



## مقاله پژوهشی

# مطالعات ارتباطی جامع ژنوم (GWAS) در برخی از توده‌های بابونه آلمانی (*Matricaria chamomilla* L.)

مهدی قنوتی<sup>۱</sup>، سعدالله هوشمند<sup>۲</sup>، سباستین آلبرشت<sup>۳</sup> و لارس گارنت اوتو<sup>۴</sup>

۱- استادیار، گروه کشاورزی دانشگاه پیام نور، تهران، ایران، (نویسنده مسوول: m\_ghanavati@pnu.ac.ir)

۲- استاد، گروه اصلاح نباتات و بیوتکنولوژی، دانشکده کشاورزی دانشگاه شهرکرد، شهرکرد، ایران

۳- شرکت گیاه دارویی فارماپلنت، آرترن، آلمان

۴- مرکز تحقیقات ژنتیک گیاهی IPK، گتزلین، آلمان

تاریخ دریافت: ۱۳۰۰/۱۲/۱۶ تاریخ پذیرش: ۱۴۰۱/۱۲/۲۷

صفحه: ۴۶ تا ۵۵

### چکیده مبسوط

**مقدمه و هدف:** بابونه آلمانی (*Matricaria recutita*) یکی از قدیمی‌ترین و پر مصرف‌ترین گیاهان دارویی در جهان است. با توجه به اهمیت این گیاه، این مطالعه با هدف تعیین ساختار ژنتیکی با استفاده از ژنوتیپ سنتجی از طریق توالی یابی (GBS) و شناسایی SNP‌های مرتبط با برخی ویژگی‌های مورفوفیزیولوژیکی در بابونه آلمانی انجام شد و سپس مطالعات ارتباطی جامع ژنوم (GWAS) انجام گردید.

**مواد و روش‌ها:** به منظور بررسی تنوع ژنتیکی و طبقه بندی اکتیپ‌های بابونه آلمانی، آزمایشی بر روی ۴۶ اکتیپ ایرانی و خارجی در قالب طرح بلوک کامل تصادفی در چهار تکرار در مزرعه تحقیقاتی شرکت فارماپلنت آلمان انجام شد. سپس با استفاده از توالی یاب Illumina HiSeq 2000/2500 ژنوتیپ سنتجی از طریق توالی یابی (GBS) صورت گرفته و تجزیه داده‌های توالی یابی و کشف SNP‌ها در غیاب یک ژنوم مرجع، به صورت *de novo* توسط پایپ لاین IPyRAD انجام شد. پس از توالی یابی نمونه‌ها، از تکنیک مطالعه ارتباطی در سطح ژنوم (GWAS) جهت بررسی ارتباط بین نشانگرها و صفات مورد مطالعه استفاده شد. در GWAS، به منظور شناسایی SNP‌های مرتبط با صفات مورفولوژی و بیوشیمیایی مختلف از آنالیز منهنن و نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU استفاده شد.

**یافته‌ها:** نتایج آنالیز منهنن و اعتبارسنجی آنها با نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU منجر به شناسایی SNP‌های مرتبط با زمان گلدهی، شدت رنگ سبز کرت، وضعیت انبساط برگ (صافی و زبری)، تراکم بخش علفی گیاه و وضعیت انشعابی شاخه‌های فرعی، قطر گل در گیاه دارویی بابونه آلمانی شد. نشانگرهای ذیل در این مطالعه شناسایی شدند: SNP با مکان ژنی locus\_161754-71 و  $-\log P\text{-value} = 10.35354$  به عنوان مرتب‌ترین نشانگر برای صفت زمان گلدهی، SNP با مکان ژنی locus\_136992-38 و  $-\log P\text{-value} = 6.5950$  به عنوان مرتب‌ترین نشانگر برای صفت تراکم بخش علفی کرت‌ها، SNP با مکان ژنی locus\_35488 و  $-\log P\text{-value} = 8.0247$  به عنوان مرتب‌ترین نشانگر برای صفت تراکم بخش علفی کرت‌ها، SNP با مکان ژنی locus\_135707-30 و  $-\log P\text{-value} = 11.4573$  به عنوان مرتب‌ترین نشانگر برای صفت ارتفاع بوته معرفی، SNP با مکان ژنی locus\_2494-70 و  $-\log P\text{-value} = 5.8867$  به عنوان نشانگر مرتبط با صفت قطر گل، به طور کلی، ساختار ژنتیکی اکتیپ‌های بابونه آلمانی با استفاده از GBS بررسی شد و همچنین SNP‌های مرتبط با برخی صفات مورفوفیزیولوژیکی شناسایی شدند.

**نتیجه‌گیری:** از نشانگرهای SNP شناسایی شده در این مطالعه GWAS، بویژه مواردی که جنبه اقتصادی بالاتری دارند، می‌توان به منظور گزینش به کمک نشانگر در برنامه‌های اصلاحی بابونه آلمانی استفاده کرد.

**واژه‌های کلیدی:** بابونه آلمانی، توالی یابی، GBS، SNP، GWAS

### مقدمه

بابونه آلمانی (*Matricaria recutita*) از پر مصرف‌ترین گیاهان دارویی در جهان است که هر ساله مقادیر فراوانی از اسانس آن در صنایع داروسازی، آرایشی، بهداشتی و صنایع غذایی استفاده می‌شود. منشأ اصلی و مرکز خزانه ژنی این گیاه منطقه مدیترانه به ویژه ایران می‌باشد. عملکرد اقتصادی گیاه بابونه آلمانی، گل‌های آن می‌باشد که حاوی ماده متابولیت‌های ثانویه متفاوتی است از جمله مهم‌ترین آن‌ها پیش‌ساز ماده کامازولن<sup>۵</sup> در ترکیبات فرار موجود در قاعده گلچه‌های لوله‌ای و آپیژنین<sup>۶</sup> موجود در گلچه‌های سفید جانبی می‌باشد (۳۸). علیرغم اهمیت این گیاه دارویی، مطالعات نچندان زیادی بروی ساختار ژنتیکی به منظور حفاظت از ژرم پلاسما آن صورت گرفته است. این در حالی است که اکتیپ‌های بابونه آلمانی موجود در طبیعت به لحاظ داشتن ژن‌های مفید برای مقاومت به تنش‌های محیطی و گسترش سازگاری ژنتیکی در برابر تغییرات محیطی دارای اهمیت می‌باشند. بنابراین، برای استفاده از این منابع، اطلاع از ماهیت و میزان تنوع موجود در ژرم پلاسما، از اهمیت ویژه‌ای برخوردار است (۱۱). به علاوه، آگاهی از تنوع ژنتیکی ژرم پلاسما معیاری مناسب برای بهره‌گیری از آن‌ها در شناسایی و انتقال ژن‌ها در بهبود گیاهان دارویی می‌باشد (۱۲). رویکردهایی که به منظور مطالعه تنوع ژنتیکی مورد استفاده قرار گرفته‌اند متنوع هستند. از جمله آن‌ها می‌توان به بررسی خصوصیات مورفولوژیکی و نشانگرهای مولکولی مثل SNP‌ها اشاره داشت (۱۴). ارزیابی ساختار ژنوتیپی بر اساس توالی یابی (GBS) به عنوان کاربرد جدیدی از فناوری NGS‌ها در شناسایی و تعیین ژنوتیپ جمعیت‌های گیاهی مورد استقبال محققان قرار گرفته است (۱). مطالعات ارتباطی جامع ژنوم (GWAS) نیز با بررسی ارتباط بین هم‌خوانی SNP و خصوصیات مثل صفات زراعی قادر است تا ارتباطات فنوتیپ-ژنوتیپ را به خوبی در مطالعات به‌نژادی آشکار سازد (۱۳). با توجه به اهمیت این گیاه، این مطالعه با هدف تعیین ساختار ژنتیکی با استفاده از GBS و شناسایی SNP‌های مرتبط با برخی ویژگی‌های مورفوفیزیولوژیکی در بابونه آلمانی انجام شد. منابع ژنتیک گیاهی پایه و اساس امنیت غذایی و یکی از ارزشمندترین ذخایر هر کشور محسوب می‌شوند. ایجاد ارقام گیاهی با عملکرد بالا جهت دستیابی به امنیت غذایی منوط به دسترسی به این منابع است. رویکردهایی که به منظور مطالعه تنوع ژنتیکی مورد استفاده قرار گرفته‌اند متنوع هستند. معیارهای تخمین تنوع ژنتیکی می‌تواند شامل صفات مورفولوژیکی یا نشانگرهای مولکولی باشد. با استفاده از نشانگرهای مورفولوژیکی و ژنتیکی می‌توان به بررسی تنوع ژنتیکی و شناسایی اجزای عملکرد پرداخت (۱۸).

