

ارزیابی اثرات اصلی و متقابل ژنوتیپ × محیط و برآورد ضرایب همبستگی ارزش‌های اصلاحی ژنومی صفات کیفیت نانوائی گندم (*T.aestivum* L.)

Evaluation of the main and genotype × environment interaction effects of and estimation of correlation coefficients of genomic breeding values of bakery quality traits of bread wheat

محمد آرمیون^۱، محمد رضا بی‌همتا^{۲*}، رضا معالی امیری^۲، منوچهر خدا رحمی^۳ و ساچیکو ایسوبه^۴

۱- استادیار، بخش تحقیقات علوم زراعی و باغی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی ایلام، سازمان

تحقیقات، آموزش و ترویج کشاورزی، ایلام، ایران

۲- استاد و دانشیار، گروه زراعت و اصلاح نباتات، دانشکده علوم و مهندسی کشاورزی دانشگاه تهران، کرج، ایران

۳- استادیار، موسسه تحقیقات اصلاح و تهیه نهال و بذر، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

۴- رئیس آزمایشگاه ژنومیکس و ژنتیک کاربردی گیاهی، موسسه تحقیقات DNA کازوسا ژاپن، کازوسا، ژاپن

Armion M¹, Bihamta MR^{*2}, Maali Amiri R², Khodarahmi M³, Isobe S⁴

1 -Assistant Professor of Crop and Horticultural Science Research Department, Ilam

Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization, AREEO, Ilam, Iran

2 -Professor & Associate Professor in plant breeding, Department of Agronomy and

Plant Breeding, the Faculty of Agricultural Science & Engineering, University of Tehran, Karaj, Iran

3 -Assistant Professor of Cereal research Department, Seed and Plant Improvement Institute, Agricultural Research, Education and Extension Organization, AREEO, Karaj, Iran

4 -Senior Scientist, Department of Applied Plant Genomics and Genetics, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

* نویسنده مسئول مکاتبات، پست الکترونیکی: mrghanad@ut.ac.ir

(تاریخ دریافت: ۱۴۰۱/۳/۱۰ - تاریخ پذیرش: ۱۴۰۱/۴/۴)

چکیده

جهت ارزیابی اثرات اصلی و متقابل ژنوتیپ × محیط و برآورد و مقایسه ضرایب همبستگی ارزش‌های اصلاحی ژنومی صفات کیفیت نانوائی گندم در یک جمعیت پیشرفته اصلاحی گندم نان، آزمایشی در قالب طرح آلفا لاتیس در سه تکرار با ۵۰ لاین و ارقام شاهد در ایستگاه تحقیقاتی چرداول مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان ایلام به مدت دو سال اجرا شد. تجزیه واریانس صفات کیفیت نانوائی گندم به صورت ساده و مرکب انجام و عامل ژنوتیپ، سال و اثر متقابل ژنوتیپ × سال بر میزان پروتئین خام، عدد فالینگ، عدد زلنی و اندازه حجم رسوب SDS معنی‌دار بود. برای برآورد ارزش‌های اصلاحی ژنومی از روش ddRAD-Seq و بستر NextSeqTM500Illumina استفاده شد. تعداد کل SNP های صحیح فراخونی شده برای ۵۰ درصد داده گم‌شده برابر با ۳۳۴۲ عدد، که تعداد ۱۳۲۲ بر ژنوم B، ۱۲۵۳ بر ژنوم A و ۷۶۷ روی ژنوم D و بیشترین SNP بر روی کروموزوم دو و سه از ژنوم B و کمترین آن روی کروموزوم چهار از ژنوم D شناسایی شد. برآورد ضریب همبستگی ارزش اصلاحی ژنومی در سال اول، دوم و متوسط دو سال برای صفات پروتئین خام، عدد فالینگ و عدد زلنی با روش‌های آماری رگرسیون LASSO و رگرسیون شبکه الاستیک صورت گرفت و برای اندازه حجم رسوب SDS هیچگونه مدلی برازش نشد.

واژه‌های کلیدی

اثرات اصلی و متقابل

روش آماری

ژنوتیپ‌سنجی به روش ddRAD-Seq

ضرایب همبستگی برآورد ارزش اصلاحی ژنومی

گندم نان

مقدمه

گندم یک محصول زراعی مهم، اصلی و عمده که در بیشتر مناطق جهان رشد می‌کند (Kristensen et al. 2018). در گندم نان (T. aestivum L. صفات کیفیت نانوائی از اهداف اولیه اصلاحی هستند (Haile 2018). علی‌رغم گسترش چند آزمون غیرمستقیم برای سنجش صفات کیفیت نانوائی در گندم همچنان این روش‌ها نیازمند صرف زمان و منابع فراوان برای سنجش رخ‌نمود^۱ صحیح بوده که مقادیر فراوان بذر در نسل‌های اولیه قابل دسترس نبوده لذا این صفات در نسل‌های پیشرفته ارزیابی می‌شوند (Battenfield et al. 2016). کاربرد مدل‌های پیش‌بین ارزش اصلاحی ژنومی در نسل‌های اولیه برنامه‌های به‌نژادی برای صفات فرآوری و کیفیت گندم باعث صرفه‌جویی در زمان و منابع می‌شود (Battenfield et al. 2016). در گندم برآورد ارزش اصلاحی ژنومی صفات عملکرد بذر، تعداد بذر در سنبله، تاریخ خوشه‌دهی، ارتفاع گیاه، خوابیدگی، جوانه‌زنی پیش از برداشت، وزن تست (حجمی)، وزن هزار بذر، عملکرد و نرمی آرد، پروتئین بذر، قدرت گلوتن، جذب آب، نشاسته‌های آسیب دیده، آرایینوزیلان، میزان گلیادین هیدراته جزئی و مقاومت به بیماری‌های قارچی (زنگ گندم، فوزرایوم خوشه و سپتوریوز) گزارش شده است (Haile 2018). از جمله روش‌های آماری که برای برآورد و مقایسه ضریب همبستگی ارزش‌های اصلاحی ژنومی استفاده شده شامل روش حداقل مربعات عادی (Elliot et al. 2009)، رگرسیون گام به گام (Elliot et al. 2009)، بهترین پیش‌بین ناریب خطی کلاسیک (Henderson 1984)، رگرسیون ریبج (Breiman 1995)، بهترین پیش‌بینی ناریب خطی ژنومی (Nejati-Javaremi et al. 1997; Van Raden 2008)، بهترین پیش‌بینی ناریب خطی وزنی (Hayashi & Iwata 2010; Gonzalez-Recio & Forni 2011)، رگرسیون شبکه الاستیک (LASSO و رگرسیون شبکه الاستیک (Zou & Hastie 2005)، رگرسیون بیز (Meuwissen et al. 2001)، رگرسیون کرنل یا هسته‌ای (Gianola & van Kaam 2008; Bennewitz et al. 2009)، روش‌های یادگیری ماشینی (Long et al. 2007; Wei et al. 2009).

