrier comparatif prévu/réalisé » présenté ci-dessus.

▪ **bilan du fonctionnement des partenariats : aspects positifs, difficultés rencontrées et points à améliorer**

Le partenariat entre l'iteipmai, le Criepmam et l'EPGV dans ce projet fait suite à un précédent partenariat identique dans le cadre du projet GENOPARFUM (lauréat CASDAR IP 2015-2017). Cet antécédent a facilité la communication entre les différents partenaires. La pandémie a cependant quelque peu rendu les échanges et le suivi du projet plus compliqué dans les premiers mois de l'année 2020, ce qui a été rétabli par la suite grâce à la prise en main d'un ensemble d'outils de communication et gestion de projet à distance.

III – Le déroulement du projet

Action 1 : Caractérisation d'une collection de population de lavande représentative de la diversité de l'espèce et recherche d'associations phénotype / génotype

- *rappel des objectifs attendus figurant dans le dossier déposé et accepté par le jury*

L'objectif de cet axe est d'identifier, à partir d'une collection d'étude, des variations du génome impliquées dans la variation de caractères d'intérêt : le rendement en huile essentielle et la qualité de l'huile essentielle (profil phytochimique), ainsi que la tolérance au dépérissement.

- *méthodes de travail utilisées*

Afin d'atteindre l'objectif cité ci-dessus, l'action 1 a été divisée en 6 tâches :

1. Construire une collection d'étude

Une collection de lavandes (sauvages et cultivées) avait été constituée lors du programme GENOPARFUM. Cette collection était présente à Mévouillon (site de montagne sèche à 900 m d'altitude dans les Baronnies (26)). La collection avait été séquencée et analysée afin d'évaluer l'étendue de la diversité génétique. Sur la base de ces résultats, une collection d'étude sera créée afin d'avoir une taille réduite de collection qui capture la diversité génétique disponible chez les lavandes. Cette collection d'étude réduite pourra ainsi plus facilement être multipliée et phénotypée.

2. Clonage et mise en place d'un essai multisites

Afin de consolider les résultats d'associations phénotype-génotype, il est prévu d'installer la collection d'étude sur différents sites aux conditions climatiques contrastées. Pour cela, les individus sélectionnés pour intégrer la collection d'études seront étiquetés à Mévouillon, des boutures seront prélevées sur ces pieds. Après une phase en pépinière à l'iteipmai, il est prévu d'installer les boutures à raison d'1 exemplaire par clone et par site dans deux nouveaux sites : à la station iteipmai de Montboucher / Jabron (26), dans la Vallée du Rhône et la station Criepmam de Manosque (04), à proximité des plateaux de Valensole et d'Albion. Les 2 sites retenus ont été choisis pour (i) leur proximité des zones de production, (ii) leur variabilité en termes de terroir et (iii) leur complémentarité avec le site de Mévouillon (26). Les plantations sur ces nouveaux sites sont prévus à l'automne 2018.

3. Phénotypage de la collection de Mévouillon (collection « adulte »)

La collection d'étude en place sur le site de Mévouillon (26) fera l'objet de phénotypage durant 2 étés consécutifs (2018 et 2019) sur un nombre de descripteurs assez exhaustif en lien avec les caractères suivants : (i) la tolérance au dépérissement, (ii) la productivité en huile essentielle et (iii) la qualité de l'huile essentielle. L'analyse des données des deux premières années de phénotypage permettra d'identifier les descripteurs phénotypiques les plus discriminants (car avec une bonne variation dans la collection). Cela permettra d'établir un catalogue réduit de descripteurs à utiliser pour les phénotypages des 3 sites à réaliser en 2020. Outre les caractères directement observables au champ, la qualité des lavandes sera évaluée. L'analyse sera réalisée par extraction des composés volatils dans l'acétate d'éthyle, suivi d'une chromatographie par GC FID. Cette méthode, tout en permettant la quantification des composés odorants qui est corrélée à la teneur en huile essentielle, permet aussi de s'affranchir des éventuels artefacts de distillation.

4. Phénotypage des collections « jeunes »

En 2020, les descripteurs pertinents issus de la tâche 1.3 seront utilisés sur les 2 sites plantés en 2018 pour décrire les collections multisites et obtenir la caractérisation de la collection d'étude dans les différents environnements.

5. Génotypage de la collection d'étude

La collection d'étude sera génotypée avec un set de marqueurs de type SNP (Single Nucleotide Polymorphism (polymorphisme d'une base)). Ces SNP seront issus des données de séquençage produites dans le projet GENOPARFUM et étant identifiés comme suffisamment polymorphe dans la collection de lavandes. Le choix des SNP se fera sur la base de :

- leur positionnement sur le génome de la lavande : un panel de 300 SNP sera sélectionné sur la base de leur répartition homogène sur la carte génétique construite en tâche 2.1.4. Ces marqueurs seront des marqueurs « neutres » car non sélectionnés pour leur localisation sur des gènes candidats impliqués dans les caractères d'intérêt retenus dans le projet. Ils permettront de contrôler la structuration de la collection d'étude, paramètre important à intégrer dans les modèles statistiques d'association génotype/phénotype ;
- leur positionnement sur des séquences candidates : le catalogue de gènes de lavande créé dans GENOPARFUM (unigene) sera comparé aux données de la bibliographie pour identifier les séquences de gènes connus pour être impliqués dans la résistance aux stress biotiques et dans la synthèse des terpènes. Ainsi, un deuxième panel de 1200 SNP positionnés sur ces séquences candidates sera ainsi sélectionné. Une annotation structurale des gènes candidats sera réalisée afin de raisonner le choix des 1200 marqueurs en fonction de leur position dans des exons ou des introns. Par ailleurs, il sera possible de sélectionner les SNP en fonction de leur impact sur l'intégrité de la séquence protéique (mutations non synonymes).

6. Analyse d'association phénotype/génotype par une approche gène-candidat

Les données de phénotypage générées en tâche 1.3 et 1.4, ainsi que les données de génotypage générées en tâche 1.5 seront analysées pour rechercher des associations phénotype/génotype. Les analyses d'associations seront réalisées indépendamment pour chaque caractère étudié et une méta-analyse permettra d'évaluer la co-localisation des loci identifiés. Un modèle mixte sera utilisé afin de tenir compte de la structuration et de l'apparement mesurés dans la collection d'étude. Les fausses associations seront également corrigées en appliquant une correction statistique (FDR). L'identification des allèles positivement associés aux caractères étudiés sera effectuée pour leur utilisation dans des programmes de croisement via l'identification des meilleurs parents. Les sets de SNPs significativement associés aux caractères ciblés constituent le livrable de cette tâche.

- *organisation mise en place, pour le chef de file et chaque réalisateur : travail réalisé, moyens humains, matériels et financiers mobilisés*

Les tâches 1.1, 1.2, 1.3 et 1.6 ont été réalisées à l'iteipmai par la chef de projet, une chargée d'expérimentation qui a suivi les collections d'études tout au long du projet, ainsi que des stagiaires (un par an en 2018, 2019 et 2020) et personnel saisonnier pour le phénotypage des collections).

La tâche 1.4 a été réalisée par l'iteipmai et le Crieppam. Chaque structure avait un binôme chargé d'expérimentation/stagiaire pour suivre et réaliser le phénotypage des collections jeunes. Du personnel saisonnier a pu occasionnellement apporter une aide pour la réalisation de cette tâche.

La tâche 1.5 a été réalisée par l'EPGV, avec l'aide d'un CDD iteipmai (février-décembre 2019) encadré par l'EPGV. De plus, le recours à un prestataire de service (Plateforme GENTYANE, Clermont-Ferrand) a été fait pour les extractions ADN haut-débit.

La tâche 1.6 a été réalisée par l'iteipmai.

- *étapes/actions et calendrier comparatif prévu/réalisé et analyse des écarts*

Cette action s'est déroulée de janvier 2018 à septembre 2021 (cf. détail par tâche dans le calendrier comparatif page 3). Comme expliqué ci-dessus, un retard a été pris sur la réalisation des tâches 1.5 et 1.6 du fait de la pandémie de COVID-19.

○ résultats obtenus

1.1 Construire une collection d'étude

Les résultats de l'analyse de diversité génétique (publiés dans la revue Innovations Agronomiques, INRAE, 2020, 79, pp.147-158. (10.15454/1vvd-sd23)).

Les résultats d'analyse des distances génétique sont résumés dans la **Figure 1**. Les données ont permis d'observer que les populations de lavandes étaient davantage structurées par leur origine géographique que par l'altitude à laquelle se situait la population. De plus, deux erreurs d'étiquetage ont été mises en évidence dans la collection adulte de Mévouillon (par exemple, 00-LA-316 et 00-LA-396 portant deux noms différents mais ayant 100 % de similitude génétique → hypothèse la plus probable est que 396 ait été transformé en 316, d'autant plus de 00-LA-316 n'existe pas dans les dossiers du CNPMAI, fournisseur du matériel végétal).

Globalement, la distance inter-population était peu importante. L'indice de différenciation F_{st} moyen était de 0.12 (ce qu'on peut grossièrement traduire par 12 % de distance génétique en moyenne entre 2 populations prises dans la collection d'origine), avec une valeur minimale de 0 et une valeur maximale de 0.27.

Ainsi, la stratégie pour la constitution de la collection d'étude a été de privilégier la conservation du plus grand nombre possible d'individus à l'intérieur d'une même population. De plus, une attention particulière a été apportée au fait de bien inclure toutes les populations ayant F_{st} supérieur à 0.12 en comparaison aux autres populations.