نشانگرهای SNP با وضوح بالا یکی از بهترین روش‌ها برای بررسی تنوع ژنتیکی هستند که از تفاوت تک نوکلئوتیدی بین دو توالی DNA ناشی می‌شوند. نشانگرهای SNP نشانگرهای دقیق و مستقیمی هستند که طبیعت واقعی تنوع اللی را نشان داده و در ژنوم گیاهان فراوانی بسیار بالای دارند. البته حذف و اضافه‌ها<sup>۷</sup> نیز به عنوان SNP ارزیابی می‌شوند. نشانگرهای SNP از نظر تکاملی، حفاظت شده بوده و به عنوان نشانگرهای

مرتبط با مکان‌های ژنی کنترل‌کننده صفات کمی در مطالعات نقشه‌یابی ارتباطی استفاده می‌شوند. به دلیل کاربردهای مختلف و بازدهی بالای اطلاعات حاصل از نشانگرهای SNP، از این نشانگر در سال‌های اخیر به عنوان نشانگر غالب در مطالعات بررسی تنوع ژنتیکی و روابط بین افراد، نقشه‌یابی ارتباطی و تجزیه عدم تعادل پیوستگی کل ژنوم در گیاهان استفاده می‌شود (۱۹). ژنوتیپ‌سنتجی از طریق توالی یابی (GBS) به عنوان کاربرد جدیدی از NGS در شناسایی و تعیین ژنوتیپ جمعیت‌های گیاهی مورد استقبال محققان قرار گرفته است (۱). در این تکنیک، استفاده از نشانگرهای مولکولی پلی‌مورفیک تک نوکلئوتیدی (SNP) رایج است (۴). به طور کلی، این تکنیک شامل چند گام به شرح ذیل است: ۱- استخراج DNA از نمونه‌های مورد مطالعه؛ ۲- هضم DNA با آنزیم‌های برشی با محدودکننده؛ ۳- ساخت کتابخانه NGS؛ ۴- توالی یابی با یکی از روش‌های توالی یابی مثل ایلومینا؛ ۴- انطباق توالی به منظور شناسایی SNP‌ها؛ ۵- آنالیز ساختاری اطلاعات (۹). مطالعات ارتباطی جامع ژنوم از نظر دانش ژنتیک، یک ارزیابی سراسری ژنوم بر روی مجموعه‌ای از تنوعات ژنتیکی فردی در افراد گوناگون است (۲۱). هدف این آزمایشات، مقایسه، نتیجه‌گیری و نهایتاً تعیین روابط مشترک، همبستگی و وابستگی، میان وجود یک تنوع ژنتیکی، و بروز یک خصوصیت مشترک در بین دارندگان آن تفاوت ژنتیکی می‌باشد. مطالعات ارتباطی جامع ژنوم عمدتاً بر روی ارتباط بین هم‌خوانی چندشکلی‌های تک-نوکلئوتیدی (SNP) و خصوصیات مثل صفات زراعی، اجزاء عملکردی و غیره متمرکز است اما داده‌های به‌دست آمده می‌تواند برای هر صفت دیگری نیز مورد استفاده قرار گیرد (۷). هم اکنون، استفاده از GWAS در گیاهان، بیوشیمیایی مرتبط با صفات مورفولوژیکی، فیزیولوژیکی، بیوشیمیایی مورد توجه محققان قرار گرفته است.

### مواد و روش‌ها

در این تحقیق، اکتیپ‌های بابونه آلمانی از برخی نقاط مختلف ایران، آلمان و برخی دیگر از کشورها که عمدتاً اروپایی بودند (جدول ۱) جمعاً به تعداد ۴۶ اکتیپ جهت ارزیابی عملکرد زراعی، خصوصیت کیفی و کاربرد آنالیزهای ژنومی جمع‌آوری و استفاده شدند. این مطالعه ژنوتیپی شامل استخراج DNA و توالی یابی در آزمایشگاه موسسه IPK در شهر گتزلین، کشور آلمان در چهار تکرار انجام شد.

#### صفات مورد بررسی در آزمایش:

**زمان گلدهی:** جهت ارزیابی این صفت، وضعیت گل‌های گیاهان کشت شده در هر کرت بررسی شد، اگر میزان گل‌ها به نحوی بود که متوجه گل دهی زودهنگام بوته‌ها می‌شدیم، نمره ۳ را می‌دادیم و اگر متوجه می‌شدیم که تازه گل‌ها تشکیل شده و متوجه گل‌دهی دیرهنگام گل‌ها می‌شدیم، نمره ۵ را می‌دادیم و زمان گل‌دهی حدوداً در نمره ۵ دادیم.

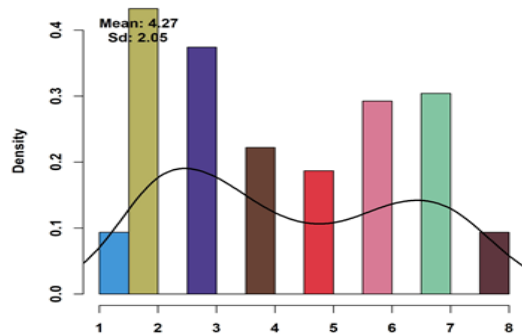
**شدت رنگ سبز در کرت:** جهت ارزیابی این صفت، وضعیت کلی رنگ گیاهان در هر کرت بررسی شد، اگر رنگ هر کرت سبزی کم رنگی داشت

1- Genotyping by Sequencing  
4- Genome wide association studies

2- Single nucleotide polymorphisms  
5- Chamazulen

3- Genome-Wide Association Study  
6- Apigenin 7- Indels

شدند. توزیع زمان گلدهی نشان داد که بیشتر کرت‌ها در دسته زودرس (با نمره‌های ۲ و ۳) قرار دارند. همچنین، در دسته دیررس، بیشتر کرت‌ها با نمره‌های ۶ و ۷ تعیین شدند (شکل ۱). همراستا با نتایج ما، اتو و همکاران (۱۶) نیز در مطالعه ساختار جمعیت بایونیه آلمانی با استفاده از GWAS نشان دادند که جمعیت‌های بایونیه آلمانی اروپایی همراه با تعدادی اکسیسون غیراروپایی از نظر زمان گلدهی در طیف متنوعی از زودگلدهی تا دیرگلدهی قرار دارند و فراوانی زودرس‌ها در آنها بیشتر از سایرین است. آنالیز منهن با دو روش MLM و FarmCPU به ترتیب منجر به شناسایی ۱ و ۵ نشانگر معنی‌دار شد که با صفت گلدهی ارتباط معنی‌داری دارند (شکل ۲). در واقع، تعداد ۵ نقطه (SNP) از حد آستانه بالاتر بودند که از بین آنها، یک نقطه (SNP) وجود داشت که منفی لگاریتم  $P$ -value بالاتری نسبت به بقیه داشت. مشخصات این نشانگرهای سیگنال شده برای صفت زمان گلدهی در جدول ۲ و ۴ آورده شده است. همانطور که در جدول‌های ذیل مشاهده می‌شود، نشانگر مولکولی SNP با مکان ژنی  $locus\_161754-71$  و  $-\log P\text{-value} = 10.35354$  می‌تواند به‌عنوان مرتبط‌ترین نشانگر برای صفت زمان گلدهی معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگرهای شناسایی‌شده با آنالیز منهن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP های بالاتر از خط آستانه (نقاط زرد و آبی رنگ به ترتیب گویای نشانگرهای شناسایی شده با روش-های MLM و FarmCPU)، دارای  $p$ -value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده‌شده و مقادیر موردانتظار برای SNP های واقع روی خط برازش وجود نداشت (شکل ۳). در توافق با مشاهدات ما، اتو و همکاران (۱۶) نیز در مطالعه خود بر روی شناسایی SNP های مرتبط با زمان گلدهی در جمعیت‌های بایونیه آلمانی اروپایی با استفاده از GWAS نشان دادند که پنج SNP مرتبط با زمان گلدهی بایونیه آلمانی وجود دارد که از آنها می‌توان به منظور ایجاد و استفاده از نشانگرهای مولکولی در برنامه‌های گزینش به نشانگر (MAS) استفاده کرد. نویسندگان پیشنهاد کردند که یکی از پنج توالی با یک ژن رمزکننده پروتئین (به نام AtXTH28 یا xyloglucan endotransglucosylase/hydrolase) ارتباط دارد که این ژن به نوبه خود با گلدهی در آرایدیوپسیس مرتبط است (۱۳). نتایج ما در کنار یافته‌های اتو و همکاران (۱۶) نشان می‌دهد که از نشانگرهای همبسته با زمان گلدهی می‌توان در راستای انتخاب ژنوتیپ‌های زودرس برای نواحی با تابستان گرم و خشک و یا انتخاب ژنوتیپ‌های دیررس برای نواحی با زمستان سرد استفاده کرد.



شکل ۱- توزیع زمان گلدهی مربوط به آنالیز GWAS اکوتیپ‌های بایونیه  
Figure 1. Flowering time distribution related to GWAS analysis of chamomile ecotypes

کرت‌ها با میزان کلروفیل برگی و سازوکارهای مقابله با تنش‌های محیطی در ارتباط است و به عبارتی شدت متوسط رنگ کرت‌ها را نتیجه گزینش برای صفاتی از قبیل میزان کلروفیل برگی و سازوکارهای مقابله با تنش‌های محیطی می‌دانند.