Beavis 1994; al. 2009) و رگرسیون‌های غیرپارامتریک (Jannink et al. 2010) است. تفاوت این روش‌ها در مدل ژنتیکی و فرضیات واریانس اثر نشانگرها می‌باشد (Jannink et al. 2010). روش بهترین پیش‌بین ناریب خطی کلاسیک برآورد ارزش اصلاحی صفات کمی را از طریق بهترین پیش‌بین ناریب خطی که مبتنی بر ماتریس مشاهده‌های رخ‌نمود و روابط خویشاوندی بوده برآورد می‌کند. در روش رگرسیون ریبج به جای ماتریس خویشاوندی از ماتریس نشانگرها استفاده شده و تمام اثرات نشانگرها را در جهت صفر کاهش می‌دهد. در فرضیه‌های این روش نشانگرها دارای اثر تصادفی با واریانس مشترک هستند. منظور از واریانس مشترک این نیست که تمام نشانگرها دارای اثر مشابه هستند بلکه تمام اثرات نشانگر به صورت مساوی در جهت صفر کاهش یافته و این روش نشانگرهای با اثر کم را تسخیر می‌کند (Bernardo & Yu 2007; Moser et al. 2009). برای اولین بار برآورد ضریب همبستگی ژنومی گندم توسط Crossa et al. (2010) برآورد و شش مدل مختلف در یک جمعیت ۵۹۹ لاین گندم برای پیش‌بینی عملکرد بذر بررسی و صحت‌ها یا ضریب همبستگی ژنومی در دامنه ۰/۳۶ الی ۰/۶۱ بر اساس روش ده بار اعتبارسنجی متقابل گزارش شد. (Heffner et al. 2011a) روش‌های BayesA، RR-BLUP، BayesB و BayesCπ مقایسه کرده و طیف وسیعی از صحت‌های پیش‌بینی بین صفات و مدل‌های مختلف از ۰/۱۷ (عملکرد دانه در روش BayesCπ) تا ۰/۷۶ (برای عملکرد آرد در BayesA) گزارش کرده و تفاوت‌های جزئی در بین مدل‌های ارزیابی شده برای هر صفت گزارش و میانگین صحت روش BayesA در بین تمام صفات بالاترین بود. (Hayes et al. 2017) مدل یک صفت برای پیش‌بینی ۱۹ صفت کیفیت نانوائی در نمونه‌های گندم نان استفاده کرده و صحت‌های ۰ تا ۰/۶۹ با استفاده از طرح‌های اعتبارسنجی مختلف گزارش کردند. (Battenfield et al. 2016) پنج روش آماری برای مدل انتخاب ژنومی برای پیش‌بینی ده صفت فرآوری و کیفیت نهایی در لاین‌های پیشرفته گندم نان بهاره استفاده و صحت‌های ۰/۳۲ الی ۰/۶۲ گزارش کردند. با ظهور فناوری‌های نسل جدید از نشانگرهای SNP جهت سنجش ژن‌نمود^۲ در بیشتر گیاهان زراعی

^۱ Phenotyping^۲ Genotyping

حسب نقشه تصادفی که به وسیله نرم افزار آلفا ژن تهیه شده، انجام شد (ALPHANAL 1986).

در این تحقیق از یک جمعیت پیشرفته اصلاحی انتخابی (جدول ۱) از لاین‌های آزمایش مقدماتی اقلیم مناطق گرم و مرطوب شمال که از موسسه تحقیقات اصلاح و تهیه نهال و بذر کشور تهیه گردیده، استفاده شد.

برای انجام آزمون‌های کیفی مقدار ۵۰ گرم از هر لاین با آسیاب آزمایشگاهی بولر موسسه تحقیقات اصلاح و تهیه نهال و بذر آسیاب شدند (AACC 26-30).

مقدار پروتئین خام و عدد زلنی نمونه‌های آرد به وسیله دستگاه NIR^۳ (Inframatic 8620) پس از تنظیم به ازای هر لاین خوانش و گزارش شد.

آزمون رسوب SDS مطابق با روش Carter et al. (1999) انجام شد. در این روش از لوله شیشه‌ای با طول ۱۵ و قطر ۱/۵ سانتی‌متر با برچسب مدرج شده و دستگاه مخلوط کن ورتکس و شیکر زلنی استفاده شد. آزمون بدین صورت انجام شد: ۰/۶ گرم آرد در لوله ۱۵ سانتی‌متر حاوی ۴ میلی‌لیتر آب مقطر افزوده شده و به مدت ۲۰ ثانیه ورتکس شد. بعد از ۵ دقیقه استراحت لوله دوباره ۱۰ ثانیه ورتکس شد. پس از ۵ دقیقه استراحت دوم مقدار ۱۲ میلی‌لیتر محلول سدیمانتاسیون افزوده شد و به مدت ۴۰ ثانیه با شیکر زلنی مخلوط شد و پس از ۵ دقیقه استراحت دوباره به مدت ۴۰ ثانیه در شیکر زلنی مخلوط شد. پس از استراحت ده دقیقه به صورت عمودی حجم رسوب با واحد میلی‌متر یادداشت شد. محلول سدیمانتاسیون SDS شامل نسبت حجمی یک به ۴۸ از SDS بیست درصد و اسید لاکتیک ۸۵ درصد در آب تهیه شد. برای اندازه‌گیری فعالیت آنزیم آلفا آمیلاز آرد یا عدد فالینگ از روش AACC 56-81.04 استفاده شد. برای خوانش مقدار عدد فالینگ حدود هفت گرم از آرد کامل گندم توزین و در ۲۵ میلی‌لیتر آب مقطر درون لوله‌های آزمایش مخصوص حل و سپس درون دستگاه قرار گرفته و عدد فالینگ بر حسب ثانیه خوانده شد.

استفاده شد (Bernardo et al. 2012). داده‌های حاصل از توالی‌یابی با روش‌های ddRAD-Seq^۱ و GBS^۲ برای دو منظور کشف SNP و ژنوتیپ‌سنجی استفاده شده‌اند (Elshire et al. 2016; Shirasawa et al. 2011). روش ddRAD-Seq به وسیله Shirasawa et al. (2016) در گوجه فرنگی استفاده و شیوه تحلیل داده‌های تجربی و in silico را معرفی کردند. Armion et al. (2020) از روش ddRAD-Seq در ژنوتیپ‌سنجی نشانگرهای SNP و گروه‌بندی لاین‌های پیشرفته گندم نان استفاده نمودند. Arafa et al. (2017) از روش ddRAD-Seq برای شناسایی ژن‌های کاندید مقاومت به بیماری بادزدگی در جمعیت نسل دوم گوجه فرنگی استفاده کردند. (DaCosta & Sorenson 2014) نتیجه گرفتند که روش ddRAD-Seq یک روش مقرون به صرفه برای تولید داده‌های فیلورنی قوی به ویژه برای تجزیه و تحلیل گونه و جنس نزدیک می‌باشد. لذا هدف از اجرای این پژوهش ارزیابی اثرات اصلی و متقابل ژنوتیپ × محیط و برآورد و مقایسه ضریب همبستگی ارزش اصلاحی ژنومی برخی از صفات موثر بر کیفیت نانوائی با استفاده از روش‌های رگرسیون LASSO، رگرسیون ریح و رگرسیون شبکه الاستیک و تعیین بهترین روش برای هر صفت بود.

مواد و روش‌ها

جهت تعیین اثرات اصلی و متقابل ژنوتیپ × محیط و ارزش‌های اصلاحی ژنومی برآورد شده در جمعیت پیشرفته اصلاحی، آزمایشی بصورت طرح آلفا لاتیس در سه تکرار با ۵۰ لاین و رقم‌های شاهد (مروارید و گنبد) در ایستگاه تحقیقاتی چرداول مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان ایلام با عرض جغرافیائی ۳۳ درجه و ۴۸ دقیقه شمالی و طول جغرافیائی ۴۶ درجه و ۳۵ دقیقه شرقی با ارتفاع ۱۰۴۵ متر از سطح دریا به مدت دو سال (فصل زراعی ۱۳۹۴-۹۵ و ۱۳۹۵-۹۶) و با آبیاری نرمال اجرا شد. هر تکرار از ۵ بلوک ناقص و هر یک دارای ده کرت و هر کرت از ۶ خط به طول ۴ متر با فاصله دو خط کشت ۲۰ سانتی‌متر بود. توزیع تیمارها به هر بلوک ناقص و کرت بر