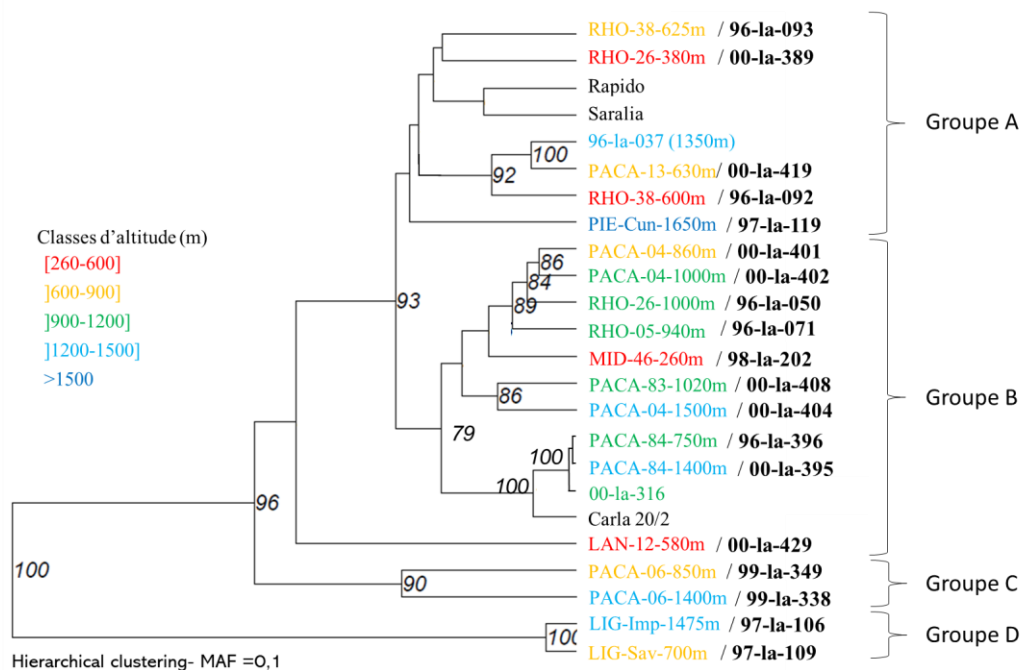


Figure 1 : Dendrogramme construit à partir des valeurs de F_{st} entre les 24 populations de lavande. Lorsque c'était possible, les populations ont été recodées avec leurs données géographiques : « région-département-altitude », / suivi de leur nom de population attribué par le Cnpmai – organisme ayant collecté ces populations dans la nature entre 1996 et 2000.

Rapido, Saralia et Carla 20/2 correspondent à des variétés population (sélection iteipmai).

Les valeurs comprises entre 70 et 100 situées au niveau des nœuds du dendrogramme indiquent la probabilité de l'organisation des populations. Par exemple, la probabilité que les populations 96-la-106 et 97-la-109 du groupe D soient très proches génétiquement et différentes des autres populations étudiées est de 100 %.

1.2 Clonage et mise en place d'un essai multisites

Au total, 282 numéros étaient en effectifs suffisant (ie, au moins 2 boutures vivantes) pour être installés sur les deux sites (Figure 2). A ceux-là s'ajoutent 32 clones issus des populations qui n'ont pu être plantés que sur le site de La Vesc et 31 clones uniquement présents sur le site près de Manosque, portant à un peu plus de 310 le nombre d'individus des collections jeunes. Afin de compléter ces collections, 38 clones issus de la collection étudiée dans GENOPARFUM et 16 clones parentaux des variétés populations de l'iteipmai ont également été plantés sur les deux sites.

Nom	Taille pop Mévouillon	Nb boutures La Vesc	Nb boutures Manosque	Nom	Taille pop Mévouillon	Nb boutures LaVesc	Nb boutures Manosque
00 LA 395 PACA-84-1400	40	14 +1	14 +2	00 LA 389 RHO-26-380	13	5 +2	5
99 LA 338 PACA-06-1400	40	1 +2	1	00 LA 419 PACA-13-630	40	19	19 +2
00 LA 429 LAN-12-580	34	1	1	97 LA 106 LIG-imp-1475	41	21	21
Saralia	40	25 +2	25 +2	96 LA 093 RHO-38-625	40	27 +2	27 +2
98 LA 202 MID-46-260	7			97 LA 119 PIE-cun-1650	40	9 +2	9 +2
97 LA 109 LIG-sav-700	40	1 +4	1	00 LA 401 PACA-04-860	40	13 +1	13 +2
96 LA 050 RHO-26-1000	40	8	8	16 LA 037	40	16	16 +4
99 LA 349 PACA-06-850	16	13	13	CARLA 20/2	40	4 +4	4 +1
00 LA 402 PACA-04-1000	40	5 +3	5 +3	RAPIDO	42	11 +3	11 +4
96 LA 071 RHO-05-940	40	1 +2	1	00 LA 316	40	19 +2	19
96 LA 092 RHO-38-600	40	18	18 +1	00 LA 404 PACA-04-1500	41	16 +1	16 +4
00 LA 396 PACA-84-750	40	13	13 +1	00 LA 408 PACA-83-1020	40	11	11 +1

Figure 2 : Liste et effectifs des collections jeunes installées à La Vesc et à côté de Manosque.

Les populations sont renseignées avec leur nom (attribué par le Cnpmai) suivi en dessous d'un code indiquant « région-département-altitude » de localisation de la population. Pour les boutures installées à La Vesc et Manosque, sont précisés :

En gras, nombre d'individus installés sur les deux sites car au moins deux boutures disponibles par individu.

En italique (+xx) : nombre d'individus de la population installés sur un seul site, faute de boutures disponibles.

En rouge : populations les plus génétiquement distantes des autres (prioritaires dans la collection d'étude)

1.3 Phénotypage de la collection de Mévouillon (collection « adulte »)

Les phénotypages de 2018 et 2019 ont permis de construire un catalogue de descripteurs qui est disponible en annexe A-5. La liste des descripteurs validés est la suivante :

- **Pour la vigueur de la plante** : Le port, la hauteur et le diamètre, la précocité de reprise végétative à la sortie de l'hiver, la précocité de floraison en été.
- **Pour l'état sanitaire de la plante** : la notation de symptômes de dépérissement (échelle de 1 à 4).
- **Pour le rendement en huile essentielle** : la biomasse fraîche à la récolte, estimation de la quantité de fleurs par pied.
- **Pour la qualité de l'huile essentielle** : profils chromatographiques sur extraits dans de l'acétate d'éthyle.

1.4 Phénotypage des collections « jeunes »

En 2020, les collections ont été phénotypées à l'aide des descripteurs présentés dans l'annexe A-6. Les phénotypages de 2020 ont permis d'identifier :

- la bonne capacité du plant à reprendre sa végétation au printemps est un bon indicateur de la vigueur et de l'état sanitaire du plant. En effet, les plants à fort reprise végétative au printemps sont moins dépérissants en été.

Les résultats de 2020 ont montré que les clones de population couvraient une large diversité de phénotypes. Peu de populations appartiennent à un unique et même groupe, ce qui semblerait souligner que la diversité intra-population ne serait pas moins importante que la diversité inter-population, excepté pour quelques populations et pour quelques descripteurs. Les clones appartenant aux populations 00-la-401, 00-la-404, 00-la-408, 00-la-419, 96-la-87, 97-la-106, 97-la-119, 99-la-349 et les variétés RAPIDO et SARALIA seraient des individus plutôt à floraison précoce, tandis que les clones appartenant aux populations 00-la-396, 00-la-93 et la variété MAILLETTE seraient des individus à floraison tardive. Concernant le volume du plant et donc le rendement, il est difficile d'établir des groupes d'individus homogènes puisque les plants dépendent des conditions locales et de l'espace dont ils disposent pour se développer.

Concernant l'état sanitaire des essais, la parcelle de Mévouillon, a été touchée par une forte mortalité. Certaines notations ne sont pas représentatives d'une population de lavandes en conditions « normales ». La notation de certains descripteurs s'est donc montrée inutile, comme le pourcentage de floraison, la biomasse sèche, ou encore le nombre de calices. Cette forte mortalité a contribué à la diminution de la diversité génétique de la collection de Mévouillon, empêchant éventuellement d'observer de nouvelles modalités. Cependant, les analyses ont permis de mettre en évidence qu'une notation de la reprise végétative pouvait être un bon indicateur de l'état sanitaire du plant : en effet, les plantes avec une forte reprise végétative au printemps ont été notées les moins dépérissantes en été (Figure 3).

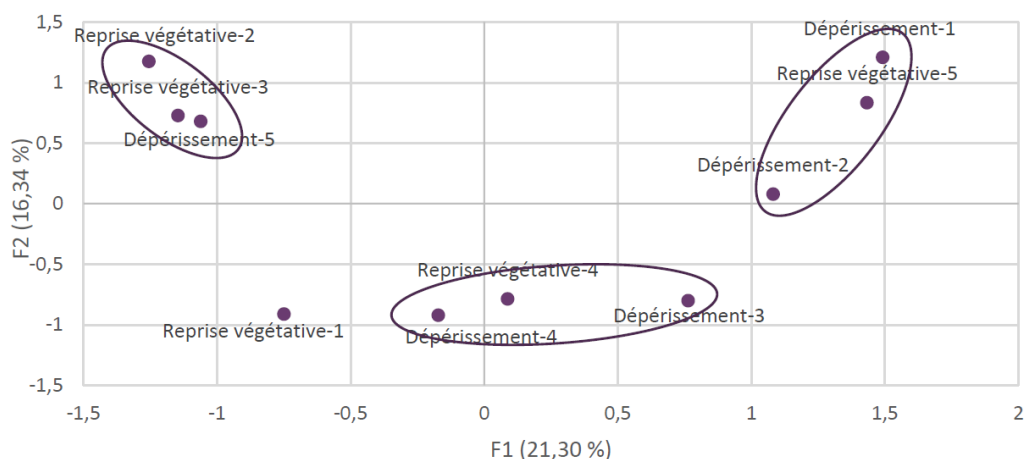


Figure 3 : Graphique des modalités pour les variables reprises végétatives et dépérissement.