آنالیز منهن با روش FarmCPU منجر به شناسایی ۴ نشانگر معنی‌دار شد که با صفت شدت رنگ سبز کرت ارتباط معنی‌داری دارند (شکل ۴). در واقع، تعداد ۴ نقطه (SNP) از حد آستانه بالاتر بودند که از بین آنها، یک نقطه (SNP) وجود داشت که منفی لگاریتم  $P$ -value بالاتری نسبت به بقیه داشت. مشخصات این نشانگرهای سیگنال شده برای صفت شدت رنگ سبز کرت در جدول ۴ آورده شده است.

نمره ۱ دادیم و اگر کرت خیلی تیره بود، نمره ۵ دادیم و اگر رنگ سبز تیره بود، نمره ۳ را لحاظ می‌کردیم.

**وضعیت تقسیمات و انشعابات برگ (صافی و زبری):** جهت ارزیابی این صفت، وضعیت تقسیمات برگ گیاهان در هر کرت بررسی شد، اگر انشعابات و تقسیمات هر برگ از هم فاصله داشت آن را به عنوان زبر و خشن در نظر گرفته و نمره ۳ داده شد و اگر هر برگ انشعابات زیادی داشت که نزدیک به هم بودند، به عنوان نرم شناخته شده و عدد ۷ داده شد و مابین این دو حالت را به عنوان متوسط شناسایی و عدد ۵ دادیم.

**تراکم بخش علفی گیاه:** جهت ارزیابی این صفت، وضعیت برگ گیاهان در هر کرت بررسی شد، اگر برگ‌ها غیر متراکم بودند، نمره ۳ داده شد، اگر برگ‌ها متراکم متوسط داشتند، نمره ۵ و اگر برگ‌ها متراکم بودند نمره ۷ داده شد.

**وضعیت شاخه‌های فرعی:** جهت ارزیابی این صفت، وضعیت انشعاب شاخه‌های گیاهان در هر کرت بررسی شد، اگر شاخه‌ها از پایین منشعب شده بودند، عدد ۱ داده شد و اگر شاخه‌ها از قسمت بالایی منشعب شده بودند، عدد ۵ و حد وسط این دو حالت را عدد ۳ دادیم.

**قطر گل:** بصورت تصادفی در هر کرت قطر ۲۰ گل را با کولیس اندازه گیری کردیم و میانگین آن را بر حسب میلی‌متر گزارش نمودیم.

**ارتفاع گیاه:** از کف زمین تا بالاترین قسمت بوته با خط کش اندازه گیری شد. پس از توالی‌یابی نمونه‌ها، از تکنیک GWAS جهت بررسی و یافتن ارتباط بین نشانگرها و صفات مورد مطالعه و در حقیقت یافتن مکان ژنی صفات مورد نظر استفاده شد. در این بخش هم از مدل متحد مخلوط (MLM) که یکی از مدل‌های روش‌های تک لوکوس، می‌باشد، استفاده شد. در روش تک لوکوس، لوکوس‌ها (SNPs) تک تک بصورت جداگانه وارد مدل می‌شوند. همچنین از روش Multilocus، که این روش اثر تمام مارکرها را همزمان می‌سنجند، نیز استفاده شد، مدل استفاده شده این روش، مدل FarmCPU بود. در آغاز این آنالیز کنترل کیفیت داده‌ها و پیرایش داده‌های ژنوتیپی انجام شد، برای این کار نشانگرهای SNP با نرخ فراوانی کمتر از ۰.۹٪ از مطالعه حذف شده و نشانگرهای باقی‌مانده برای آنالیز GWAS به کار برده شدند. در این راستا، آنالیز GWAS توسط پکیج rMVP در برنامه Rstudio انجام شد. در رابطه با اینکه کدام SNPs سیگنال هستند، از دو روش کیوکیویلات و منهن استفاده شد.

## نتایج و بحث

**زمان گلدهی:** در مطالعه GWAS، به‌منظور شناسایی SNP های مرتبط با صفت زمان گلدهی از آنالیز منهن و نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU استفاده شد. جهت تعیین توزیع صفت، زمان گلدهی کرت‌ها بر اساس شماره ۱ (زودرس‌ترین) تا ۸ (دیررس‌ترین) امتیازدهی

**شدت رنگ سبز کرت:** در مطالعه GWAS، به‌منظور شناسایی SNP های مرتبط با صفت شدت رنگ سبز کرت‌ها از آنالیز منهن و نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU استفاده شد. جهت تعیین توزیع صفت، شدت رنگ سبز کرت‌ها بر اساس شماره ۱ (سبز کم رنگ) تا ۷ (سبز پر رنگ) امتیازدهی شدند. توزیع شدت رنگ سبز کرت نشان داد که بیشتر کرت‌ها در دسته سبز کم‌رنگ تا شدت رنگ متوسط (با نمره‌های ۲ و ۳) قرار دارند (شکل ۴). همراستا با نتایج ما، راتهور و کومار (۱۷) نیز در مطالعه جمعیت‌های بایونیه آلمانی تحت کشت‌وکار در هند نشان دادند که این جمعیت‌ها از نظر شدت رنگ سبز کرت‌ها در طیف متنوعی از سبز کم رنگ تا سبز پررنگ قرار دارند و فراوانی شدت رنگ متوسط در آنها بیشتر از سایرین است. نویسندگان پیشنهاد کردند که شدت متوسط رنگ

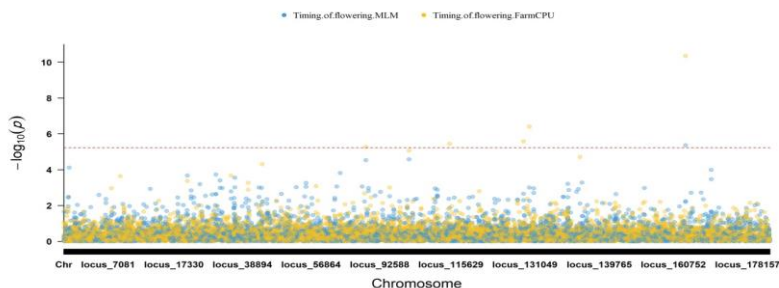
جدول ۱- جمعيت‌هاي بابونه مورد استفاده در پژوهش

Table 1. Chamomile populations used in the research

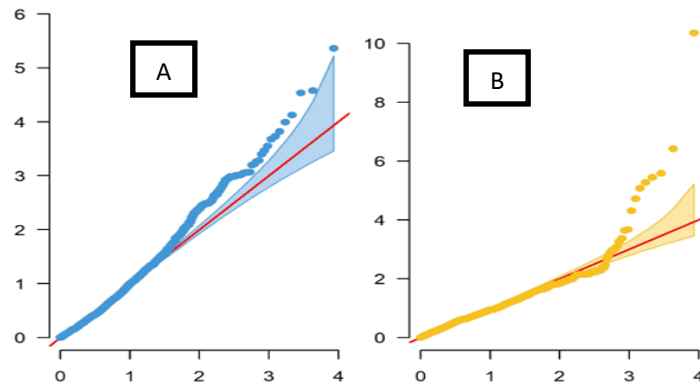
شهرانها	کشور	اکوتیپ	ردیف
نامشخص	اسلواکی	Lutea von Vilora	۱
نامشخص	آلمان	259 1 Zü FNR	۲
نامشخص	هند	HK Indien	۳
گچساران	ایران	MG 23	۴
چرام	ایران	MG17	۵
ایذه	ایران	MG07	۶
دره شهر	ایران	MG38	۷
دزفول	ایران	MG31	۸
نامشخص	کره شمالي	MAT19	۹
	آلمان	Manzana	۱۰
مسجد سلیمان-اندیکا	ایران	MG26	۱۱
خرم آباد	ایران	MG41	۱۲
باشت	ایران	M 30	۱۳
شوشتر	ایران	MG06	۱۴
لهستان	لهستان	ZlotyLan2018 LeGutko	۱۵
چرام	ایران	MG18	۱۶
بهبهان	ایران	MG02	۱۷
چرام	ایران	MG19	۱۸
ایذه	ایران	MG08	۱۹
هفتکل	ایران	MG10	۲۰
اسلواکی	اسلواکی	Bona von Vilora PG18	۲۱
	آلمان	215/08 Son_Standard aus E17 isoFNR	۲۲
گتوند-روستای ایستادگی	ایران	MG28	۲۳
هفتکل-روستای افتادگی	ایران	MG09	۲۴
مسجد سلیمان-اندیکا	ایران	MG25	۲۵
	چک-اسلواکی	MAT15	۲۶
شوشتر-روستای میاناب	ایران	MG05	۲۷
	ایران	MG32	۲۸
چرام-روستای بلدان	ایران	MG16	۲۹
	فرانسه	Ferme de St.Marthe von Rose vPG16	۳۰
	کرواسی	HK Kroatien	۳۱
شیراز	ایران	MG14	۳۲
چرام	ایران	MG15	۳۳
اصفهان	ایران	MG42	۳۴
	ارژانتین	Argenmilla	۳۵
باغملک-میداوود	ایران	MG03	۳۶
ایذه	ایران	MG37	۳۷
	انگلستان	HK Dover	۳۸
	آلمان	CamoFlora	۳۹
	آلمان (مصر)	213/06 SCH_Standard FNRIso E17	۴۰
	ایتالیا	Aromitalien	۴۱
رامهرمز	ایران	MG39	۴۲
دهدشت	ایران	MG29	۴۳
دهنز	ایران	MG40	۴۴
	آلمان (مصر)	Germania	۴۵
زیدون-روستای لنگیر	ایران	MG33	۴۶

آبی رنگ به ترتیب گویای نشانگرهای شناسایی شده با روش‌های MLM و FarmCPU، دارای p-value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده شده و مقادیر مورد انتظار برای SNP‌های واقع روی خط برازش وجود نداشت (شکل ۲).