¹ Double digest restriction-site associated DNA sequencing

² Genotyping by sequencing

³ Near Infrared Reflectance

جدول ۱- شجرنامه لاین و ارقام شاهد زراعی جمعیت اصلاحی

Table 1. Lines and checked cultivars pedigrees of the breeding population

شماره	لاین / ژنوتیپ	شماره	لاین / ژنوتیپ
1	Morvarid	26	SUP152*2/TECUE#1
2	Gonbad	27	SUP152*2/TECUE #1...
3	SITTE/MO//PASTOR/3/TILHI....	28	SITTE/MO//PASTOR/3/TILHI/4/M..
4	REEDLING#1	29	SAAR//INQALAB 91*2/KUKUNA...
5	ALTAR84/AE.SQUARROSA.....	30	WHEAR/VIVITSI//WHEAR/3/WH...
6	NAC/TH.AC//3*PVN/3/MIRL....	31	MILAN/KAUZ//BABAX/3/BAV92...
7	CHIBIA//PRLII/CM65531/3/S....	32	PAURAQ//ND643/2*WBLL1/3/PA...
8	KACHU//WBLL*2/BRAMBLING	33	WHEAR/VIVITSI//WHEAR*2/3/K...
9	BAJ#1*2/WHEAR	34	SOKOLL/3/PASTOR//HXL7573/2....
10	PBW343*2/KUKUNA/3/PAST....	35	SOKOLL/3/PASTOR//HXL7573/2....
11	KACHU/BECARD//WBLL1*2....	36	ND643/2*WBLL1//HEILO
12	SUP152*2/TECUE#1	37	SUP152*2/TINKIO#1
13	WHEAR/KUKUNA/3/C80.1/3.....	38	KACHU*2/3/ND643//2*PRL/2*PAS..
14	QUAIU*2/KINDE	39	ND643/2*WBLL1/4/CHIBIA//PRLII..
15	FRNCLN/NIINI #1//FRANCOL...	40	BECARD/3/PASTOR//MUNIA/ALT..
16	FRNCLN*2/TECUE#1	41	PFAU/MILAN//FISCALL/3/VORB...
17	MUTUS*/TECUE#1	42	KIRITATI//ATTILA*2/PASTOR/3...
18	FRNCLIN#1/AKURI#1//FRN...	43	CHIBIA//PRLII/CM65531/3/SKA....
19	CHIBIA//PRLII/CM65531/3/SK...	44	CHIR3/4/SIREN//ALTAR 84/AE.....
20	KACHU#1//WBLL1*2/KUKUNA	45	PREMIO/BERKUT
21	CHIBIA//PRLII/CM65531/3/SKA.	46	MILAN/SHA7/3/THB"S"/TON"S"...
22	CHIBIA//PRLII/CM65531/3/SW89.	47	GASPARD//MILAN/SHA7/3/MIL...
23	KACHU/PVN/KACHU	48	GASPARD//MILAN/SHA7/3/MIL...
24	PCAFLR/KINGBIRD#1//KIRITAT..	49	SW89.3064/STAR//INIA/3/MILAN...
25	PCAFLR/KINGBIRD #1//KIRITAT..	50	MV17/6/ATRAK/5/4777/FKN/GB/...

جفت باز از محلول گوی‌های مغناطیسی AMPure beads (Beckman Coulter, Brea, CA, USA) استفاده شد. پس از رقیق سازی DNA ژنومیک به نسبت ۱:۱۰۰ و قبل از تشکیل کتابخانه مشترک، اندیکس Read1_N5**/Read2_N7**(0.5uM) و مخلوط ترکیبات KOD -MgSO4, dNTP, 10x KOD buf -plus- (Toyobo, Osaka, Japan) plus- ver.2 افزوده شد. برنامه ddRAD-Seq برای دستگاه PCR و تکثیر قطعات به صورت؛ آغاز عمل واسرشت‌سازی در دمای ۹۵°C به مدت سه دقیقه، ۲۰ مرحله بصورت عمل واسرشت: دمای ۹۴°C در ۳۰ ثانیه، اتصال: دمای ۵۵°C در ۳۰ ثانیه، تکثیر: دمای ۷۲°C در یک دقیقه و سه دقیقه انتهایی بعد از اتمام مرحله عمل تکثیر دمای ۷۲°C انتخاب شد. برای تعیین غلظت DNA ژنومیک از بافر Qubit® dsDNA HS Reagent و محلول واکنش گر Qubit® dsDNA HS با حساسیت بالا و دو استاندارد ۱ (Qubit™ ds DNA HS) و ۲ (Qubit™ ds DNA HS) استفاده شد. پس از عمل PCR، از هر

به منظور استخراج DNA، از بافت برگ گیاهچه‌های با سن چهارده روز نمونه‌برداری و فوراً به آزمایشگاه موسسه تحقیقات DNA کازوسا (www.kazusa.or.jp/en/) منتقل و در فریزر ۸۰°C قرار گرفت. از دستگاه TissueLyser II (Qiagen Inc., Hilden, Germany) برای آسیاب و از پروتکل شرکت کیاژن (The DNeasy Plant Mini kit) برای استخراج DNA نمونه‌ها استفاده شد. تعیین غلظت اولیه DNA با دستگاه Nano drop ND-۱۰۰۰ (Thermo Fisher Scientific Inc., Waltham, USA) و با دستگاه Qubit® 30 Fluorometer تعیین شد. در دمای ۱۶°C به مدت ۳۰ دقیقه و ۳۷°C به مدت ۱۰ دقیقه و برای ۲۵ دوره تنظیم شد. برای حذف قطعات با طول کمتر از ۳۵۰

تجزیه واریانس روی داده‌های بدست آمده از طرح آلفا لاتیس برای هر یک از صفات در هر دو سال و میانگین سال‌ها انجام گرفت. مقایسه میانگین در مقایسه با شاهد با استفاده از آزمون LSD در سطح پنج درصد برای تمام صفات انجام گرفت. با توجه به همسان بودن نتایج تجزیه واریانس آلفا لاتیس و طرح بلوک‌های کامل تصادفی، برای تعیین اثر متقابل ژنوتیپ × محیط (سال) تجزیه مرکب در قالب طرح بلوک‌های کامل تصادفی برای دو سال انجام گرفت (SAS 9.2).

تمام روش‌های برآورد ارزش‌های اصلاحی ژنومی¹ بر اساس مدل پیش‌بین پایه (Moser et al. 2009) ذیل بوده:

$$y_i = \sum_{k=1}^p x_{ik} \beta_k + e_i$$

متغیر y_i فنوتیپ فرد i و x_{ik} بردار ستونی $1 \times p$ ژنوتیپ‌های نشانگر SNP فرد i در مکان ژنی k از p مکان ژنی و β_k اثر k امین نشانگر SNP و e_i مقدار باقیمانده است.

$$y_i = g(X_i) + e_i \text{ یا}$$

$$y_i = \text{میانگین فنوتیپ افراد یا لاین‌ها}$$

$$X_i = \text{بردار } 1 \times p \text{ ژنوتیپ‌های SNP روی فرد } i$$

$$g(X_i) = \text{تابع رابط ژنوتیپ‌ها به فنوتیپ‌ها}$$

$$e_i = \text{مقدار باقیمانده}$$

به طور کلی مقدار GEBV برابر با $g(X_i)$ است. شباهت دیگر بین روش‌ها این است که همه آنها به دنبال حداقل کردن یک تابع هزینه معین هستند. در آنالیز حداقل مربعات تابع هزینه برابر با حداقل مربعات باقیمانده $\sum e_i^2$ است. در این تحقیق از روش‌های آماری مختلف که اثر نشانگر را به‌عنوان پیش‌بین تصادفی در نظر گرفته استفاده شد.