Reprise végétative notée de 1 (très faible) à 5 (très forte). Dépérissement noté de 1 (plant non dépérissant) à 5 (plant très dépérissant quasi mort)

1.5 Génotypage de la collection d'étude

L'identification et la sélection des SNP a été réalisée par Boris Demenou (février 2019-décembre 2019), CDD sur le projet Genolavande, Les étapes de la sélection des SNP sont détaillées dans le compte-rendu technique intermédiaire de 2019. Brièvement, les étapes ont été les suivantes :

1. **Identification de gènes candidats associés aux caractères d'intérêt** : afin d'identifier des gènes dont la fonction a été démontrée comme impliquée dans l'expression des caractères qui nous intéressent (qualité et rendement huile essentielle, stress biotique, stress abiotique), des recherches bibliographiques ont été effectuées dans les bases de données d'articles en ligne comme Pubmed,

2. Recherche des homologues des gènes candidats chez la lavande en comparant les séquences candidates à l'unigène construit dans GENOPARFUM
3. Sélection d'autres séquences de lavande pour définir des SNPs répartis sur l'ensemble du génome, de manière homogène et aléatoire, notamment pour la tâche de construction de la carte génétique.
4. Détection de SNP sur les séquences d'intérêt sélectionnées (candidates et neutres)
5. Design et validation du panel de sondes de capture (baits) pour répondre aux spécifications NEB pour le design du panel de baits,

La sélection finale comptait 3765 SNP présents dans les gènes candidats et des gènes bien distribués sur l'ensemble du génome (Figure 4). Après contrôle et design des amorces par New England Biolabs, seuls 3716 SNP ont été retenus dans le panel final, ce qui représente tout de même 85 % de SNP en plus que les 2000 initialement espérés au moment de la construction du projet.

Un prélèvement foliaire a été réalisé sur les 258 individus issus de la collection de Mévouillon (population sauvage + clones parentaux des variétés population iteipmai) et toujours vivants à la date du prélèvement (août 2018). Les feuilles ont été mises à sécher sur billes de silicagel. Une fois sèches, 20 à 40 mg de matière végétale par individu ont été conditionnés dans des plaques 96 et envoyés pour extraction ADN sur la plateforme INRA GENTYANE, à Clermont-Ferrand. Les ADNs extraits ont finalement été envoyés à l'EPGV pour le génotypage. L'annexe A-7 reprend l'ensemble des individus génotypés dans le cadre du projet. Le livrable a été une matrice de génotypage contenant 231 individus et 3552 marqueurs ayant été génotypés avec succès. Ainsi entre le design des amorces et la matrice de génotypage finale, seul 4 % des SNPs ont été « abandonnés » et 30 % de la collection de Mévouillon.

Catégorie de gène	Nombre de gènes identifiés	Nombres de SNPs positionnés sur les gènes
Gène candidat impliqué dans la tolérance aux stress abiotiques	280	562
Gène candidat impliqué dans la tolérance aux stress biotiques	97	133
Gène candidat impliqué dans la production d'huile essentielle	68	361
Gène « neutre » distribué aléatoirement sur le génome	1379	2709
Total	1824	3765

Figure 4 : La répartition des SNP sur les séquences candidates et neutres

1.6 Analyse d'association phénotype/génotype par une approche gène-candidat

En premier lieu, la matrice de génotypage obtenue à l'issue de la tâche 1.5 a été nettoyée pour enlever les données (marqueurs ou individus) avec trop de données manquantes, pouvant générer des biais dans l'analyse d'association. Finalement, 202 individus et 3150 marqueurs ont été conservés. Au total, 6 modalités ont été considérées pour les données de phénotypage :

- 2018 – Mévouillon
- 2019 – Mévouillon
- 2020 – Mévouillon
- 2020 – La Vesc
- 2020 – Revest-du-Bion

Et la modalité BLUP (obtenue avec le package R statgenGxE) qui permet de résumer les données obtenues dans les 5 précédents environnements.

Ensuite, la structuration de la population a été évaluée avec les marqueurs cartographiés sur la carte génétique de lavande (tâche 2.1.3) (Figure 5). Globalement, les individus issus d'une même population sont regroupés ensemble. De plus, comme dans GENOPARFUM, on arrive à la conclusion que la structuration des lavandes est liée à leur origine géographique (lavandes provençales génétiquement distantes des lavandes du Nord de l'Italie et du parc du Mercantour). Par ailleurs, les variétés populations et autres clones sélectionnées par l'iteipmai et les membres du réseau PPAM constituaient un groupe génétique distinct, ce qui est cohérent avec le fait que ces variétés ont été obtenues à partir de sélections dans une base génétique restreinte.

Les données de phénotypage ont été analysées en se focalisant sur les données quantitatives qui sont les seules à pouvoir être traitées par une approche de statistiques en génétique d'association. Cela inclut les variables notées avec des classes (de 1 à 5 par exemple comme pour le dépérissement). Les variables ne suivant pas une distribution gaussienne ont été transformées avec une loi logarithmique en base 10 (\log_{10}). L'analyse de l'héritabilité des caractères phénotypés (Figure 6) a permis de mettre en évidence que globalement, le profil en huile essentielle (i.e la proportion de tel ou tel terpène) dans les COV émis est un caractère très héritable, contrairement au nombre de fleurs, dépérissement ou la masse sèche qui ont des valeurs d'héritabilité entre 0.3 et 0.55 et sont donc beaucoup plus dépendants de l'environnement.

Finalement, l'analyse d'association a été réalisée à l'aide du logiciel GAPIT. Au total, 1 à 20 associations ont été identifiées par caractère étudié (Figure 7). Dans certains cas, le BLUP n'a pas permis d'identifier de nouvelles associations, dans d'autres cas elles ont permis de confirmer des précédentes associations (trouvées dans les analyses par environnement) ou d'en identifier de nouvelles. Certaines associations semblent être de bons candidats à utiliser en sélection, notamment celles situées sur des séquences classées comme gènes candidats pour le caractère suivi : par exemple, les 26 marqueurs associés identifiés avec les caractères COV et précocité de floraison, qui sont situés sur des séquences identifiées comme impliquées dans la synthèse de l'huile essentielle dans l'action 1.5.

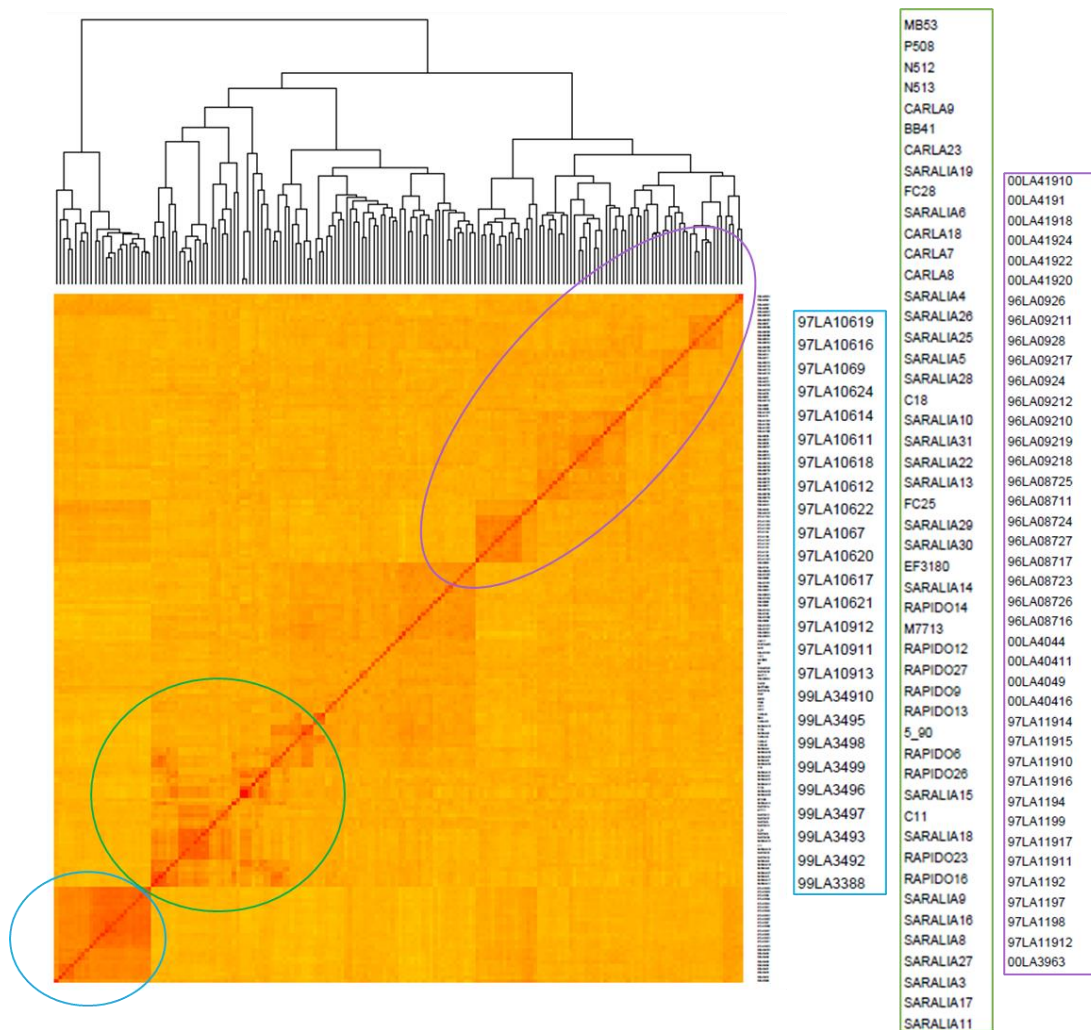


Figure 5 : Matrice d'apparentement entre les individus de la collection adulte

Représentation graphique des distances génétiques entre les individus, calculées à partir des données de génotypage auprès de 400 marqueurs SNP cartographiés sur la carte génétique de la lavande.