همانطور که در جدول مشاهده می‌شود، SNP با مکان ژنی locus 38-136992 با داشتن مقدار  $\log P\text{-value} = 6.5950$  می‌تواند به‌عنوان مرتبط‌ترین نشانگر برای صفت شدت رنگ سبز کرت معرفی شود. نمودار QQ-پلات با دو روش FarmCPU و MLM نیز منجر به تایید نشانگرهای شناسایی شده با آنالیز منهتن شد. به عبارتی، QQ-پلات با دو روش MLM و FarmCPU نشان داد که SNP‌های بالاتر از خط آستانه (نقاط زرد و



شکل ۲- مطالعه گسترده ارتباط ژنومی (GWAS) برای زمان گلدهی اکوتیپ‌های بابونه با استفاده از روش‌های MLM و FarmCPU  
Figure 2. Extensive study of genomic association (GWAS) for flowering time of chamomile ecotypes using MLM and FarmCPU methods



شکل ۳- نمودار QQ پلات با دو روش (A) MLM و (B) FarmCPU در GWAS زمان گلدهی اکوتیپ‌های بابونه  
Figure 3. QQ plot plot with two methods MLM (A) and FarmCPU (B) in GWAS flowering time of chamomile ecotypes

جدول ۲- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت زمان گلدهی اکوتیپ‌های بابونه

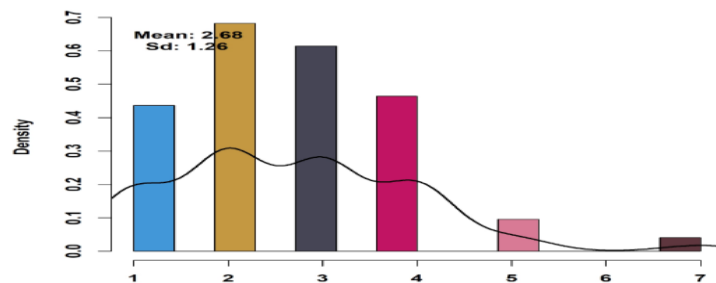
Table 2. SNP markers identified by FarmCPU approach for flowering time trait of chamomile ecotypes

SNP	REF	ALT	Effect	SE	-log pvalue
locus_89951-69	T	A	۱/۱۹	-۰/۲۹	۵/۲۶
locus_115004-15	G	A	-۰/۴۶	-۰/۱۱	۵/۴۴
locus_130938-23	G	C	-۰/۴۰	-۰/۰۹	۵/۵۸
locus_131494-101	C	T	-۰/۷۹	-۰/۱۷	۶/۴۰
locus_161754-71	T	A	-۱/۲۶	-۰/۱۹	۱۰/۳۵

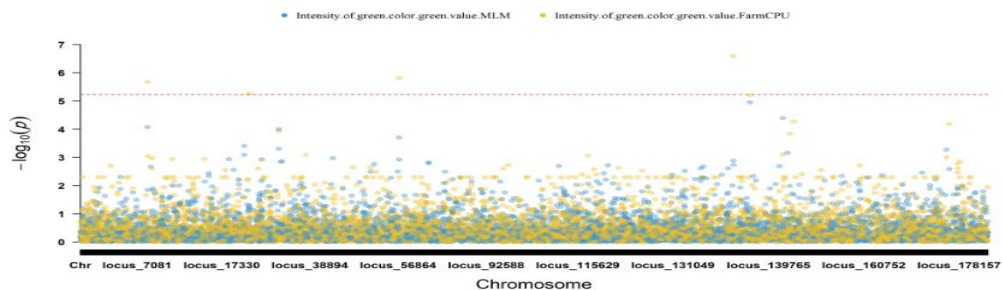
جدول ۳- نشانگرهای SNP شناسایی شده با رویکرد MLM برای صفت زمان گلدهی اکوتیپ‌های بابونه

Table 3. SNP markers identified by MLM approach for flowering time of chamomile ecotypes

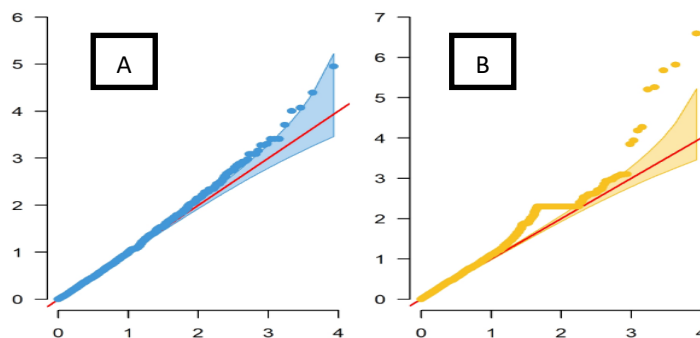
SNP	REF	ALT	Effect	SE	-log pvalue
locus_161754-71	T	A	-۱/۳۸	-۰/۲۹	۶/۳۶



شکل ۴- توزیع شدت رنگ سبز کرت مربوط به آنالیز GWAS اکوتیپ‌های بابونه  
Figure 4. Cretan green intensity intensity distribution related to GWAS analysis of chamomile ecotypes



شکل ۵- مطالعه گسترده ارتباط ژنومی (GWAS) برای شدت رنگ سبز کرت اکوتیپ‌های بابونه با استفاده از روش MLM و FarmCPU  
Figure 5. Extensive study of genomic association (GWAS) for green color intensity of chamomile ecotype plots using MLM and FarmCPU methods



شکل ۶- نمودار QQ پلات با دو روش MLM (A) و FarmCPU (B) در GWAS شدت رنگ سبز کرت اکوتیپ‌های بایونه  
Figure 6. QQ plot diagram with two methods MLM (A) and FarmCPU (B) in GWAS Intensity of the green color of the chamomile ecotypes

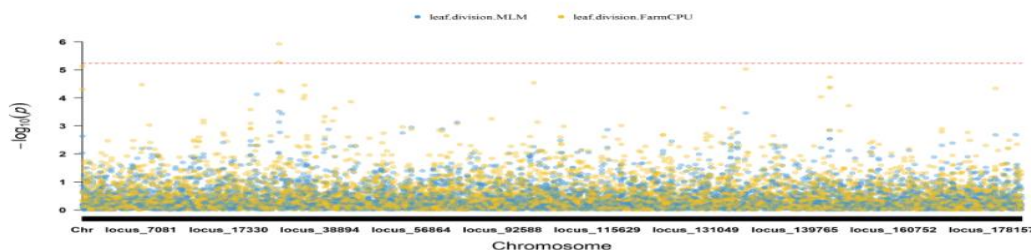
جدول ۴- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت شدت رنگ سبز کرت اکوتیپ‌های بایونه  
Table 4. SNP markers identified by the FarmCPU approach for the green color intensity trait of chamomile ecotypes

SNP	REF	ALT	Effect	SE	-log pvalue
locus_7885-44	A	G	-۰/۴۹	-۰/۱۲	۵/۶۷
locus_20832-20	C	T	-۰/۳۵	-۰/۰۹	۵/۲۶
locus_57030-41	G	A	-۰/۵۰	-۰/۱۱	۵/۸۲
locus_136992-38	G	C	-۰/۷۶	-۰/۱۷	۶/۵۹

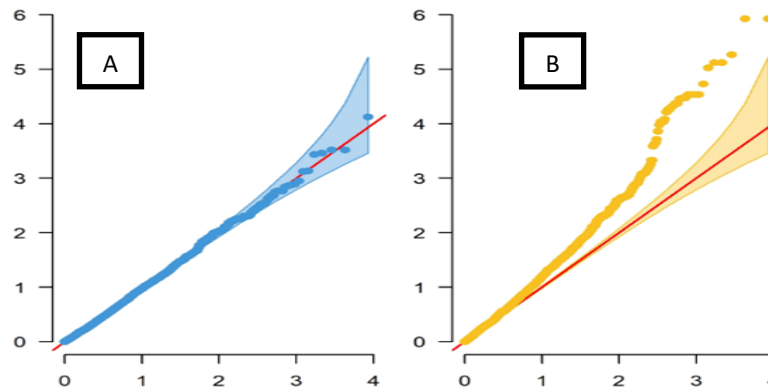
تاکنون در خصوص شناسایی نشانگرهای SNP مرتبط با شدت رنگ سبز کرت‌ها با بهره‌گیری از GWAS گزارشی منتشر نشده است و مطالعه ما اولین گزارش درباره تعیین نشانگرهای SNP مرتبط با شدت رنگ سبز کرت‌ها توسط آنالیز ارتباطی ژنوم بود. در مطالعه‌ای، تان و همکاران (۲۰) به شناسایی QTL های مرتبط با رنگ بوته‌های چای بر اساس نشانگرهای SSR پرداختند. آنها نشان دادند که دو QTL به نام‌های qYSC8 و qYSC3 با رنگ بوته‌های چای ارتباط معنی‌داری دارند و از آنها می‌توان به منظور انتخاب رنگ ژنوتیپ‌های چایی استفاده کرد. ضمناً نویسندگان پیشنهاد کردند که این QTL ها احتمالاً با فاکتورهای رونویسی در ارتباط هستند که بایستی مطالعات بیشتری در راستای آشکارسازی نقش آنها انجام شود. در کل، از نشانگرهای SNP شناسایی شده در این مطالعه می‌توان به منظور گزینش ژنوتیپ‌ها با شدت سبزی متفاوت برگ به کمک نشانگر استفاده کرد.