روش رگرسیون ریح ابتدا به وسیله Whittaker et al. (2000) در زمینه تلاقی‌های دو والدی در روش انتخاب به کمک نشانگر پیشنهاد و به صورت ذیل محاسبه می‌شود:

$$g(X_i) = \sum_{k=1}^p x_{ik} \beta_k$$

$$x_{ik} = \text{امتیاز } k \text{ نشانگر SNP در فرد } i \text{ ام}$$

$$\beta_k = \text{اثر مرتبط با نشانگر } k \text{ ام}$$

نمونه حدود چهار میکرولیتر برداشته و در هشت الی نه میکروسانتریفیوژ ۲۰۰ میکرولیتر جمع‌آوری و سپس همه مقادیر در یک میکروسانتریفیوژ ۱/۵ میلی‌لیتری جمع و کتابخانه تشکیل شد و توسط دستگاه کیوبیت غلظت آن سنجیده شد. با استفاده از دستگاه BluePippin قطعات DNA ژنومیک در محدوده 500~800 جفت باز انتخاب و خالص‌سازی و کتابخانه آزمایشگاهی برای عمل توالی‌یابی آماده شد. تعیین کمیت و کیفیت DNA ژنومیک کتابخانه بوسیله دستگاه‌های Qubit (ds DNA HS) و Tape Station Hs D1000 انجام شد. برای تعیین غلظت دقیق DNA ژنومیک برحسب پیکومولار از کیت Kapa Library Quantification Kit (KAPA Kit) استفاده و غلظت بوسیله دستگاه 7900HT Fast Real-Time PCR system شرکت Life Technologies تعیین شد. پس از تصحیح رقت کتابخانه به یک نانو مولار واسرشت و رقیق سازی بوسیله محلول 0.2N NaOH، بافر HT1 و PhiX control انجام گرفت. عملیات بارگذاری کتابخانه در استوانه واکنش‌گر آماده‌سازی و با دستگاه NextSeq™ 500 Sequencer توالی‌یابی شد (et al. 2020).

پس از خاتمه عمل توالی‌یابی کتابخانه، اطلاعات در قالب فایل qseq file format (qseq) ارائه شد و بر حسب نوع فناوری هر دستگاه یک فایل تعیین کیفیت توالی به نام fastQC بر مبنای نمره Phred تولید شد. پس از این مرحله برای استخراج توالی‌های صحیح از فرآیند بیوانفورماتیک استفاده شد (Shirasawa et al. 2016). ابتدا توالی‌های با کیفیت خوانش پایین بوسیله نرم افزار PRINSEQ و سپس آداپتورها به وسیله fastx_clipper حذف شدند. پس از آن، خوانش‌های فیلتر شده نسبت به توالی مرجع به نام IWGSC RefSeq v1.0 با استفاده از نرم افزار bowtie2 نقشه‌یابی شدند (IWGSC 2018). فایل فرمت نتیجه همردیف/ نقشه SAM (Sequence Alignment Map) به فایل فرمت همردیف/ نقشه باینری تبدیل و سپس برای خوانش SNP استفاده و در نهایت فایل VCF (variant call format) تولید شد. برای مشاهده موقعیت توالی‌های SNP برای هر ژنوم و کروموزوم و تعداد آنها از نرم افزار TASSEL نسخه 5.2.33 (Bradbury et al. 2007) استفاده شد.

¹ Genomic estimated breeding value= GEBV

حقیقت، توابع هزینه مدل‌های RR-BLUP و LASSO را می‌توان به صورت ذیل تعمیم داد:

$$\frac{1}{2} \sum_i e_i^2 + \frac{\lambda}{2} \sum_k |\beta_k|^q$$

اگر مقدار $q=1$ باشد، رابطه برابر با مدل LASSO و $q=2$ برابر با رگرسیون ریج خواهد شد. به دلیل تفاوت عامل تنظیم λ در مدل‌ها، روش LASSO ضرایب را شدیدتر از روش ریج رگرسیون منقبض کرده و حتی بعضی از آنها برابر با صفر شده و باعث تهیه یک مدل متخلخل شده و بنابراین رگرسیون LASSO یک روش انتخاب متغیر است.

روش رگرسیون شبکه الاستیک^۲ که اگر مقدار q بین $1 \leq q \leq 2$ باشد تابع LASSO تبدیل به مدل رگرسیون EN می‌شود. این روش از مزایای دو روش رگرسیون ریج و LASSO استفاده کرده و عامل جریمه هر دو روش (L1 (عبارت جریمه رگرسیون ریج) و L2 (LASSO)) در ترکیب خطی خود دارد (Zou & Hastie, 2005). رگرسیون شبکه الاستیک با ترکیب رگرسیون LASSO و رگرسیون ریج بر معایب آن‌ها غلبه کرده و جایگزین مطمئنی برای آنها است. به این ترتیب اگر با مدلی مواجه بوده که متغیرهای توصیفی (نشانگرها) آن با یکدیگر همبستگی داشته مناسب است از رگرسیون شبکه الاستیک استفاده شود.

بسته‌های نرم افزاری: آنالیز روش‌های رگرسیون ریج بهترین برآوردگر غیر اریب خطی RR-BLUP با استفاده از بسته "rrBLUP" که به وسیله Endelman (2011) توصیف شده و همچنین روش LASSO و Elastic net با استفاده از بسته "Glmnet" (Friedman et al. 2009) انجام گرفت.

ارزیابی برآورد صحت پیش‌بینی‌های ژنومی: در این تحقیق از فرآیند اعتبار سنجی متقابل ده بار اجرا برای برآورد صحت‌های پیش‌بینی ژنومی استفاده شد (Pérez et al. 2010). در این روش جمعیت به ده قسمت مساوی تقسیم شده و لاین‌ها به صورت تصادفی به هر قسمت منتسب می‌شوند. از اطلاعات نشانگرها و فنوتیپ نه قسمت به عنوان جمعیت اصلی یا برآورد برای تهیه مدل برآورد که آثار نشانگرها در آن محاسبه شده و برای پیش‌بینی مقدار ارزش اصلاحی ژنومی برآورد شده قسمت دهم که به عنوان

ارزش اصلاحی برابر با جمع اثرات p نشانگر است. برآوردگر حداقل مربعات نرمال در این روش تعدیل شده و در نتیجه مقادیر β با استفاده از رابطه زیر برآورد می‌شود.

$$\beta^{\wedge} = (X'X + \lambda I)^{-1} X'y$$

X = ماتریس برخورداری یا وقوع یا همبندی رابط نشانگرها به افراد

I = ماتریس یک

y = بردار ارزش‌های اصلاحی برآورد شده

تفاوت بین برآوردگر روش رگرسیون ریج β و برآوردگر حداقل مربعات معمولی، جمله λI است. که باعث ناویژه شدن ماتریس $X'X$ و کاهش همبستگی بین پیش‌بین‌گرها می‌شود. بنابراین تفاوت روش ریج با روش حداقل مربعات کلاسیک در وجود پارامتری به نام پارامتر λ که به پارامتر جریمه یا محدود کننده یا تعدیل یا تنظیم کننده معروف است. بنابراین تابع هزینه رگرسیون ریج برابر است با:

$$\frac{1}{2} \sum_i e_i^2 + \frac{\lambda}{2} b'b$$

b = بردار β است

λ = عامل جریمه

در این روش، اثرات نشانگر را به صورت افزایش بر اساس مدل بی‌نهایت در نظر گرفته و تمام نشانگرها دارای یک واریانس مشترک (همگن) و اثر آنها در جهت صفر منقبض یا کاهش شده، اما به تمام نشانگرها اجازه اثرات غیر یکنواخت می‌دهد (Meuwissen et al. 2001; Heffner et al. 2011a; He et al. 2016).