En haut de la figure, la représentation graphique est un dendrogramme et en bas une heatmap. La heatmap est une matrice de corrélation (i.e on retrouve la même liste d'individus en ligne et en colonne). Plus un carré est rouge, plus la paire d'individus à l'intersection ligne/colonne est génétiquement proche, plus le carré est jaune, plus les individus sont distants.

Les cercles bleu, vert et violet indiquent les groupes génétiques les plus proches, avec le détail des individus qui les constituent présentés sur la droite de la figure.

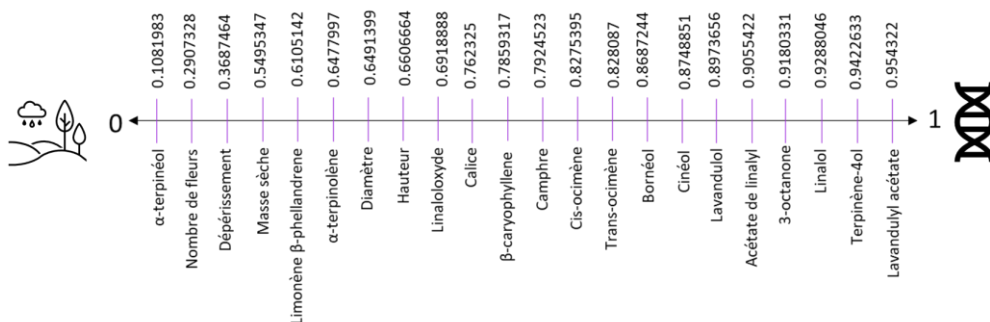


Figure 6 : Héritabilité au sens large calculées pour l'ensemble des variables étudiées dans la tâche 1.6

	2018 Mévouillon	2019 Mévouillon	2020 Mévouillon	2020 La Vesc	2020 Revest	BLUP	Total
3-octanone (log)	1	1		1		3	6
a-terpinéol (log)					2	2	4
a-terpinolène (log)	4	1				3	8
acétate de linalyle (log)	2	8		7		3	20
b-caryophyllene (log)	6						6
bornéol (log)			2				2
camphre (log)			1				1
cinéol (log)			1				1
cis-ocimène (log)			1			7	8
lavandulol (log)	3	2		3		6	14
lavandulyl acétate (log)	3	4	4	5	1	7	24
limonene b-phellandrene (log)					22		22
linalol (log)	3	5	2	1		1	12
linaloloxyle (log)							0
terpinène-4-ol (log)						2	2
trans-ocimène (log)	5			2		5	12
masse sèche (log)	4					5	9
Nombre de fleurs (log)	4		3	1			8
dépérissement		1	3	10			14
diamètre plant						1	1
hauteur plant							0
hauteur calice					1		1
Précocité de floraison					14		14
							189

Figure 7 : Nombre d'associations identifiées pour les différents caractères étudiés et selon les différentes modalités (environnements)

De la même façon, les associations identifiées pour le même caractère dans plusieurs environnements sont de bons candidats. Par exemple, le marqueur GS236_3656 identifié dans 5 des 6 modalités étudiées comme étant associé à la teneur en linalol. Enfin, dans certains cas nous avons identifié des marqueurs en association avec deux caractères, comme les marqueurs GS5698_1081 en association avec les teneurs en camphre et en bornéol. La liste des marqueurs associés identifiés est présentée en annexe A-8.

Enfin, pour les marqueurs représentant les meilleurs candidats comme cité ci-dessus, nous avons recherché les allèles favorables au marqueur dans le panel d'étude (cf Annexe A-8). Ainsi, ces allèles pourront être criblés dans les programmes de sélection pour recherches des individus les moins bons.

- *indicateurs de suivi*
→ cf tableau à suivre
- *indicateurs de résultats*
→ cf tableau à suivre
- *indiquer le pourcentage de réalisation des actions*
→ cf tableau ci-dessous

Tâche	Objectif	Actions	Livrables et indicateurs de suivi	Indicateurs d'évaluation	Pourcentage réalisation
Axe 1	Caractérisation de la collection de lavande	Phénotype et génotyper une collection d'étude	Association des marqueurs SNP avec des caractères agronomiques	Bonne corrélation génotype / phénotype	100%
1.1.	Création la collection d'étude	Choisir les accessions les plus diversifier au sein de la collection Génoparfum	Collection maintenue par clonage	Bonne diversité génotypique	100%
1.2.	Clonage et mise en place multisite de la collection d'étude	Implanter et maintenir un dupliqua de la collection d'étude sur 3 sites	3 collections d'étude identiques implantées sur 3 terroirs différents	Bonne traçabilité des accessions et variété des terroirs	100%
1.3.	Phénotypage de la collection d'étude "adulte"	Identifier les descripteurs pertinents pour décrire la collection d'étude sur la collection la plus âgée	Liste des descripteurs utiles avec échelle de notation	Nombre de descripteurs	100%
1.4.	Phénotypage des collections d'étude "jeunes"	Utiliser les descripteurs choisis pour décrire les 3 collections	Catalogue des descriptions de la collection d'étude	Pertinence de l'échelle de notation pour décrire la variabilité des terroirs	100%
1.5.	Génotypage de la collection d'étude	Génotyper toutes les accessions de la collection grâce aux marqueurs SNP de la carte génétique de référence	Génotypage de la collection d'étude	Variabilité des SNP d'une accession à l'autre	100%
1.6.	Association génotypes / phénotypes	Associer les marqueurs SNP et les descriptions indépendamment de l'effet terroir	Liste de correspondance entre les SNP et les caractères d'intérêt	Bonne corrélation entre les marqueurs et les caractères Evaluation de l'effet terroir sur l'expression des caractères	100%

Action 2 : Développement d'outils méthodologiques

- *rappel des objectifs attendus figurant dans le dossier déposé et accepté par le jury*

Afin de pouvoir appliquer les marqueurs identifiés dans l'action dans un programme de sélection assistée par marqueurs, en vue de l'amélioration variétale des lavandes, il convient de (i) construire une carte génétique pour avoir une première vue linéaire et globale du génome afin de suivre la ségrégation des allèles favorables dans les schémas de croisements et (ii) en amont de cela, disposer de méthode de croisements dirigés pour construire les schémas de sélection. Cet axe sera donc divisé en 2 tâches elles-mêmes divisées en sous tâches.

- *méthodes de travail utilisées*

2.1 Etablissement d'une carte génétique du génome de la lavande

2.1.1 *Autofécondation*

Au démarrage du projet, la technique du croisement contrôlé n'était pas maîtrisée. En revanche, l'utilisation de bourdons en conditions insect-proof permettait de s'assurer de l'autofécondation d'une variété clonale cultivée. Cette technique sera utilisée pour obtenir une descendance issue de l'autofécondation de la lavande Maillette. Maillette est une variété élite déjà utilisée comme dans Génoparfum. La détection de SNP effectuée dans ce précédent projet avait mis en évidence un taux d'hétérozygotie de 0.50, indiquant qu'il était possible de

s'attendre à une ségrégation dans une descendance de Maillette, permettant ainsi la réalisation d'une carte génétique. Les graines ainsi obtenues seront semées et cultivées pour constituer une descendance qui sera limitée à 200 individus.

2.1.2 Choix des SNP pour établir la carte génétique de Maillette

Il était prévu de choisir un set de 1000 SNP pour établir une carte génétique, à partir de la liste des 200 000 SNP trouvés hétérozygotes chez Maillette sur les séquences de l'unigène. Finalement, compte-tenu de l'outil de génotypage identifié pour l'action 1, il a été décidé en 2^{ème} année de projet d'utiliser le même outil de génotypage pour l'ensemble du projet, offrant ainsi la possibilité de génotyper la descendance de Maillette avec plus de 3000 SNP.

2.1.3 Génotypage de la descendance à l'aide des marqueurs SNP et construction d'une carte génétique de Maillette

A notre connaissance, il n'existe pas de carte génétique de la lavande. Dans cette tâche nous souhaitons construire une carte génétique pour avoir une représentation graphique du génome de la lavande avec la répartition de marqueurs SNP et faire une première estimation de la décroissance du déséquilibre de liaison en fonction des distances génétiques. Ces informations seront également utilisées pour sélectionner 300 marqueurs distribués de manière homogène sur le génome pour évaluer la structuration du panel d'étude pour la génétique d'association et pour estimer la résolution de la génétique d'association réalisée en tâche 1.6. Avec le génotypage de 200 individus on attend une résolution de carte de 1 SNP/0.5cM.