وضعیت تقسیمات و انشعابات برگ (صافی و زبری)  
آنالیز منهتن با روش FarmCPU منجر به شناسایی ۳ نشانگر معنی‌دار شد که با صفت وضعیت سطحی برگ (صافی و زبری) کرت‌ها ارتباط معنی‌داری دارند (شکل ۷). در واقع، تعداد ۳ نقطه (SNP) از حد آستانه بالاتر بودند که از بین آنها، یک نقطه (SNP) وجود داشت که منفی لگاریتم P-value بالاتری نسبت به بقیه داشت. مشخصات این نشانگرهای سیگنال شده برای صفت وضعیت سطحی برگ (صافی و زبری) کرت‌ها در جدول ۵ آورده شده است. همانطور که در جدول مشاهده می‌شود، SNP با مکان ژنی locus\_31850-95 با داشتن مقدار  $\log P$ -

value = 5.92567- می‌تواند به عنوان مرتب‌ترین نشانگر برای صفت وضعیت سطحی برگ (صافی و زبری) کرت‌ها معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگرهای شناسایی شده با آنالیز منهتن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP های بالاتر از خط آستانه (نقاط زرد و آبی رنگ به ترتیب گویای نشانگرهای شناسایی شده با روش‌های MLM و FarmCPU) دارای P-value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده شده و مقادیر موردانتظار برای SNP های واقع روی خط برازش وجود نداشت (شکل ۸).  
تاکنون، مطالعات اندکی در خصوص شناسایی مکان‌های ژنی کنترل‌کننده وضعیت سطحی برگ (صافی و زبری) و نشانگرهای مرتبط با آنها در گیاهان انجام شده است و لذا مطالعه ما برای اولین بار منجر به شناسایی نشانگرهای SNP مرتبط با صافی و زبری در گیاه دارویی بایونه آلمانی شد. در یک مطالعه، فراری و همکاران (۵) با استفاده از نشانگرهای RFLP موفق به شناسایی یک QTL مرتبط با وضعیت سطحی برگ سب‌زمینی شدند. بعد از این مطالعه، گزارشی جدیدی در خصوص شناسایی QTL ها و نشانگرهای مرتبط با وضعیت ظاهری برگ از نظر صافی و زبری منتشر نشد. در کل، از نشانگرهای SNP شناسایی شده در این مطالعه می‌توان به منظور گزینش ژنوتیپ‌ها با سطح صافی و زبری مورد نظر برگ به کمک نشانگر استفاده کرد.



شکل ۷- مطالعه گسترده ارتباط ژنومی (GWAS) برای وضعیت سطحی برگ (صافی و زبری) اکوتیپ‌های بایونه با استفاده از روش MLM و FarmCPU  
Figure 7. Extensive study of genomic association (GWAS) for leaf surface condition (smoothness and roughness) of chamomile ecotypes using MLM and FarmCPU methods



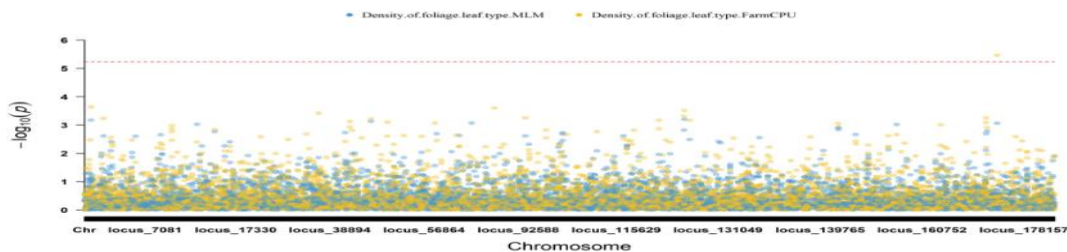
شکل ۸- نمودار QQ پلات با دو روش (A) MLM و (B) FarmCPU در GWAS وضعیت سطحی برگ (صافی و زبری) اکوتیپ‌های بابونه  
Figure 8. QQ plot plot with two methods MLM (A) and FarmCPU (B) in GWAS Leaf surface condition (smoothness and roughness) of chamomile ecotypes

جدول ۵- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت وضعیت سطحی برگ (صافی و زبری) اکوتیپ‌های بابونه  
Table 5. SNP markers identified by FarmCPU approach for leaf surface branch (fine and rough) of chamomile ecotypes

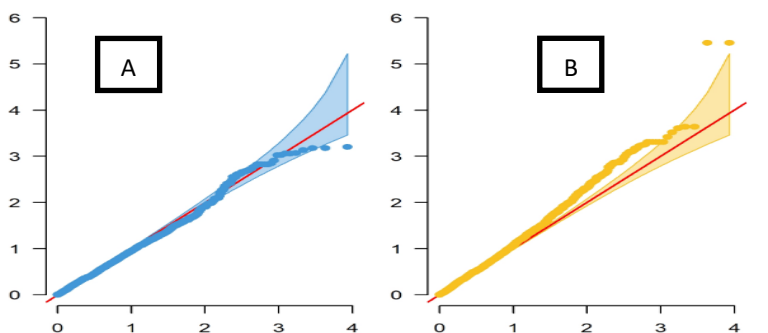
SNP	REF	ALT	Effect	SE	-log pvalue
locus_31850-21	G	A	-۰/۹۰	-/۱۹	۵/۲۶
locus_31850-94	A	G	-۰/۹۴	-/۱۸	۵/۹۲
locus_31850-95	T	G	-۰/۹۴	-/۱۸	۵/۹۲

p-value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده شده و مقادیر مورد انتظار برای SNP‌های واقع روی خط برازش وجود نداشت (شکل ۱۰). تاکنون، پژوهشی در حیطه شناسایی مکان‌های ژنی کنترل کننده تراکم بخش علفی و نشانگرهای مرتبط با آنها در گیاهان انجام نشده است و بنابراین مطالعه ما برای اولین بار منجر به شناسایی نشانگر SNP مرتبط با تراکم بخش علفی در گیاه دارویی بابونه آلمانی شد. در یک مطالعه، Guo و همکاران (۲۰۱۱) با استفاده از نشانگرهای SSR موفق به شناسایی پنج QTL مرتبط با تراکم بخش هوایی ذرت شدند. بعد از این پژوهش، گزارشی جدیدی در حیطه تعیین QTL ها و نشانگرهای مرتبط با تراکم بخش هوایی و علفی گیاهان منتشر نشد. در کل، از نشانگر SNP شناسایی شده در مطالعه ما می‌توان به منظور گزینش ژنوتیپ‌ها با تراکم موردنظر بخش علفی استفاده کرد.

**تراکم بخش علفی گیاه**  
آنالیز منتهن با روش FarmCPU منجر به شناسایی ۱ نشانگر معنی‌دار شد که با صفت تراکم بخش علفی کرت‌ها ارتباط معنی‌داری دارد (شکل ۹). در واقع، تعداد ۱ نقطه (SNP) از حد آستانه بالاتر بود. مشخصات این نشانگر سیگنال شده برای صفت تراکم بخش علفی کرت‌ها در جدول ۶ آورده شده است. همانطور که در جدول مشاهده می‌شود، SNP با مکان ژنی locus\_136992-38 با داشتن مقدار  $\log P$ -value = 6.5950- می‌تواند به عنوان مرتبط‌ترین نشانگر برای صفت تراکم بخش علفی کرت‌ها معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگرهای شناسایی شده با آنالیز منتهن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP‌های بالاتر از خط آستانه (نقاط زرد و آبی رنگ به ترتیب گویای نشانگرهای شناسایی شده با روش‌های MLM و FarmCPU) دارای



شکل ۹- مطالعه گسترده ارتباط ژنومی (GWAS) برای تراکم بخش علفی اکوتیپ‌های بابونه با استفاده از روش MLM و FarmCPU  
Figure 9. Extensive study of genomic association (GWAS) for herbaceous densification of chamomile ecotypes using MLM and FarmCPU methods



شکل ۱۰- نمودار QQ پلات با دو روش (A) MLM و (B) FarmCPU در GWAS تراکم بخش علفی اکوتیپ‌های بابونه  
Figure 10. QQ plot plot with MLM (A) and FarmCPU (B) methods in GWAS Density of herbaceous part of chamomile ecotypes