روش رگرسیون LASSO^۱ که رابطه بین LASSO و رگرسیون ریج به واسطه رابطه تابع هزینه LASSO به صورت زیر است:

$$\frac{1}{2} \sum_i e_i^2 + \frac{\lambda}{2} \sum_k |\beta_k|$$

در مدل رگرسیون ریج (RR-BLUP) یک عامل جریمه برای نشانگر با اثرات بزرگ (ضرایب رگرسیون) وجود دارد. در

¹ Least absolute shrinkage and selection operator = LASSO

² Elastic net

جمعیت اعتبار یا آزمون از اطلاعات نشانگری لاین‌های آن استفاده شد. این فرآیند ده بار به صورت تصادفی اجرا شد و در مرحله صحت پیش‌بین نشانگرها توسط ضریب همبستگی پیرسون بین مقادیر پیش‌بینی شده و رخ‌نمودها (مشاهدات حقیقی) ارزیابی شد. متوسط ده بار اجرا به عنوان صحت پیش‌بین ضریب همبستگی برآورد ارزش‌های اصلاحی ژنومی گزارش شد.

$$r_G = (GEBV: EBV)$$

r = ضریب همبستگی پیرسون بین ارزش اصلاحی ژنومی برآورد شده و ارزش اصلاحی حقیقی یا مشاهده‌ای
 $GEBV$ = ارزش اصلاحی ژنومی برآورد شده: از طریق اطلاعات رخ‌نمود صفت مورد مطالعه و نشانگر
 EBV = ارزش اصلاحی حقیقی برآورد شده: از طریق اطلاعات رخ‌نمود صفت مربوطه (مشاهده)

نتایج و بحث

تجزیه واریانس: پس از جمع‌آوری داده‌های سنجش رخ‌نمود

جدول ۲- میانگین صفات کیفیت نانوائی در سال اول، دوم و متوسط دو سال

صفت آرد بذر	سال اول (زراعی ۹۵-۱۳۹۴)	سال دوم (زراعی ۹۶-۱۳۹۵)	میانگین دو سال
پروتئین خام (%)	۱۰/۵۷	۱۰/۷۶	۱۰/۶۹
عدد فالینگ (ثانیه)	۳۸۵/۲	۴۷۰	۴۲۸
عدد زلنی	۲۴/۹۹	۲۱/۹۳	۲۳/۷۲
اندازه حجم رسوب SDS (میلی‌متر)	۴۴/۷۹	۴۳/۹۶	۴۴/۶۸

نسبت به رقم گنبد تولید و در گروه a قرار گرفتند. لاین‌های شماره ۶، ۱۵، ۲۱، ۲۲، ۲۳ و ۲۴ به طور معنی‌دار بالاترین عدد زلنی نسبت به ارقام مروارید و گنبد تولید و در گروه a قرار گرفته، همچنین لاین‌های شماره ۵، ۶، ۸، ۱۰، ۱۱، ۱۲، ۱۷، ۱۸، ۲۰، ۲۹، ۳۰، ۳۲، ۳۶، ۳۹، ۴۴ و ۴۷ بصورت معنی‌دار بیشترین اندازه حجم رسوب SDS نسبت به ارقام مروارید و گنبد تولید و در گروه a قرار گرفتند. تجزیه مرکب صفات مورد تحقیق برای عامل لاین معنی‌دار بود که بیانگر وجود تنوع ژنتیکی بین آن‌ها

است. همچنین اثر سال و اثر متقابل لاین × سال برای همه صفات معنی‌دار بود که معنی‌دار بودن اثر سال نشان دهنده وجود اختلاف بین میزان و پراکنش باران، میزان درجه حرارت و شرایط اقلیمی در دو سال زراعی متوالی و عوامل غیر قابل کنترل محیط است. لذا این یافته بیانگر آن است که محیط (سال) تاثیر شدیدی بر مقدار هر صفت داشته و بنابراین برای تعیین مقدار رخ‌نمود بایستی حداقل از اطلاعات دو سال استفاده کرد. علاوه بر آن عامل اثر متقابل لاین × سال بر صفات کیفیت نانوائی از قبیل

میزان پروتئین خام، عدد فالینگ، عدد زلنی و اندازه حجم رسوب SDS به شدت اثر گذار بوده و میزان رخ نمود هر یک از لاین‌ها برای هر کدام از صفات از سالی به سال دیگر اختلاف داشت (جدول ۴). (Comstock and Moll 1963) محیط را به دو دسته میکرو؛ یک مکان در یک سال و ماکرو؛ یک یا بیش یک مکان در دو یا چند سال طبقه‌بندی کردند. Comstock and Robinson (1952) و Matzinger (1963) بر اهمیت ارزیابی اثر متقابل ژنوتیپ × محیط تاکید کردند. جهت انتخاب موثر برای یک صفت خاص، تعیین کمیت هر عامل ژنوتیپ (G)، محیط (E) و اثر متقابل ژنوتیپ در محیط (GEI) سهم در تنوع صفات کمی و کیفی گندم اهمیت دارد (Mladenov, 2001; Abugalieva and Peña, 2010). لذا درک این عوامل به وسیله به‌نژادگران ضروری بوده و در طراحی اهداف و راهبردهای مناسب برای تهیه و اصلاح ارقام گندم با ظرفیت عملکرد بالا و همچنین با خصوصیات کیفی ثابت و ویژه در جهت تامین نیازهای بازار راهگشا است (Williams et al. 2008). اجزای عملکرد بصورت واضح تعیین کننده اصلی درآمد کشاورز، در صورتیکه کیفیت بذر برای کیفیت به‌نژادی اهمیت دارد (Mladenov et al. 2016). کیفیت بذر یک صفت مرکب و پیچیده است که به تعدادی از صفات وابسته است و سهم هر صفت با توجه به عکس‌العمل اختصاصی در شرایط محیطی متغیر است (Peterson et al. 1998). تعیین ارتباط بین شاخص‌های کیفی در تعریف مقادیر بهینه از کیفیت بذر برای یک ناحیه خاص مورد توجه زیادی قرار گرفته است و به به‌نژادگران در ایجاد ارقام با کیفیت مناسب کمک می‌کند (Desheva 2016). همچنین سهم هر یک از عوامل موثر بر تنوع از جمله لاین، محیط (سال) و اثر متقابل لاین × محیط در جدول شماره ۵ گزارش شده که بزرگی اثر هر یک از عوامل بسته به نوع هر صفت متغیر بود. سهم هر یک از اجزای واریانس موثر در ایجاد تنوع یعنی ژنوتیپ (لاین)، محیط (سال) و ژنوتیپ در محیط برای صفات کیفی به ترتیب برای پروتئین؛ ۶۷/۳۰، ۱/۴۷ و ۹/۳۳، عدد فالینگ؛ ۳۹/۹۸، ۱۴/۳۳ و ۳۷/۶۱، عدد زلنی؛ ۴۶/۹۸، ۲۵/۱۶ و ۴/۹۰ و اندازه حجم رسوب (SDS)؛ ۷۳/۴۹، ۰/۴۱ و ۹/۰۵ بود. این نتایج نشان داد که در صفات پروتئین عامل ژنوتیپ، عدد فالینگ؛ ژنوتیپ و اثر متقابل (تقریباً به طور

مساوی)، عدد زلنی؛ در درجه اول ژنوتیپ و سپس اثر متقابل و در اندازه حجم رسوب SDS ژنوتیپ بیشترین سهم در ایجاد تنوع داشتند. ترکیب خصوصیات مورفولوژیکی ژنوتیپ و بر همکنش با محیط، بیان رخ نمود صفات تولیدی و کیفیت نانویی را تعیین کرده و اثرات ژنوتیپ، محیط و متقابل نقش مهمی در ایجاد ارقام جدید و موفقیت برنامه‌های به نژادی دارد (Desheva 2016). فرآیند به نژادی و تولید نه‌تنها به عوامل ژنتیک، بلکه به عوامل محیطی نیز وابسته است. تفاوت معنی‌دار رخ نمود بین ژنوتیپ‌های گندم با توجه به بذر و کیفیت آرد وجود دارد (Mladenov et al. 2001). برای به‌نژادگر علاوه بر آن‌که میانگین ژنوتیپ برای هر صفت، پایداری آن صفت نیز که به اثر متقابل ژنوتیپ در محیط وابسته بوده مورد توجه است (Zecevicz et al. 2009). تحقیقات فراوانی بر روی اثر متقابل ژنوتیپ در محیط شاخص‌های کیفیت نان انجام شده است (Desheva 2016). (Šimićet et al. 2006) گزارش کردند که عامل محیط به شدت روی مقدار پروتئین خام بذر موثر است. ژنوتیپ و محیط اثر قوی‌تری بر واریانس مقدار حجم رسوب نسبت به اثر متقابل ژنوتیپ در محیط دارد (Zecevicz et al. 2009). (Desheva 2016) نشان داد که اثرات ژنوتیپ، محیط و ژنوتیپ در محیط برای تمام صفات کیفیت نانویی بذر بررسی شده معنی‌دار و بزرگی آن‌ها متغیر بوده که عوامل ژنوتیپ بر روی میزان حجم رسوب و متقابل روی واریانس پروتئین خام قوی‌ترین اثر را داشتند. پروتئین بذر یک شاخص کیفی مهم است که میزان فراوان آن باعث کیفیت بهتر می‌شود (Vázquez et al. 2012). بر اساس نتایج Williams et al. (2008) افزایش میزان پروتئین تحت تاثیر اثر متقابل ژنوتیپ در محیط و اثر محیط قرار می‌گیرد.