Des échantillons foliaires ont été prélevés sur les descendants au stade plantule, ainsi que les plantes parentales. Les échantillons ont été mis à sécher dans des sachets krafts sur billes de silicagel. Des pesées ont été réalisées pour normaliser les échantillons secs avant extraction des ADNs. Ainsi, 20 mg de chaque descendant et parent a été pesé et conditionné en plaque de 96 puits. Les plaques ont été envoyées sur la plateforme INRAe Gentyane (Clermont Ferrand), où une extraction robotisée a été réalisée. Une fois extraits, les ADNs ont été envoyés à l'équipe EPGV pour le génotypage.

L'outil de génotypage contenant les 3 716 marqueurs SNPs cibles a été utilisé. Le génotypage par séquençage des SNP a généré des reads de 100 paires de bases pour chaque individu et chaque marqueur. Différents filtres de qualité sont appliqués pour ne conserver que les données les plus « solides » :

- un seuil de profondeur minimal est établi à 5,
- la qualité de génotypage $GQ > 20$
- un filtre sur la fréquence des allèles pour les individus hétérozygotes dont le seuil appliqué est : $0.2 < AF < 0.8$.

Seuls les marqueurs ayant passé l'ensemble des filtres sont conservés pour la construction de la carte.

Trois outils de bio-informatiques gratuits (LepMAP3, Package R | OneMap et Package R | R/qtl) ont été utilisés pour construire des cartes génétiques de lavande. La méthode de construction est assez similaire pour les trois logiciels. Tout d'abord, une étape de tri qualitatif des données est effectuée, suivie d'un test de ségrégation. Ensuite, les groupes de liaison sont formés avec uniquement les marqueurs de bonne qualité. Une fois les groupes de liaisons formés, les marqueurs sont ordonnés dans les groupes pour construire les cartes finalisées. Cependant, chacun des outils varie plus ou moins sur les algorithmes sous-jacents à ces étapes. Ainsi, la comparaison des résultats obtenus avec les différents outils permet d'identifier les marqueurs cartographiés avec « robustesse » statistique.

2.2 Mise au point d'une méthode de croisement contrôlé de la lavande

2.2.1: Etude de la pollinisation de la lavande

Une étude bibliographique sur la biologie florale et la pollinisation de la lavande et plus généralement des labiées permettra de déterminer le meilleur stade de prélèvement du pollen, de castration et de pollinisation sur lavande. Cette étude sera réalisée par l'iteipmai et permettra de proposer un ou plusieurs protocoles de croisement contrôlé de la lavande.

Différents paramètres ont été étudiés afin de mieux comprendre la floraison de la lavande et tester des méthodes de croisements contrôlés :

- Sur l'autofécondation et la geitonogamie en plaçant une lavande variété Diva (castrée ou non) au milieu de lavande Maillette et en présence de bourdons
- Sur l'autogamie en plaçant des épis sous sachets d'autofécondation avant la période de floraison
- Sur les stades de floraison en observant l'ouverture des fleurs, l'ouverture des sacs polliniques, la viabilité du pollen à l'aide de carmin acétate, le développement et la réceptivité du pistil à l'aide d'eau oxygénée et la cinétique de floraison d'un épi.

2.2.2 Développement de la méthode de croisement contrôlé

Des croisements d'expérimentation seront mis en place sur la base des données recueillies en tâche 2.2.1. Ils permettront d'obtenir des graines supposées interfécondées qui seront ensuite semées. Les clones choisis pour l'interfécondation seront Diva et Maillette, deux clones élites utilisés dans GENOPARFUM pour lesquels des SNP (~200000 sur chaque clone) ont été détectés. Plusieurs méthodes de castration ont été testées lors de la floraison :

- 1/ on retire simplement la corolle
- 2/ suite au retrait de la corolle, on découpe le haut du calice
- 3/ coupe franche : on découpe le haut du calice

De la même manière, plusieurs modes de pollinisation ont été testés :

- pollinisation manuelle à 24h, 48h et 72h après castration
- pollinisation une seule fois ou jusqu'à 36 fois
- pollinisation médiée par insectes (mouche, bourdon ou osmie)

2.2.3 Validation du croisement contrôlé par génotypage

Il était prévu de choisir un set de SNP réduit (<96 SNP), pour contrôler les croisements. Finalement, compte-tenu de l'outil de génotypage identifié pour l'action 1, il a été décidé en 2^{ème} année de projet d'utiliser le même outil de génotypage pour les deux actions, offrant ainsi une économie de coût sur le génotypage des quelques descendant issus des croisements contrôlés. En effet, il existait une économie d'échelle avec l'outil de génotypage NEB sélectionné et moins il y avait d'échantillons à génotyper, plus cela était cher.

2.2.4: Rédaction de la procédure

En fonction des résultats du génotypage, le meilleur des protocoles sera validé. Cette tâche, conduite par l'iteipmai, a pour but de produire une procédure de pollinisation destinée à être utilisée en routine pour les futurs croisements de lavande et lavandins.

- *organisation mise en place, pour le chef de file et chaque réalisateur : travail réalisé, moyens humains, matériels et financiers mobilisés*

Les tâches 2.1.1, 2.2.1, 2.2.2 et 2.2.4 ont été réalisées à l'iteipmai par deux chargées d'expérimentation aidées ponctuellement de stagiaires.

Les tâches 2.1.2 et 2.2.3 ont été réalisées par l'EPGV, avec l'aide d'un CDD iteipmai (février-décembre 2019) encadré par l'EPGV. De plus, le recours à un prestataire de service (Plateforme GENTYANE, Clermont-Ferrand) a été fait pour les extractions ADNs haut-débit.

La tâche 2.1.3 a été réalisée par l'iteipmai et l'EPGV avec l'aide d'une stagiaire de 4 mois.

- *étapes/actions et calendrier comparatif prévu/réalisé et analyse des écarts*

Cette action s'est déroulée de janvier 2018 à septembre 2021 (cf. détail par tâche dans le calendrier comparatif page 3). Un retard a été pris sur la réalisation des tâches 2.1.3 et 2.2.3 du fait de la modification de la stratégie de génotypage en année 2 pour bénéficier d'un génotypage plus dense et à coût moindre, puis le retard a été pris en 2020 du fait de la pandémie de COVID-19. Les autres tâches de l'action se sont déroulées conformément au calendrier prévisionnel.

- *résultats obtenus*

2.3 Etablissement d'une carte génétique du génome de la lavande

2.3.1 *Autofécondation*

Le livrable attendu pour cette tâche était les 200 accessions issues de l'autofécondation de Maillette. Finalement, ce sont 220 descendants viables qui ont été obtenus. Les individus, ainsi que les clones parentaux de Maillette ayant servi aux autofécondations ont été prélevés en juillet 2019 (prélèvements foliaires) pour les extractions ADN et génotypages prévus en tâche 2.1.3. L'annexe A-7 reprend l'ensemble des individus génotypés dans le cadre du projet.

2.3.2 *Choix des SNP pour établir la carte génétique de Maillette*

L'outil de génotypage utilisé a été le même que celui présenté dans les résultats de la tâche 1.5.

2.3.3 *Génotypage de la descendance à l'aide des marqueurs SNP et construction d'une carte génétique de Maillette*

Le génotypage des ADN a été réalisé sur 215 descendants et 10 des 20 clones parentaux Maillette. Ce choix a été fait par l'Iteipmai afin de réduire le nombre d'individus à séquencer, d'autant plus qu'une variabilité inexistante à faible était attendue entre les différents parents. Les résultats obtenus par l'EPGV ont indiqué un manque d'informations de génotypage chez certains individus (principalement chez les parents), très certainement liée à une moins bonne qualité des ADN. Ceci s'est traduit par une quantité insuffisante de reads lors de l'alignement des séquences d'ADN et donc une difficulté à déterminer le génotype des individus concernés. Du fait du modèle génétique utilisé (famille F2), les logiciels de cartographie ne requéraient que deux individus identifiés comme parents de la descendance. Ainsi, les deux parents ayant le plus d'informations génétiques ont été conservés pour la suite des analyses : parent_AF3 et parent_AF4.

Au niveau des marqueurs moléculaires, l'EPGV a identifié 148 SNPs parmi les 3716 dépourvus de données de génotypage. Par ailleurs, un excès d'hétérozygotie a été observé : en moyenne 85,6 % des SNPs étaient hétérozygotes chez les individus issus de l'autofécondation (AF).

Finalement, la matrice de génotypage utilisée pour la construction de la carte comprenait :

- Les meilleurs individus et parents (soit au total 206 AF et 2 parents Maillette)
- Et les marqueurs hétérozygotes chez les parents (soient 1947 marqueurs), seuls marqueurs pour lesquels on s'attend à observer une ségrégation dans la descendance.

Pour la construction des cartes génétiques, dans l'ensemble on constate qu'il y a le même nombre de groupes de liaisons obtenus avec les logiciels (Figure 8). A noter cependant que pour Rqtl nous avons de base 59 groupes de liaisons proposés, mais au-delà du groupe de liaison 25, les groupes de liaisons restants n'avaient plus que 2 marqueurs positionnés, ce qui était peu pertinent à conserver.

Les différents logiciels ont globalement cartographié un nombre de marqueurs équivalent, R/qtl étant celui qui en ait positionné le plus. Toutefois, le nombre de marqueurs en position unique est assez variable. Le logiciel R (autant avec qtl qu'Onemap) a positionné l'ensemble des marqueurs a des sites uniques alors que la carte mâle construite avec LepMAP3 n'a que 56% de ses marqueurs positionnés à des sites uniques. C'est-à-dire qu'une part significative des marqueurs co-localisent sur cette carte.