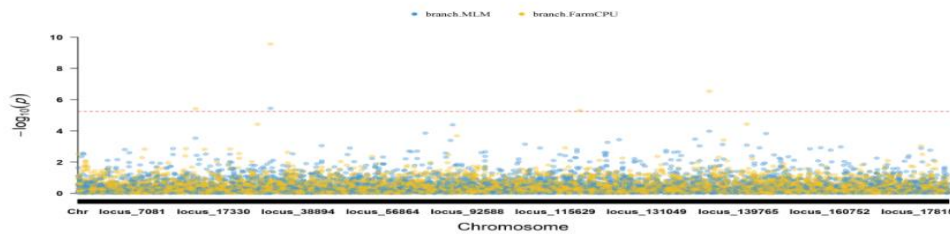
جدول ۶- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت تراکم بخش علفی اکوتیپ‌های بابونه

Table 6. SNP markers identified by the FarmCPU approach for foliage density of chamomile ecotypes

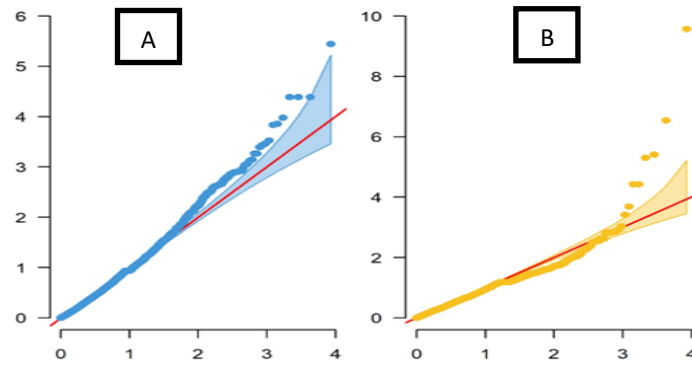
SNP	REF	ALT	Effect	SE	-log pvalue
locus_136992-38	T	G	۰/۶۰	۰/۱۱	۶/۵۹

بین مقادیر مشاهده شده و مقادیر موردانتظار برای SNP‌های واقع روی خط برازش وجود نداشت (شکل ۱۲). تاکنون، مطالعات اندکی در خصوص شناسایی مکان‌های ژنی کنترل کننده وضعیت شاخه‌های فرعی و نشانگرهای مرتبط با آنها در گیاهان انجام شده است و لذا مطالعه ما برای اولین بار منجر به تعیین نشانگرهای SNP مرتبط با وضعیت شاخه‌های فرعی در گیاه دارویی بابونه آلمانی شد. در یک مطالعه، هووانگ و همکاران (۱۰) با استفاده از نشانگرهای SSR موفق به شناسایی دو QTL به نام RSB7 و RSB6 شدند که مرتبط با وضعیت وضعیت شاخه‌های فرعی گیاه آراییدوپسیس بودند. نویسندگان نشان دادند که QTL های RSB7 و RSB6 به ترتیب با ژن های FLC و FRI همولوگ هستند و به طریق ایپساتیک، وضعیت انشعابی شاخه‌ها را تعیین می‌کنند. بعد از این تحقیق، گزارشی جدید درباره شناسایی QTL ها و نشانگرهای مرتبط با وضعیت شاخه‌های فرعی گیاهان منتشر نشد. در کل، از نشانگرهای SNP شناسایی شده در این مطالعه می‌توان به منظور گزینش ژنوتیپ‌ها با وضعیت موردنظر شاخه‌ها (انشعاب شاخه‌ها از بالا یا پایین) استفاده کرد.

**وضعیت شاخه‌های فرعی**  
 آنالیز منهن با روش FarmCPU منجر به شناسایی ۳ نشانگر معنی‌دار شد که با صفت وضعیت شاخه‌های فرعی کرت‌ها ارتباط معنی‌داری دارند (شکل ۱۱). در واقع، تعداد ۳ نقطه (SNP) از حد آستانه بالاتر بودند که از بین آنها، یک نقطه (SNP) وجود داشت که منفی لگاریتم P-value بالاتری نسبت به بقیه داشت. مشخصات این نشانگرهای سیگنال شده برای صفت وضعیت شاخه‌های فرعی کرت‌ها در جدول ۷ آورده شده است. همانطور که در جدول مشاهده می‌شود، SNP مکان ژنی locus\_35488-32 با داشتن مقدار  $-\log P\text{-value} = 8.0247$  می‌تواند به عنوان مرتبط‌ترین نشانگر برای صفت وضعیت شاخه‌های فرعی کرت‌ها معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگرهای شناسایی شده با آنالیز منهن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP های بالاتر از خط آستانه (نقاط زرد و آبی رنگ) به ترتیب گویای نشانگرهای شناسایی شده با روش های MLM و FarmCPU، دارای P-value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر رد می‌شود. در مقابل، تفاوت معنی‌داری



شکل ۱۱- مطالعه گسترده ارتباط ژنومی (GWAS) برای وضعیت شاخه‌های فرعی اکوتیپ‌های بابونه با استفاده از روش MLM و FarmCPU  
 Figure 11. Extensive Study of Genomic Relationship (GWAS) for the Status of Chamomile Ecotype Subsidia Using MLM and FarmCPU Methods



شکل ۱۲- نمودار QQ پلات با دو روش MLM (A) و FarmCPU (B) در GWAS وضعیت شاخه‌های فرعی اکوتیپ‌های بابونه  
 Figure 12. QQ plot plot with MLM (A) and FarmCPU (B) methods in GWAS Status of chamomile ecotype sub-branches

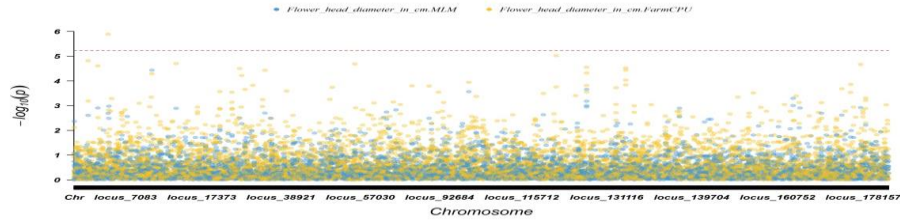
جدول ۷- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت وضعیت شاخه‌های فرعی اکوتیپ‌های بابونه

Table 7. SNP markers identified by the FarmCPU approach for the status of chamomile ecotype subbranches

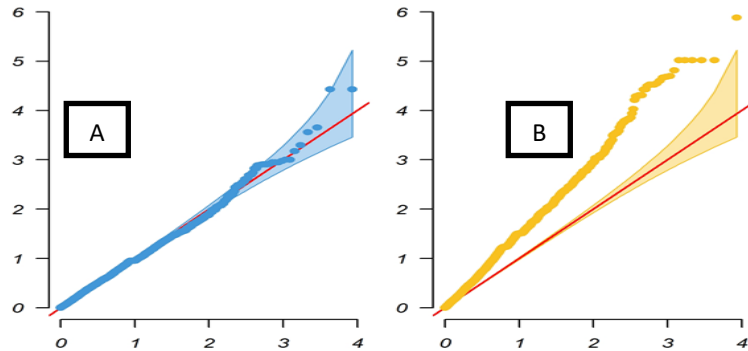
SNP	REF	ALT	Effect	SE	-log pvalue
locus_12071-29	T	G	۰/۶۰	۰/۱۱	۷/۸۰
locus_35488-32	C	T	۰/۴۲	۰/۰۸	۸/۰۲
locus_117292-63	C	T	۰/۶۳	۰/۱۳	۶/۵۸

هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده شده و مقادیر موردانتظار برای SNP‌های واقع روی خط برازش وجود نداشت (شکل ۱۴). مطالعه ما برای اولین بار منجر به تعیین نشانگر SNP مرتبط با قطر گل در گیاه دارویی بابونه آلمانی شد. در یک پژوهش، کاوا و همکاران (۲) با استفاده از نشانگرهای SNP موفق به شناسایی دو QTL به نام‌های qFlower\_D3.1 و qFlower\_D2.1 شدند که مرتبط با تعداد گل در نتاج حاصل از تلاقی گیاهان دارویی *Petunia integrifolia* × *P. axillaris* بودند. در کل، از نشانگر SNP شناسایی شده در این مطالعه می‌توان به منظور گزینش ژنوتیپ‌ها بر اساس قطر گل استفاده کرد.