مطالعات دیگر نشان داد که شرایط محیطی دارای اثر بیشتری روی میزان پروتئین نسبت به ژنوتیپ می‌باشد (Mut et al. 2010; Surma et al. 2012). Drezner et al. (2010) در تفسیر اثر متقابل ژنوتیپ در محیط برای صفات کیفی (رسوب زلنی و عدد فالینگ) بر روی ده واریته گندم نان زمستانه در آزمایشی که در یک ایستگاه و برای شش فصل زراعی انجام شد، بیان کردند که اندازه اثر GEI روی عدد فالینگ کمتر از اثر ژنوتیپ است. بر اساس

تعدادی از فعالیتهای پژوهشی که به وسیله محققین مختلف انجام و توسط Peña et al. (2002) گزارش شد اساساً سهم تنوع اثر متقابل GEI در صفات کیفیت نانوائی بذر کمتر از ژنوتیپ است. لذا نتایج این پژوهش برای صفات کیفیت نانوائی بذر در تطابق کامل با Peña et al. (2002) و Drezner et al. (2010) بود.

جدول ۳- تجزیه واریانس صفات کیفیت نانوائی بذر گندم نان در قالب طرح آلفا لاتیس

میانگین مربعات سال اول					
منابع تغییر	درجه آزادی	پروتئین (%)	عدد فالینگ (ثانیه)	عدد زلنی	اندازه حجم رسوب SDS (میلی متر)
تکرار	۲	۰/۰۳۸۶	۴۳۲۴/۱	۷۱/۲۵	۹/۷۸
بلوک (تصحیح شده)	۱۲	۰/۱۳۲۵	۱۵۷۳/۷۱	۷/۱۳	۷/۲۹
تیمار	۴۹	۱/۲۱۶۲	۲۶۵۰۷/۵۵	۹۴/۷۳	۷۹/۰۷
تیمار (تصحیح شده)	۴۹	۱/۲۰۹۱***	۲۴۳۵۶/۴۱***	۹۳/۶۸***	۷۲/۸۶***
باقیمانده (خطا)	۸۶	۰/۱۳۷	۲۷۰۱/۴۷	۷/۹۹	۹/۹۶
کل	۱۴۹	۰/۴۸۶۶	۱۰۴۶۱/۲۸	۳۷/۳	۳۲/۴۷
ضریب تغییرات (%)		۳/۴	۱۱	۱۲/۹	۷/۲
LSD (۰/۰۵)		۰/۵۸۸۷	۸۴/۳۶۷۴	۴/۵۸۸۷	۵/۱۲۱۸
میانگین مربعات سال دوم					
تکرار	۲	۰/۱۸۰۹	۱۳۵۶/۵۱	۱/۶۴۶۷	۸/۹۹
بلوک (تصحیح شده)	۱۲	۰/۲۷۰۲	۵۲۶/۶۸	۲۴/۴۵	۱۰/۳۸
تیمار	۴۹	۱/۵۱۹۹	۳۳۲۵۳/۳۴	۱۱۶/۲۷	۱۳۲/۸۶
تیمار (تصحیح شده)	۴۹	۱/۴۳۰۴***	۲۸۹۰۰/۷***	۱۱۴/۶۹***	۱۱۵/۳۸***
باقیمانده (خطا)	۸۶	۰/۲۵۳۵	۴۱۷/۸۵۹۷	۲۴/۱۸	۱۲/۰۷
کل	۱۴۹	۰/۶۷۰۴	۱۱۲۳۷/۴۷	۵۴/۱۸	۵۱/۶۱
ضریب تغییرات (%)		۴/۸	۵/۳	۱۹/۷	۷/۸
LSD (۰/۰۵)		۰/۸۲۹۳	۳۴/۷۸۱۱	۸/۰۰۲۶	۵/۶۳۹۷
میانگین مربعات متوسط دوسال					
تکرار	۲	۰/۰۸۶۹	۲۲۳۳/۶۵	۱۱/۵۴	۷/۷۶
بلوک (تصحیح شده)	۱۲	۰/۱۲۰۱	۴۹۲/۶۳	۹/۸۱	۶/۹۱
تیمار	۴۹	۱/۲۰۴۲	۱۵۴۰۱/۵۲	۶۸/۱۳	۹۵/۵۸
تیمار (تصحیح شده)	۴۹	۱/۱۶۰۵***	۱۳۵۸۳/۴۴***	۶۷/۱۹***	۸۴/۴۲***
باقیمانده (خطا)	۸۶	۰/۱۲۱	۷۲۰/۴۲	۸/۸۷	۶/۴۸
کل	۱۴۹	۰/۴۷۶۷	۵۵۵۰/۴	۲۸/۴۷	۳۵/۸۴
ضریب تغییرات (%)		۳/۳	۶/۳	۱۲/۶	۵/۷
LSD (۰/۰۵)		۰/۵۶۴۶	۴۳/۵۶۷۹	۴/۹۴۴۵	۴/۱۹۳۴

جدول ۴- تجزیه واریانس مرکب صفات کیفیت بذر گندم نان برای دوسال

میانگین مربعات				درجه آزادی	منابع تغییر
اندازه حجم رسوب SDS (میلی متر)	عدد زلنی	عدد فالینگ (ثانیه)	پروتئین (%)		
۵۲/۰۸۳۳۳*	۷۰۲/۲۷***	۵۴۱۰۲۵/۳۳***	۲/۵۷۶۱۳۳۳***	۱	سال
۹/۳۸۳۳۳ ^{ns}	۳۶/۴۴۶۷ ^{ns}	۲۸۴۰/۷۴۷ ^{ns}	۰/۱۰۹۷۳۳۳ ^{ns}	۴	تکرار در سال
۱۸۸/۶۸۵۷۸۲***	۱۳۷/۴۱۰۱۳۶***	۳۰۷۹۲/۳۶۱***	۲/۴۰۲۹۲۲۴***	۴۹	لاین
۲۳/۲۳۹۷۹۶***	۷۳/۵۸۹۷***	۲۸۹۶۸/۵۴۴***	۰/۳۳۳۲۰۸۲**	۴۹	لاین × سال
۱۰/۷۴۷۲۸	۱۶/۰۴۸۷۱	۱۴۹۷/۲۷۷	۰/۱۹۳۲۳۶۷	۱۹۶	خطای دوم
۴/۹۲۳۳	۶/۰۱۶۲	۵۸/۱۱۱	۰/۶۶۰۲	ضریب تغییرات (%)	
۳/۷۳۲۷	۴/۵۶۱۴	۴۴/۰۵۸	۰/۵۰۰۵	LSD (0.01)	
۷/۳۸۷۵	۱۷/۰۷۴	۹/۰۳۸۷	۴/۱۲۰۹	LSD (0.05)	

ns, *, **, و *** به ترتیب عدم اختلاف معنی دار و تفاوت معنی دار در سطح ۵، ۱ و ۰/۱ درصد دارند.