Cartes	LepMap M	LepMap F	OneMap	Rqtl
Nombre de Groupes de Liaison	25	25	25	25 affectés
Nombre de marqueurs	425	425	414	461
Position unique des marqueurs	238	311	414	461
Taille de la carte (cM)	452,98	1583,34	1741,00	2883,23
Distance minimale (cM)	0	0	0,0001	0,0000001
Distance maximale (cM)	16,534	48,558	72,646	80,865
Médiane (cM)	0,98	0,4875	1,959	2,2
Ecart-type (cM)	7,532	2,403	7,124	11,16
Moyenne de fréquence des SNPs (cM)	1,13	3,96	4,476	6,61

Figure 8 : Statistiques de bases sur les données de chacune des cartes créées.

On a observé que la carte de plus grande taille a été obtenue avec le logiciel R/qtl (2883.23 cM) et la carte la plus petite correspondait à la carte femelle de LepMAP3 (452.98 cM). La carte femelle de LepMAP3 et la carte de OneMap ont quant à elles présenté des tailles comparables autour de 1600-1700 cM. Initialement, il était attendu en génotypant avec les 3000 marqueurs de pouvoir en positionner environ 1000 avec une densité d'un SNP tous les 0,5 centiMorgan (cM). Or, seuls 414 à 425 SNPs ont pu être répartis sur tout le génome de la lavande avec une densité de marquage d'un SNP tous les 1.13 cM à 6.61 cM, même si les valeurs médianes comprises entre 1 cM et 2 cM sont un résultat davantage encourageant vis-à-vis de la qualité des cartes obtenues.

Pour la suite de la comparaison des cartes nous n'avons pris en compte que la carte de femelle de LepMAP3. En effet, les deux cartes générées par cet outil présentaient exactement les mêmes marqueurs et dans le même ordre. La seule différence résidait dans les distances entre les marqueurs cartographiés. La densité de marquage ayant déjà été traitée dans les paragraphes précédents et dans la mesure où l'objectif, ici, était de comparer l'ordonnement des marqueurs le long des groupe de liaison, cette carte mâle a été écartée pour la suite des analyses.

Ainsi, les cartes de LepMap et OneMap et R/qtl partageaient 381 SNPs en commun (Figure 9). Ainsi, 82 %, 89 % et 92 % des marqueurs positionnés sur les cartes Rqtl, LepMAP et OneMap, respectivement, ont été communs aux 3 outils.

Ce ratio est plutôt satisfaisant ainsi, peu de marqueurs ont été cartographiés mais leurs positionnements semblent « robustes ».

Résultats LepMap			Résultats ONEMAP			Résultats Rqtl		
FONCTION	SNP_ABIOTIC1	43	SNP_ABIOTIC1	37	SNP_ABIOTIC1	45		
	SNP_ABIOTIC2	33	SNP_ABIOTIC2	34	SNP_ABIOTIC2	37		
	SNP_SCO	176	SNP_SCO	174	SNP_SCO	188		
	SNP_AUTRES	115	SNP_AUTRES	114	SNP_AUTRES	127		
	SNP_EO1	19	SNP_EO1	18	SNP_EO1	20		
	SNP_EO2	7	SNP_EO2	8	SNP_EO2	9		
	SNP_BIOTIC1	12	SNP_BIOTIC1	11	SNP_BIOTIC1	13		
	SNP_BIOTIC2	13	SNP_BIOTIC2	11	SNP_BIOTIC2	13		
	SNP_MxD	1	SNP_MxD	1	SNP_MxD	3		
	SNP_17EO2BIOTIC	6	SNP_17EO2BIOTIC	6	SNP_17EO2BIOTIC	6		
Localisation	Exon	57	Exon	58	Exon	68		
	Intron	34	Intron	30	Intron	33		
	NSP	40	NSP	35	NSP	40		
		0	294		0	320		

Résultats 3 cartes			Légende : (Mk = marqueurs)
FONCTION	SNP_ABIOTIC1	36	
	SNP_ABIOTIC2	30	- SNP_ABIOTIC2 = Mk sur gène abiotique & dupliqué dans le génome (2)
	SNP_SCO	157	- SNP_SCO = Mk sur gène communs à Embryophyta & simple copie
	SNP_AUTRES	106	- SNP_AUTRES = Mk répartis sur le génome d'une espèce apparentée <i>Scutellaria baicalensis</i>
	SNP_EO1	17	- SNP_EO1 = Mk sur gène de production d'huiles essentielles & simple copie (1)
	SNP_EO2	6	- SNP_EO2 = Mk sur gène de production d'huiles essentielles & dupliqué dans le génome (2)
	SNP_BIOTIC1	11	- SNP_BIOTIC1 = Mk sur gène biotique & simple copie (1)
	SNP_BIOTIC2	11	- SNP_BIOTIC2 = Mk sur gène biotique & dupliqué dans le génome (2)
	SNP_MxD	1	- SNP_MxD = Mk diagnostique pour le croisement contrôlé Maille x Diva (pour distinguer les descendants)
	SNP_17EO2BIOTIC	6	- SNP_17EO2BIOTIC = Mk sur gène HE ou biotique qui servent de marge pour atteindre les 3700 SNPs pour la cartographie
Localisation	Exon	53	
	Intron	29	
	NSP	33	
		0	266
Marqueurs en commun		381	

Figure 9 : Nombre de SNP cartographiés pour chaque carte et pour les marqueurs en communs

En continuant dans l'analyse, nous avons cherché à déterminer quels types de SNPs avaient été positionnés (Figure 9). La majorité des SNPs cartographiés (68 à 70 % selon les cartes) correspondent à des séquences sélectionnées pour être réparties aléatoirement et de manière homogène sur l'ensemble du génome de la lavande (séquences provenant de la comparaison à la base de données BUSCO notées « SCO » ou séquences issues de la comparaison au génome de *S. baicalensis* notées « autres»). Ce résultat est cohérent dans la mesure où ces SNPs représentent 70 % des SNPs compris dans l'outil de génotypage.

Les 30 % restants des SNPs cartographiés sont positionnés sur des séquences candidates, proportionnellement au nombre de SNPs initiaux compris dans l'outil de génotypage (13-15 % de SNPs liés au stress abiotique, 9-10 % de SNPs liés à la synthèse de l'huile essentielle, 3-4 % des SNPs liés à la synthèse des terpènes).

Finalement, seul environ 100 des 400 marqueurs positionnés disposaient d'une information de localisation dans la partie intronique ou exonique de la séquence de gène.

Dans l'ensemble, les groupes de liaisons ont montré des similitudes sur le positionnement des marqueurs (colinéarité observée) (Figure 10). Cependant, quelques inversions locales concernant des marqueurs voisins ont été constatées (Annexe 3). Deux ruptures de colinéarité plus importantes ont été observées. Le groupe de liaison 25 établi par Rqtl ne présentait aucune équivalence avec les groupes de liaison construits par les deux autres outils. Par ailleurs, le groupe de liaison 2 formé par LepMap correspondait à la somme des groupes 2 et 24 formés par OneMap et R/qtl. Cette situation décrite permet d'expliquer le manque de cohérence pour le groupe 24.

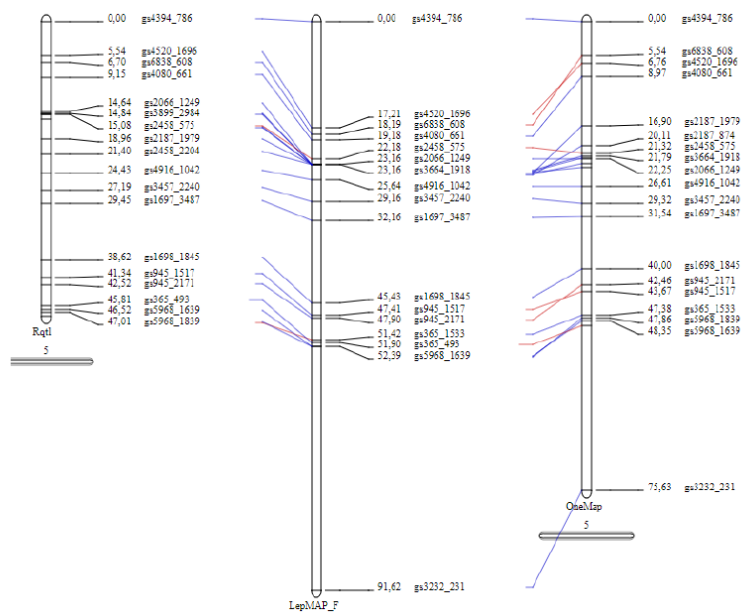


Figure 10 : Comparaison de 3 cartes pour le groupe de liaison 5

2.4 Mise au point d'une méthode de croisement contrôlé de la lavande

2.4.1: Etude de la pollinisation de la lavande

La synthèse des travaux sur les stades de floraison de la lavande fine est présentée Figure 11. Ces résultats confirment la protandrie des fleurs de lavande.

iteipmai

Stades de floraison et de développement des organes reproducteurs de la lavande vraie (*Lavandula angustifolia*)