آنالیز منهن با روش FarmCPU منجر به شناسایی ۱ نشانگر معنی‌دار شد که با صفت قطر گل ارتباط معنی‌داری دارد (شکل ۱۳). در واقع، تعداد ۱ نقطه (SNP) از حد آستانه بالاتر بود. مشخصات این نشانگر سیگنال شده برای صفت قطر گل در جدول ۸ آورده شده است. همانطور که در جدول مشاهده می‌شود، نشانگر SNP مکان ژنی locus\_2494-70 با داشتن مقدار  $-\log P\text{-value} = 5.8867$  می‌تواند به عنوان نشانگر مرتبط با صفت قطر گل معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگر شناسایی شده با آنالیز منهن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP های بالاتر از خط آستانه (نقاط زرد و آبی رنگ) به ترتیب گویای نشانگرهای شناسایی شده با روش های MLM و FarmCPU، دارای P-value بالا و معنی‌دار



شکل ۱۳- مطالعه گسترده ارتباط ژنومی (GWAS) برای قطر گل اکوتیپ‌های بابونه با استفاده از روش MLM و FarmCPU  
Figure 13. Extensive study of genomic association (GWAS) for flower diameters of chamomile ecotypes using MLM and FarmCPU methods



شکل ۱۴- نمودار QQ پلات با دو روش (A) MLM و (B) FarmCPU در GWAS قطر گل اکوتیپ‌های بابونه  
Figure 14. QQ plot plot with two methods MLM (A) and FarmCPU (B) in GWAS diameter of chamomile ecotypes

جدول ۸- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت قطر گل اکوتیپ‌های بابونه

Table 8. SNP markers identified by FarmCPU approach for flower diameter trait of chamomile ecotypes

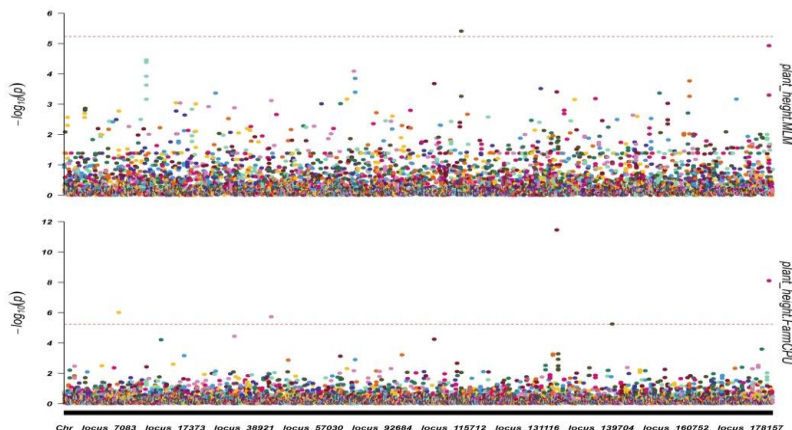
SNP	REF	ALT	Effect	SE	-log pvalue
locus_2494-70	C	T	۰/۵۴	۰/۱۰	۵/۸۸

SNP رد می‌شود. در مقابل، تفاوت معنی‌داری بین مقادیر مشاهده شده و مقادیر موردانتظار برای SNP‌های واقع روی خط برازش وجود نداشت (شکل ۱۶).

مطالعه ما برای اولین بار منجر به تعیین نشانگرهای SNP مرتبط ارتفاع بوته در گیاه داروئی بابونه آلمانی شد. ارتفاع یکی از صفات مهم زراعی است که نه تنها بر کنترل آفات و بیماری‌ها بلکه بر عملیات زراعی برداشت نیز تاثیرگذار است و به همین دلیل نیز مطالعات متنوعی در راستای شناسایی نشانگرها و QTL‌های مرتبط با این صفت صورت گرفته است. برای مثال، دانگ و همکاران (۴) با استفاده از توالی‌یابی مجدد ژنوم موفق به شناسایی یک QTL به نام qPHA10 شدند که مرتبط با ارتفاع بوته در گیاه کلزا بود. نویسندگان پیشنهاد کردند که از این QTL می‌توان به منظور انتخاب بوته‌هایی استفاده کرد که از ارتفاع مناسبی در مزرعه برخوردار هستند. در کل، از نشانگرهای SNP شناسایی شده در این مطالعه می‌توان به منظور گزینش ژنوتیپ‌ها با ارتفاع مناسب بوته به‌منظور تسهیل عملیات زراعی استفاده کرد.

### ارتفاع بوته

آنالیز منهن با دو روش MLM و FarmCPU به ترتیب منجر به شناسایی ۱ و ۵ نشانگر معنی‌دار شد که با صفت ارتفاع بوته ارتباط معنی‌داری دارند (شکل ۱۵). در واقع، تعداد چند نقطه (SNP) از حد آستانه بالاتر بودند که از بین آنها، یک نقطه (SNP) وجود داشت که منفی لگاریتم P-value بالاتری نسبت به بقیه داشت. مشخصات این نشانگرهای سیگنال شده برای صفت ارتفاع بوته در جدول ۹ و ۱۰ آورده شده است. همانطور که در جدول مشاهده می‌شود، SNP با مکان ژنی locus\_135707-30 با داشتن مقدار  $\log P\text{-value} = 11.4573$  می‌تواند به عنوان مرتبط‌ترین نشانگر برای صفت ارتفاع بوته معرفی شود. نمودار QQ پلات با دو روش MLM و FarmCPU نیز منجر به تایید نشانگرهای شناسایی شده با آنالیز منهن شد. به عبارتی، QQ پلات با دو روش MLM و FarmCPU نشان داد که SNP‌های بالاتر از خط آستانه (نقاط زرد و آبی رنگ به ترتیب گویای نشانگرهای شناسایی شده با روش‌های MLM و FarmCPU، دارای P-value بالا و معنی‌دار هستند. در چنین حالتی، فرض صفر مبنی بر عدم ارتباط بین صفت و نشانگر

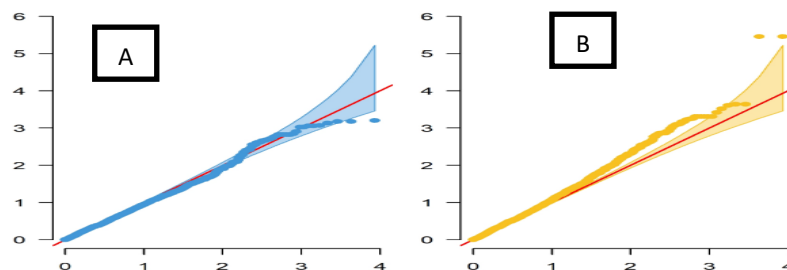


شکل ۱۵- مطالعه گسترده ارتباط ژنومی (GWAS) برای ارتفاع بوته اکوتیپ‌های بابونه با استفاده از روش MLM و FarmCPU  
Figure 15. Extensive study of genomic association (GWAS) for plant height of chamomile ecotypes using MLM and FarmCPU methods

جدول ۹- نشانگرهای SNP شناسایی شده با رویکرد FarmCPU برای صفت ارتفاع بوته اکوتیپ‌های بابونه

Table 9. SNP markers identified by FarmCPU approach for plant height trait of chamomile ecotypes

SNP	REF	ALT	Effect	SE	-log pvalue
locus_8296-25	A	G	۱/۴۲	۰/۳۲	۶/۰۱
locus_49847-19	G	A	-۴/۳۵	۱/۰۰	۵/۷۳
locus_135707-30	A	T	۴/۵۰	۰/۶۵	۱۱/۴۵
locus_141117-22	C	T	-۲/۷۸	۰/۶۷	۵/۲۵
locus_182069-110	A	T	-۲/۶۲	۰/۵۲	۸/۱۰



شکل ۱۶- نمودار QQ پلات با دو روش MLM (A) و FarmCPU (B) در GWAS ارتفاع بوته اکوتیپ‌های بابونه  
Figure 16. QQ plot plot with two methods MLM (A) and FarmCPU (B) in GWAS Plant height of chamomile ecotypes

جدول ۱۰- نشانگرهای SNP شناسایی شده با رویکرد MLM برای صفت ارتفاع بوته اکوتیپ‌های بابونه

Table 10. SNP markers identified by MLM approach for plant height trait of chamomile ecotypes

SNP	REF	ALT	Effect	SE	-log pvalue
locus_116415-5	T	C	۳/۰۱	۰/۶۳	۵/۴۱

گل کرت، تعداد روز تا گلدهی، وزن تر گیاه، وزن خشک گیاه، میزان نسبی آنتوسیانین ساقه، میزان اسانس، ماتریسین، آلفایزابلول، بیزابولول اکساید A و نسبت بیزابولون اکساید به آلفایزابلول در گیاه دارویی بابونه آلمانی شد. در مقابل، نشانگرهای SNP مرتبط با وضعیت ظاهری تک بوته، تعداد گل و بیزابولول اکساید B در این مطالعه یافت نشدند. درکل، از نشانگرهای SNP شناسایی و سیگنال شده در این مطالعه می‌توان به منظور گزینش به کمک نشانگر در برنامه‌های اصلاحی آتی بابونه آلمانی استفاده کرد. با توجه به اینکه جهت انتخاب صفات مرتبط با اسانس زمان و هزینه زیادی لازم است، لذا انتخاب بر اساس SNP‌های سیگنال شده، می‌تواند در هزینه و زمان صرفه جویی کند.