جدول ۵- سهم هر یک از اجزای واریانس لاین، محیط (سال) و اثر متقابل در مقدار تنوع صفات کیفیت نانویی بذر

میانگین مربعات				درجه آزادی	منابع تغییر
حجم رسوب (میلی متر)	عدد زلنی	عدد فالینگ (ثانیه)	پروتئین (%)		
۰/۴۱	۴/۹۰	۱۴/۳۴	۱/۴۷	۱	سال
۰/۳۰	۱/۰۲	۰/۳۰	۰/۲۵	۴	تکرار در سال
۷۳/۴۹	۴۶/۹۸	۳۹/۹۸	۶۷/۳۰	۴۹	لاین
۹/۰۵	۲۵/۱۶	۳۷/۶۱	۹/۳۳	۴۹	لاین × سال
۱۶/۷۴	۲۱/۹۵	۷/۷۸	۲۱/۶۵	۱۹۶	خطای دوم

گزارش شد. همچنین تعداد کل SNP های صحیح فراخوانی شده برای ۵۰ درصد داده گم شده برابر با ۳۳۴۲ عدد، که تعداد ۱۳۲۲ معادل ۳۹/۵۶ درصد روی ژنوم B، ۱۲۵۳ معادل ۳۷/۴۹ درصد روی ژنوم A و ۷۶۷ معادل ۲۲/۹۵ درصد روی ژنوم D شناسایی شد (Armion et al. 2020).

تجزیه ژنومی: بر روی جمعیت مورد تحقیق با انجام فیلتر اول و دوم برای فراوانی آلل های نادر کمتر از ۵ درصد و میزان ناخالصی بالاتر از ۱۰ درصد برای داده های گم شده ۲۰٪، ۳۰٪، ۴۰٪ و ۵۰٪ نسبت به ژنوم مرجع پایگاه داده IWGSC RefSeq v1.0. بیشترین SNP خوانش شده برای ۵۰ درصد داده های گم شده

جدول ۶- خصوصیات توصیف آماری صفات کیفیت نانوائی بذر گندم نان

صفت	سال	کمینه	اولین چارک	میانه	میانگین	سومین چارک	بیشینه	انحراف	
								معیار	معیار
عدد زنی	۱	۱۱/۳۲	۲۰/۳۳	۲۵/۶۶	۲۴/۹۹	۲۸/۹۹	۳۹/۳۴	۶/۲۲۶۰	۰/۱۶
	۲	۹	۱۷/۶۷	۲۲/۵۰	۲۱/۹۳	۲۶/۹۲	۳۰	۵/۶۱۹۵	۲/۲۷
متوسط دو سال		۱۲/۶۹	۲۰/۲۸	۲۵/۱۴	۲۳/۷۲	۲۶/۶۷	۳۲/۳۵	۴/۸۷۵۲	۲/۴۵
	۱	۲۹/۳۳	۴۰/۱۷	۴۵	۴۴/۷۹	۵۰	۵۷/۶۷	۶/۶۵۵۱	۰/۸۳
اندازه حجم رسوب SDS	۲	۳۱/۶۷	۴۰/۱۷	۴۴/۳۳	۴۳/۹۶	۴۶/۵۸	۵۸	۵/۱۳۳۳	۰/۹۰
		۳۱/۹۸	۴۱/۴۰	۴۴/۸۶	۴۴/۶۸	۴۹/۳۱	۵۷/۳۴	۵/۶۲۳۴	۰/۲۷
عدد فالینگ	۱	۱۴۸	۳۲۰/۵	۳۹۶	۳۸۵/۲	۴۶۳/۵	۶۵۱	۱۰۵/۱۲۰۲	۰/۴۵
	۲	۲۸۷	۴۰۶/۸	۴۶۷	۴۷۰/۳	۵۳۲/۸	۷۱۳	۹۴/۰۲۹۱	۰/۶۵
متوسط دو سال		۲۹۳	۳۶۶/۸	۴۲۸/۵	۴۲۸	۴۷۵	۶۲۷	۷۱/۶۴۲۴	۲/۰۴
	۱	۸/۹۳	۱۰/۲۳	۱۰/۶۸	۱۰/۵۷	۱۰/۹۸	۱۲/۲۴	۰/۷۱۱۸	۲/۱۱
پروتئین	۲	۹/۲۳	۱۰/۴۷	۱۰/۸۳	۱۰/۷۶	۱۱/۱۶	۱۲/۳۷	۰/۶۳۷۱	۳/۶۲
		۹/۱۰	۱۰/۳۸	۱۰/۸۷	۱۰/۶۹	۱۱/۰۷	۱۱/۹۷	۰/۶۳۳۳	۵/۰۳

برآورد همبستگی ارزش های اصلاحی ژنومی: مقادیر توصیف آماری متوسط هر یک از صفات کیفیت نانوائی برای سال اول، دوم و متوسط دو سال در جدول ۶ گزارش شده است. برای آزمون توزیع نرمال داده ها، از دو آزمون نکوئی برازش Jarque-Bera (1980) و D'Agostino (1990) جهت تعیین انطباق میزان چولگی و کشیدگی داده ها با توزیع نرمال انجام شد. بر حسب مقدار آماره و سطح معنی دار تمام صفات از توزیع نرمال هم در سال اول و دوم و هم متوسط دو سال تبعیت کردند. میزان ارزش اصلاحی ژنومی برای هر یک از صفات کیفیت نانوائی بر اساس مقدار میانگین هر صفت با روش های آماری رگرسیون LASSO، رگرسیون ریدج و رگرسیون شبکه الاستیک برآورد شده و سپس میانگین ضریب همبستگی ارزش اصلاحی ژنومی مقادیر میانگین برآورد شده با رخ نمود مشاهده شده با روش ده بار اعتبار سنجی متقابل برای سال اول، دوم و متوسط محاسبه شد (جدول ۷). دامنه تغییرات متوسط میزان ضریب همبستگی پیرسون ژنومی در صفت پروتئین در روش های آماری مختلف برای سال اول، دوم و متوسط دو سال به ترتیب بین (۰/۰۶۵۷ الی ۰/۱۰۰۳)، (۰/۰۱۴۰ - الی ۰/۰۸۰۵) و (۰/۱۳۶۰ الی ۰/۳۷۹۲)، عدد فالینگ (۰/۲۰۷۳ الی ۰/۲۲۲۸)، (۰/۴۲۲۳ الی ۰/۶۱۵۱) و (۰/۲۲۲۴ الی ۰/۲۳۶۲)، عدد زنی (۰/۱۰۹۷ الی ۰/۲۹۳۳)، (۰/۳۲۹۰ -، ۰/۰۷۵۹ -) و (۰/۰۳۹۱، ۰/۳۲۵۸) و اندازه حجم رسوب SDS (۰/۳۰۸۹ - الی ۰/۰۱۳۴ -)، (۰/۳۱۱۶ - الی ۰/۲۴۵۰ -) و (۰/۲۴۹۳ - الی ۰/۱۳۰۰ -) بود.