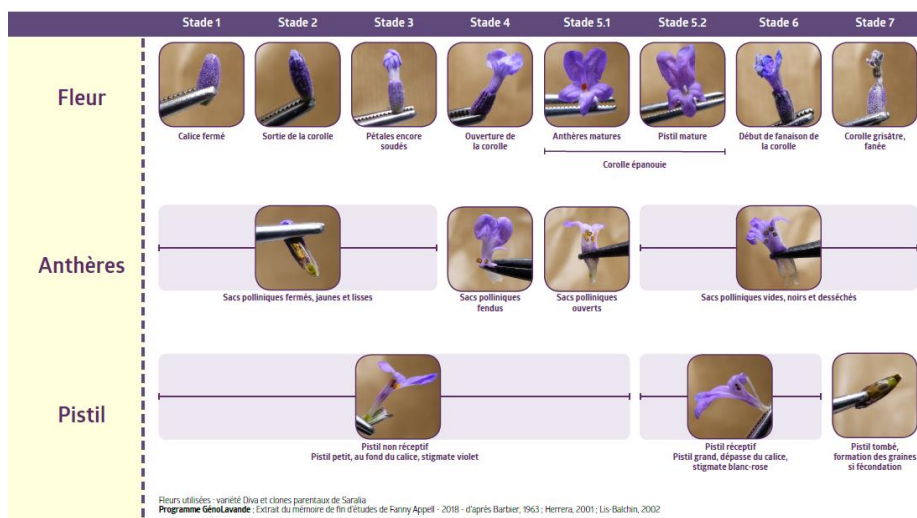


Figure 11 : Poster illustrant les stades de floraison et de développement des organes reproducteurs

Par ailleurs, les travaux permettant d'identifier les meilleurs stades de castration et de pollinisation ont permis de déterminer que :

- Les méthodes consistant au retrait de la corolle à la pince suivi de la coupe du haut du calice lorsque les fleurs sont au stade de développement 3, correspondent à la meilleure méthode de castration.
- le suivi du développement du pistil (Figure 12) a permis de déterminer une plage optimale pour la pollinisation pour les fleurs au stade 2, la phase la plus propice à la pollinisation se trouve entre 51 et 120 heures après la pollinisation. Au stade 3, la

période optimale de pollinisation se trouve entre 30 et 80 heures après la castration ; au-delà de 145 heures, il n'y a quasiment plus de pistils réceptifs pour la pollinisation.

Les optimisations réalisées en 2021 ont permis d'affiner ces recommandations :

- On obtient plus de graines en pollinisant une seule fois 72 h après castration. Un résultat moindre mais quand même intéressant est obtenu en pollinisant 2 fois, 48 h et 72 h après castration. La modalité de pollinisation une fois 48 h après castration a définitivement été écartée.
- Le sens de pollinisation a un impact sur le succès de la fécondation : en effet, plus de graines étaient obtenus lorsque Diva était la mère et Maillette le père.

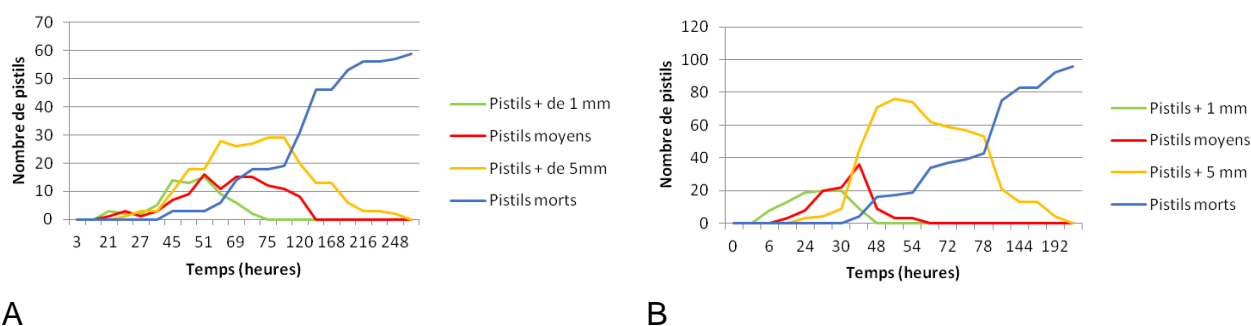


Figure 12 : Graphique de l'évolution des pistils sur des fleurs castrées et pollinisées au stade 2 (Fig. A) ou a stade 3 (Fig. B)

2.4.2: Développement de la méthode de croisement contrôlé

L'ensemble des tests et optimisations réalisés en tâche 2.2.1 ont permis de réaliser une série de croisement. Les graines obtenues ont été mises en pépinière à la station iteipmai de Chemillé afin de réaliser des tests en génotypage. Au total, 45 descendants viables ont été obtenus à l'issue des 3 campagnes de croisements.

Les résultats de ce projet ont été utilisés dans un autre projet de l'institut, SELAV 2, démarré en juin 2021 qui est un programme d'amélioration génétique lavande lavandin. L'utilisation du protocole de croisement à plus grande échelle y est utilisée pour générer des descendances plus importantes (400 lavandes issues de croisement avec castration obtenues en une seule campagne).

2.4.3: Validation du croisement contrôlé par génotypage

A ce jour, seul 15 des 45 descendants obtenus avec les différentes méthodes de croisement ont été génotypés. Ainsi, sur cette tâche ; seuls 33 % des travaux ont été réalisés. Cela est dû au retard pris sur le génotypage : les plaques ADNs contenant les descendants de croisements n'ont pas pu être toutes génotypées dans le temps du projet.

Par ailleurs, pour les 15 individus génotypés, la quantité de données manquantes était très importante (>80 % des 3716) et les marqueurs présentant des données de génotypage n'étaient pas des marqueurs diagnostiques (monomorphes entre Diva et Maillette). La mauvaise qualité de génotypage est très probablement liée à la qualité des ADNs extraits qui n'étaient pas bonne. En perspective, il est prévu de réanalyser ces échantillons dans le cadre du projet d'amélioration génétique lavande/lavandin SELAV initié à l'iteipmai en juin 2021, en y associant les 35 individus restants et les 400 individus générés dans SELAV.

2.4.4: Rédaction de la procédure

Le meilleur protocole de croisement a été validé et présenté dans le compte-rendu technique intermédiaire de 2020 (Annexe A-10). Ce protocole a été utilisé avec succès dans le cadre du projet d'amélioration génétique lavande/lavandin SELAV initié à l'iteipmai en juin 2021.

- *indicateurs de suivi*
→ cf tableau à suivre

- *indicateurs de résultats*
→ cf tableau à suivre
- *indiquer le pourcentage de réalisation des actions*
→ cf tableau ci-dessous

Tâche	Objectif	Actions	Livrables et indicateurs de suivi	Indicateurs d'évaluation	Pourcentage réalisation
Axe 2	Développement d'outils méthodologiques	Construire une carte génétique de référence et maîtriser les croisements contrôlés	Cartes génétiques de référence pour la lavande et procédure de pollinisation contrôlée de la lavande	Bonne répartition des marqueurs SNP sur la carte génétique Obtention d'une descendance hybride	92%
2.1.	Construction d'une carte génétique de référence de la lavande				100.00%
2.1.1.	Autofécondations	Mettre en culture et maintenir 2 descendances issue de l'autofécondation des lavandes Maillette et Diva	Descendances (2)	Nombre de graines Nombre de descendants	100.00%
2.1.2.	Génotypage des descendants	Génotyper 200 individus par descendances grace aux SNP de Génoparfum	Génotypage des descendance	Bonne diversité génétique	100.00%
2.1.3.	Carte génétique de Maillette	Construire de la carte génétique de chacune des variétés parents	Carte génétique de chaque parents avec les marqueurs SNP les plus pertinents	Nombre et pertinence des SNP des cartes génétiques	100.00%
2.2.	Mise au point d'une méthode de croisement contrôlé de la lavande				
2.2.1.	Etude de la pollinisation de la lavande	Réaliser une étude bibliographique sur la biologie florale et la fécondation des labiées et plus spécifiquement des lavandes	Protocole(s) de croisements contrôlés à tester	Nombre de protocoles	100.00%
2.2.2.	Développement de la méthode de croisement	Tester les différents protocoles par le croisement des lavandes Diva et Maillette	Descendance pour chaque croisement contrôlé	Nombre de graine pour chaque croisement Nombre de plants descendants de chaque croisement	100.00%
2.2.3.	Validation du croisement contrôlé par génotypage	Valider la maîtrise de l'interfécondation par analyse des génotypes de quelques descendants issus de chaque croisement	Descendants en terre validée comme issu d'un croisement contrôlé	% de descendants issus de l'interfécondation	33.00%
2.2.4.	Rédaction de la procédure	Compiler les données obtenues	Procédure de croisement contrôlé de la lavande	Intégration de la procédure dans les schémas de création variétale lavande Utilisation de la méthode pour créer des lavandins	100.00%

IV - Les modalités de valorisation du projet

- *les différents modes de diffusions, les diffusions, le public concerné (données chiffrées si possible)*

Les résultats obtenus dans ce projet ont été communiqués à la filière au fur et à mesure de l'avancement des travaux : des comptes-rendus techniques intermédiaires détaillant les méthodes utilisés et les résultats par année sont disponibles aux adhérents de l'institut (cf. annexe A-10 à A-13). Une version résumée, également réalisée tous les ans est librement accessible depuis le site de l'iteipmai (<https://www.iteipmai.fr/nos-publications/articles-publies>)

Les résultats ont également été présentés sur le terrain aux membres de la filière :