### نتیجه‌گیری کلی

در مطالعه GWAS، به منظور شناسایی SNP های مرتبط با صفات مورفولوژی و بیوشیمیایی مختلف از آنالیز منهتن و نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU استفاده شد. نتایج آنالیز منهتن و اعتبارسنجی آنها با نمودار QQ پلات مبتنی بر مدل‌های MLM و FarmCPU منجر به شناسایی SNP های معنی دار و مرتبط با صفاتی همچون زمان گلدهی، شدت رنگ سبز کرت، وضعیت سطحی برگ (صافی و زبری)، تراکم بخش علفی گیاه، وضعیت انشعابی شاخه‌های فرعی، وضعیت شاخه‌های پایین بوته، هموزنی یا یکسانی گیاهان کرت، ارتفاع بوته، وجود و یا عدم وجود گل بزرگ، قطر گل، میزان

### منابع

- Bastien, M., H. Sanah and F. Belzile. 2014. Genome wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean with a genotyping-by-sequencing approach. *The Plant Gen*, 7: 1-13.
- Cao, Z., Y. Guo and Q. Yang. 2019. Genome-wide identification of quantitative trait loci for important plant and flower traits in petunia using a high-density linkage map and an interspecific recombinant inbred population derived from *Petunia integrifolia* and *P. axillaris*. *Hortic Res*, 6: 27.
- Das, M. 2014. Chamomile: medicinal, biochemical, and agricultural aspects. Boca Raton: CRC Press Taylor and Francis Group.
- Dong, Z., M. Khorshed Alam, M. Xie, L. Yang, J. Liu, M. Helal, J. Huang, X. Cheng, Y. Liu, C. Tong, and C. Zhao. 2021. Mapping of a major QTL controlling plant height using a high-density genetic map and QTL-seq methods based on whole-genome resequencing in *Brassica napus*. *G3 Genes/Genomes/Genetics*, 11: 33-45.
- Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler and S.E. Mitchell. 2016. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5): e19379.
- Frery, A., S. Doganlar and M.C. Daunay. 2003. QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of solanaceous species. *Theor Appl Genet*, 107: 359-370.
- Frayling, T.M. 2016. Genome-wide association studies: the good, the bad and the ugly. *Clin Med (Lond)*, 14(4): 428-431.
- Ghanavati, M., S. Houshmand, H. Zainali and F. Abrahimpour. 2010. Chemical Composition of the Essential Oils of *Matricaria recutita* L. Belonging to Central and South Parts of Iran. *J. Med. Plants*, 9 (34): 102-108
- He, J., X. Zhao, A. Laroche, Lu. Z-X, H. Liu and Z. Li. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front Plant Sci*, 5:484.
- Huang, Y.F., J.A. Poland, C.P. Wight, E.W. Jackson and N.A. Tinker. 2014. Using genotyping by- sequencing (GBS) for genomic discovery in cultivated oat. *PLoS One*, 9(7): e102448.
- Kameswara, N. 2014. Plant genetic resources: Advancing conservation and use through biotechnology. *Afr. J. biotechnol.* 3(2): 136-145.
- Kilian, B. and A. Graner. 2012. NGS technologies for analyzing germplasm diversity in genebanks. *Brief Funct Genomics*, 11(1): 38-50.
- Kurasawa, K., A. Matsui, R. Yokoyama, T. Kuriyama, T. Yoshizumi, M. Matsui, K. Suwabe, M. Watanabe and K. Nishitani. 2009. The AtXTH28 gene, a xyloglucan endotransglucosylase /hydrolase, is involved in automatic self-pollination in *Arabidopsis thaliana*. *Plant Cell Physiol*, 50(2): 413-22.
- Lu, F., AE. Lipka, J. Glaubitz, R. Elshire, J.H. Cherney, MD. Casler, ES. Buckler and D.E. Costich. 2017. Switchgrass genomic diversity, Ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet*, 9(1): e1003215.
- Mehrabi, A. and S. Miri. 2019. Identification of informative SNP markers associated to root traits and indices, water use efficiency and transpiration efficiency in durum wheat using GWAS. *MGJ*, 14(1) : 39-47
- Otto, L.G., P. Mondal, J. Brassac, S. Preiss, J. Degenhardt, S. He and T.F. Sharbel. 2017. Use of genotyping-by-sequencing to determine the genetic structure in the medicinal plant chamomile, and to identify flowering time and alpha-bisabolol associated SNP-loci by genome-wide association mapping. *BMC genomics*, 18(1): 599.
- Rathore, S. and R. Kumar. 2021. Agronomic interventions affect the growth, yield, and essential oil composition of German chamomile (*Matricaria chamomilla* L.) in the western Himalaya. *Ind Crops Prod*, 171: 4-9.
- Ruzicka, J., M. Hacek and J. Novak. 2021. Genetic variation and mitochondrial relationships between various chamomile accessions. *J Appl Genet*, 62(1): 73-84.
- Ruzicka, J. and J. Novak. 2020. Mitochondrial genome variation between different accessions of *Matricaria chamomilla* L. (Asteraceae) based on SNP mutation analysis. *Genet Resour Crop Evol*. 67: 853-864.
- Tan, L.Q., L.Y. Wang and L.Y. Xu. 2016. SSR-based genetic mapping and QTL analysis for timing of spring bud flush, young shoot color, and mature leaf size in tea plant (*Camellia sinensis*). *Tree Genet. Genomes*, 12:52.
- Uitdewilligen, J., A.M.A Wolters, B.B. D'hoop, T. Borm, R. Visser and H.J. van Eck. 2013. A next-generation sequencing method for genotyping-bysequencing of highly heterozygous Autotetraploid potato. *PLoS One*, 8(5): e62355.

## Genome Wide Assosiation Study (GWAS) in some Chamomile Genotypes (*Matricaria chamomilla* L.) using Morphological, Phenological Traits

Mehdi Ghanavati<sup>1</sup>, Sadollah Houshmand<sup>2</sup>, Sebastian Albrecht<sup>3</sup> and Lars-Gernot Otto<sup>4</sup>

1- Department of Agriculture, Payame Noor University, Tehran, Iran, (Corresponding author: m\_ghanavati@pnu.ac.ir)

2- Shahrekord University, Professor, Biotechnology and plant breeding, Faculty of Agriculture, Chaharmahal and baktiyari province, Iran

3- Pharmaplant - Medicinal and Aromatic Plant Research and Seed Production GmbH, Artern, Germany

4- IPK Leibniz Institute of Plant Genetics and Crop Plant Research· Department of Breeding Research, Gatersleben, Germany

Received: 7 March, 2022

Accepted: 17 May, 2022

### Extended Abstract

**Introduction and Objective:** German chamomile is one of the most widely used medicinal plants in the world, and large amounts of its essential oil are used every year in the pharmaceutical, cosmetic, health, and food industries.

Medicinal plants have many applications in the treatment of humans, among which chamomile is one of the most valuable and widely used medicinal plants in the world. One of the most important breeding programs is the recognition of genetic diversity for the initial evaluation of plant masses. For this purpose, a study of genetic diversity and classification of experimental chamomile ecotypes was designed and implemented.

**Material and methods:** 46 native and foreign ecotypes of chamomile were evaluated in a randomized complete block design with four replications in the research farm of German Pharmaplant. Some of these ecotypes were collected from their natural growing areas and some were chamomile cultivars. For this purpose, some morphological, phenological and were evaluated. GWAS technique was used to investigate the relationship between markers and studied traits. In GWAS, Manhattan analysis and QQ plot diagram based on MLM and FarmCPU models were used to identify SNPs related to different morphological and biochemical traits.

**Results:** The results of Manhattan analysis and their validation with QQ plot diagram based on MLM and FarmCPU models led to the identification of SNPs related to the following traits in chamomile: Flowering time, intensity of plot green color, leaf surface state (smoothness and roughness), density of foliage part, branching style of sub-branches, type of lower branches of the plant, homogeneity of plot plants, plant height, presence or absence of large flower, flower diameter, plot flower number, number of days to flowering, plant fresh weight, plant dry weight, relative amount of stem anthocyanin, oil content, matricin, alpha-bisabolol, bisabolol oxide A, and bisabolone oxide to alpha-bisabolol ratio. The following markers were identified in this study: SNP with locus\_161754-71 and log -  $P_{value}=10.35354$  as the most relevant marker for flowering time trait, SNP with locus\_136992-38 and log -  $P_{value}=6.5950$  as the most relevant marker for green color of plot, SNP with locus\_31850-95 and log -  $P_{value}=5.92567$  as the most relevant marker for leaf surface (fine and rough) of plot, SNP with locus\_136992-38 and log -  $P_{value}=6.5950$  as the most relevant marker for density of the foliage part of plot, SNP with locus\_35488-32 and log -  $P_{value}=8.0247$  as the most relevant marker for status of sub-branches of plot, SNP with locus\_139800-67 and log -  $P_{value}=8.438303$  as the most relevant marker for status of lower branches of the plant, SNP with locus\_135707-30 and log -  $P_{value}=11.4573$  as the most relevant marker for height of plant, , SNP with locus\_2494-70 and log -  $P_{value}=5.8867$  as the most relevant marker for the flower diameter, Overall, we discovered the genetic structure in the chamomile ecotypes by using GBS and also identified SNPs associated with some morpho-physiological. Therefore, the SNP markers identified in this GWAS study can be used for marker-assisted selection in chamomile breeding programs.

**Conclusion:** The SNP markers identified in this GWAS study, especially those with a higher economic aspect, can be used for marker-assisted selection in German chamomile breeding programs.

**Keywords:** Chamomile, GBS, GWAS, Sequencing, SNPs