 - lors de journées techniques dédiées au programme GENOLAVANDE : le vendredi 18 juin 2021 à Montboucher sur Jabron (matin) et le mardi 29 juin pour Revest du Bion (matin),
 - et lors de journées portes ouvertes de l'iteipmai. Dans ce cadre, des posters de présentation des résultats ont été réalisés (cf. Annexes A-15 à A-17),
 - un article scientifique est en cours de finalisation (cf. Annexe-18) et est prévu pour publication dans le journal *Frontiers in Plants Sciences*.
 -
- *exploitation et valorisation des résultats, de l'expérience acquise*

Les résultats acquis dans l'action 2 ont directement nourri le programme d'amélioration génétique lavande/lavandin du réseau PPAM (iteipmai, Crieppam, Chambre d'Agriculture de la Drôme) porté par l'iteipmai. Les résultats ont notamment permis de réaliser pour la première fois des croisement interspécifiques lavande fine x lavande aspic pour régénérer du lavandin, espèce pour laquelle seuls 3 clones sont actuellement cultivés et pour laquelle il n'existe pas de collection de ressources génétique, faute de lavandin sauvage. Ainsi ces travaux de croisements interspécifiques sont primordiaux. De plus, l'outil de génotypage développé va être utilisé dans ce programme SELAV afin de caractériser les descendants de croisement, notamment aux marqueurs identifiés comme impliqués dans des caractères d'intérêt. Ces génotypages permettront de valider ou non le caractère prédictif/ diagnostique des marqueurs identifiés dans GENOLAVANDE.
- *les transferts prévus*

L'objectif est de pouvoir utiliser l'outil de génotypage en routine dans les programmes de sélection mais également pour les tâches de conservation des ressources génétiques des lamiacées. Des discussions sont en cours auprès de BioGEVES avec Arnaud REMAY pour utiliser l'outil de génotypage pour caractériser la collection de référence de lavande au GEVES.
- *indiquer si un site Internet ou une base de données a été réalisé nom, modalités d'accès, et perspectives de développement et maintien*

Nous n'avons pas réalisé de site internet ou base de données.
- *modules de formations éventuels et OAD : préciser les conditions dans lesquelles ces transferts pourront être réalisés (dates, publics, modèle économique, gratuité...)*

Il n'y a pas eu de module de formation réalisé.
- *valorisation européenne ou internationale*
 - un article scientifique est en cours de finalisation (cf. Annexe) et est prévu pour publication dans le journal *Frontiers in Plants Sciences*.
 - Il est prévu de présenter les résultats obtenus dans le projet lors du Symposium MAP de l'IHC 2022 (International Horticultural Congress) à Angers le 15/08/2022.

V – Les perspectives

▪ *les points forts et les points faibles du projet*

Le projet a réellement permis de développer des innovations pour la lavande :

- le développement d'un outil de génotypage avec un usage relativement transversal et l'identification de gènes candidats pour des critères de sélection clés sont de vrais points forts du projet
- de même, le développement de méthodes de croisements contrôlés, adapté aux croisements contrôlés interspécifiques est une vraie innovation pour le champ de la création variétale lavande/lavandin.
- Il reste à améliorer la carte génétique de la lavande, difficulté qui sera facilement levée grâce à la maîtrise des croisements contrôlés qui permettra d'obtenir des populations de cartographie de meilleure qualité que la descendance de Maillette utilisée dans ce projet (on espère moins de distorsion de ségrégation).

▪ *les difficultés rencontrées*

La difficulté majeure rencontrée vient du retard pris dans les actions de génotypage du fait de la pandémie. De plus, la mortalité importante sur certaines parcelles d'essais n'a pas permis d'évaluer une diversité génétique aussi large que celle qu'on avait imaginé en début de projet.

▪ *les suites envisagées*

Pour aller plus loin dans le développement d'outils génomiques, il est prévu de faire un séquençage d'un génome de référence de lavande. Cela permettrait notamment d'avoir accès à un plus grand nombre de marqueurs.

Nous souhaitons également tester l'outil de génotypage pour identifier des marqueurs associés à la tolérance à la sécheresse. En effet, dans le cadre du projet RECITAL (lauréat Casdar 2016), des indicateurs pertinents de la tolérance au stress hydrique ont été identifiés chez les lavandes, permettant ainsi une caractérisation plus fine des collections et ouvrant ainsi la possibilité à la détection d'association phénotype/génotype.

Les marqueurs identifiés dans ce projet seront testés dans le cadre de SELAV pour valider leur pertinence comme marqueurs diagnostiques du phénotype de descendants issus de croisements.

Par ailleurs, nous envisageons d'intégrer la lavande aspic dans le développement de connaissance et d'outils génomiques afin d'accompagner la création de nouveaux lavandins.

▪ *les modalités envisagées pour pouvoir analyser l'impact du projet sur les différents publics cibles à court et moyen terme*

A long terme, ce projet aura été un succès s'il nous permet de sortir de nouvelles variétés obtenues via un processus de croisement dirigé. A moyen terme, le succès sera mesuré au nombre de nouveau génotypes créés à évaluer, notamment dans le cadre du lavandin où la disponibilité en ressources génétiques sauvages est très restreinte et l'enjeu de sélection très important compte tenu de la part importante de lavandin dans la filière lavandicole.

Annexes

Les annexes suivantes sont à joindre :

- Comptes rendus des comités de pilotage
 - Annexe A-1** CR COPIL de lancement du 5 mars 2018
 - Annexe A-2** CR COPIL du 31 janvier 2019
 - Annexe A-3** COPIL du 13 février 2020
 - Annexe A-4** COPIL du 3 février 2021

▪ **Les principaux livrables**

- Livrable 1.1 - Collection d'étude créée dès le début du projet : livrable présenté dans le texte, page 7
- Livrable 1.2 - Implantation multisite de la collection d'association : livrable présenté dans le texte, page 7
- Livrable 1.3- La liste de descripteurs pertinents : Livrable présenté en **annexe A-5**

- Livrable 1.4 - La description de la collection d'étude pour les 3 sites : Livrable sous forme de rapport de stage présentant le détail du phénotypage 2020 des 3 sites présenté **en annexe A-6**
 - Livrable 1.5- Données de génotypage de la collection d'étude : livrable intégré à l'**annexe A-7** qui reprend l'ensemble du matériel génotypé dans le projet. Les matrices des données de génotypage étant trop lourde à intégrer ici
 - Livrable 1.6 – Liste des SNPs associés aux caractères d'intérêt : livrable présenté en **annexe A-8**
 - Livrable 2.1.1 – Les 200 accessions issus de l'autofécondation : livrable intégré à l'**annexe A-7** qui reprend l'ensemble du matériel génotypé dans le projet. Les matrices des données de génotypage étant trop lourde à intégrer ici
 - Livrable 2.1.2 – Choix des SNPs pour la carte génétique : ce livrable est présenté dans le texte pages 8-9
 - Livrable 2.1.3 – Carte génétique de la lavande Maillette : Livrable sous forme de rapport de stage présentant le détail de la construction de la carte génétique présenté **en annexe A-9**
 - Livrable 2.2.1 - Protocoles de pollinisation des lavandes : ces protocoles sont présentés dans le texte pages 13, 17 et 18.
 - Livrable 2.2.2 – Nombre de descendants issus des croisements contrôlés. livrable intégré à l'**annexe A-7** qui reprend l'ensemble du matériel génotypé dans le projet
 - Livrable 2.2.3 – Données de génotypage de descendants issus des croisements contrôlés. livrable intégré à l'**annexe A-7** qui reprend l'ensemble du matériel génotypé dans le projet, les matrices des données de génotypage étant trop lourde à intégrer ici
 - Livrable 2.2.4 – Procédure de réalisation des croisements contrôlés : livrable intégré au CRT 2020 présenté en **annexe A-10**
- Publications réalisées dans le cadre du projet (articles scientifiques, articles techniques,...)
 - A-11 – Compte-Rendu Technique intermédiaire 2018
 - A-12 – Compte-Rendu Technique intermédiaire 2019
 - A-13 – Compte-Rendu Technique intermédiaire 2021
 - A-14 – Poster présentation Génolavande 2018 sur la biologie florale de la lavande
 - A-15 – Poster présentation Génolavande 2019 sur les distances génétiques entre lavandes
 - A-16 – Poster présentation Génolavande 2021 sur la construction de la carte génétique
 - A-17 --Article scientifique en cours de finalisation
 - Autres supports de communications mentionnant les résultats du projet

○ C – Compte rendu financier

Les pièces à fournir sont les suivantes :

Par le chef de file :

- compte de réalisation consolidé, de l'ensemble du projet, action par action, établi par le chef de file
- le compte du chef de file faisant apparaître le versement aux tiers
- la copie des conventions (et avenants le cas échéant) avec montants des travaux prévus et des subventions CASDAR prévues, établie par le chef de file
- la copie éventuelle de l'accord de propriété intellectuelle du projet

Pour chaque partenaire recevant des financements CASDAR (chef de file + chaque organisme tiers) ou devant effectuer des travaux (même s'ils ne sont pas subventionnés) :

- le compte de réalisation global du projet (signé du commissaire aux comptes ou de l'agent comptable)
- le tableau des agents mobilisés sur le projet (modèle DGAL)
- pour les organismes privés : le rapport du commissaire aux comptes de la structure établi spécifiquement pour le projet en question et qui doit viser le respect de la convention et du guide des modalités financières
- le tableau des dépenses directes
- le tableau des prestations de service (cf décision en comité de pilotage)

- en cas de non assujettissement à la TVA, une attestation du représentant légal de la structure ou de l'administration fiscale (ne pas faire figurer si la convention est établie en HT)
- Recettes obtenues complétant la subvention du CASDAR : source (collectivités, organismes français..., modalités d'obtention, années concernées)
- Octroi de subvention européenne