همچنین بهترین روش های آماری برآورد ارزش اصلاحی ژنومی صفت پروتئین (شکل ۱- الف) در سال اول؛ EN (۰/۱۰۰۳) دوم؛ EN (۰/۰۸۰۵) و متوسط؛ LASSO (۰/۳۷۹۲) و EN (۰/۳۷۴۹)، عدد فالینگ (شکل ۱- ب) در سال اول؛ EN (۰/۲۲۲۸) و LASSO (۰/۲۱۱۸)، دوم؛ EN (۰/۶۱۵۱) و LASSO (۰/۵۲۳۶)، متوسط RR (۰/۲۳۶۲)، EN (۰/۲۲۳۰) و LASSO (۰/۲۲۲۴)، عدد زنی (شکل ۱- ج) در سال اول؛ LASSO (۰/۲۹۵۳) و EN (۰/۲۶۸۱) دوم؛ هیچکدام و متوسط؛ EN (۰/۳۲۵۸) و برای صفت حجم رسوب SDS هیچگونه مدلی برازش نشد (شکل ۱- د). متوسط صحت های پیش بینی بر اساس ضریب همبستگی پیرسون در بین ترکیب روش - صفت و صفات ارزیابی شده متفاوت بود. زمانی که از یک جمعیت مرجع مشابه برای پیش بینی چندین صفت استفاده شده صحت های پیش بینی برای بعضی از صفات کمتر از سایر صفات همان طوری که وراثت پذیری برای تعدادی از صفات پایین تر از تعداد دیگر خواهد شد و لذا این یافته مورد انتظار بود (Combs and Bernardo 2013). بیشترین صحت به ازای هر صفت در سال اول؛ عدد زنی (۰/۲۹۳۳) سال دوم؛ عدد فالینگ (۰/۶۱۵۱) و در متوسط دو سال؛ میزان پروتئین (۰/۳۷۹۲) بود. این نتایج نشان می دهد که ممکن است تعدادی از صفات برای اجرای روش انتخاب ژنومی در برنامه های به نژادی مطلوب تر باشند (Haile 2018). مقایسه سه روش آماری مختلف به ازای هر صفت و سال برای برآورد متوسط صحت پیش بینی در این

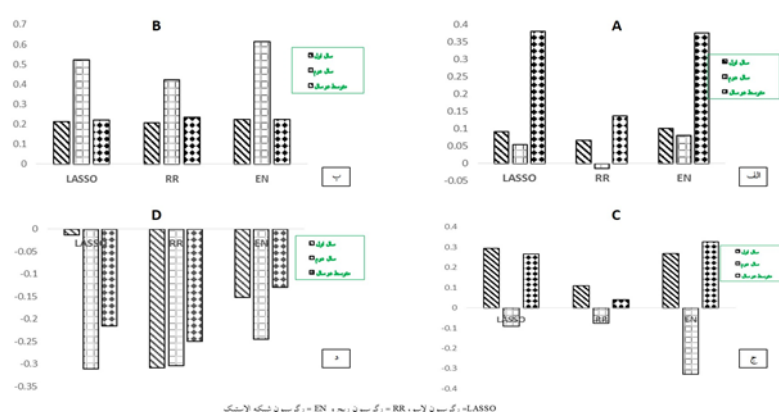
تحقیق نشان داد که با توجه به سال متغیر و عدم ثبات هر مدل در سال برای هر صفت نشانه وجود اثر متقابل بین اثر نشانگر SNP و محیط ($M \times E$) است (Haile 2018). روش رگرسیون ريج به عنوان یک روش جایگزین روش انتخاب متغیر به وسیله Whittaker و همکاران (۲۰۰۰) برای پیش‌بینی ژنومی پیشنهاد شد و روشی است که تمام نشانگرها (با اثر کوچک، بزرگ و بدون اثر) را در مدل رگرسیون در نظر می‌گیرد. در این روش از طریق پارامتر جریمه λ که به برآوردگر روش رگرسیون عادی افزوده شده اثرات نشانگرها را به طور یکنواخت در جهت عدد صفر منقبض می‌کند و فرض بر این است که هر نشانگر سهم مساوی در واریانس ژنتیکی دارد. همچنین فرض بر این است که تمام مکان‌های ژنی به مقدار مساوی میزان واریانس را توجیه می‌کند و برای محاسبه واریانس هر ژن فقط نیاز به برآورد کل واریانس بوده و واریانس هر مکان ژنی از طریق تقسیم واریانس کل به تعداد مکان‌های ژنی بدست می‌آید (Meuwissen et al. 2001). اگر چه این فرض غیر واقعی است اما روش ريج رگرسیون زمانی که پیش‌بینی برای صفاتی که به وسیله تعداد زیادی ژن جای با اثر کم کنترل شده به صورت مناسب عمل می‌کند (Lorenz et al. 2011). Tibshirani et al. (1996) گزارش کرد که صحت

پیش‌بینی روش LASSO زمانیکه تعداد کم تا متوسط از ژن‌ها با اثر متوسط وجود داشته باشد، در مقایسه با انتخاب متغیر و رگرسیون ريج برتر بوده و در صورتیکه تعداد زیادی از ژن‌ها با اثرات کم وجود داشته باشند رگرسیون ريج پیشنهاد می‌شود. به طور کلی تحقیقات سابق گزارش کردند که یک تفاوت در صحت‌های پیش‌بین ژنومی برای صفات، جمعیت‌ها و روش‌های آماری مشاهده شده و ارزیابی مدل‌های مختلف آماری برای هر جمعیت و محیط لازم است (Haile 2018). نتایج این پژوهش نیز بیانگر آن است که صفات کیفیت نانویی بذر گندم در این جمعیت مورد تحقیق تحت کنترل ژن‌ها و نشانگرهای بزرگ و متوسط اثر بوده و با شناسایی نشانگرهای SNP مرتبط با ژن‌های این صفات و صرفاً با انجام سنجش ژن‌نمود و تعیین امتیاز عددی نشانگرهای SNP با استفاده از روش انتخاب ژنومی و برآورد ارزش اصلاحی ژنومی از طریق رگرسیون‌های LASSO و رگرسیون EN در نسل‌های ابتدایی اقدام کرده و در انتخاب ژنوتیپ‌های با بیشترین ارزش اصلاحی ژنومی که دارای کیفیت بالاتر نانویی است گام موثری در شتاب بخشیدن برنامه‌های به‌نژادی و کاهش هزینه و زمان معرفی رقم با عملکرد بالا و کیفیت مناسب گام برداشت.

جدول ۷- متوسط ضرایب همبستگی صحت پیش‌بینی ($r_{GEBV:EBV}$) صفات کیفیت نانویی

ردیف	سال	روش آماری	پروتئین (%)	عدد فالینگ (ثانیه)	عدد زلنی	اندازه حجم رسوب SDS (میلی متر)
۱		*LASSO	۰/۰۹۲۰	۰/۲۱۱۸	۰/۲۹۳۳	-۰/۰۱۳۴
۲	۱	RR	۰/۰۶۵۷	۰/۲۰۷۳	۰/۱۰۹۷	-۰/۳۰۸۹
۳		EN	۰/۱۰۰۳	۰/۲۲۲۸	۰/۲۶۸۱	-۰/۱۵۲۰
۱۲		LASSO	۰/۰۵۴۰	۰/۵۲۳۶	-۰/۰۹۱۱	-۰/۳۱۱۶
۱۳	۲	RR	-۰/۰۱۴۰	۰/۴۲۲۳	-۰/۰۷۵۹	-۰/۳۰۳۵
۱۴		EN	۰/۰۸۰۵	۰/۶۱۵۱	-۰/۳۲۹۰	-۰/۲۴۵۰
۲۳		LASSO	۰/۳۷۹۲	۰/۲۲۲۴	۰/۲۶۵۲	-۰/۲۱۶۰
۲۴	متوسط	RR	۰/۱۳۶۰	۰/۲۳۶۲	۰/۰۳۹۱	-۰/۲۴۹۳
۲۵		EN	۰/۳۷۴۹	۰/۲۲۳۰	۰/۳۲۵۸	-۰/۱۳۰۰

* - LASSO = رگرسیون لاسو، RR = رگرسیون ريج و EN = رگرسیون شبکه الاستیک



شکل ۱- ضرایب همبستگی ژنومی الف- پروتئین خام، ب- عدد فالینگ، ج- عدد زنی و د- حجم رسوب در روش‌های مختلف آماری

منابع

Abugalieva A, Peña R J (2010) Grain quality of spring and winter wheat of Kazakhstan. *Asian Australasian Journal of Plant Science and Biotechnology* 4:87-90.

ALPHANAL. A program for analysis of alpha lattice incomplete block experiments version 1.1. Edinburgh: AFRC unit of statistics, 1986.

Arafa RA, Rakha MT, Soliman NEK, Moussa OM, Kamel SM, Shirasawa K (2017) Rapid identification of candidate genes for resistance to tomato late blight disease using next-generation sequencing technologies. *PLoS ONE* 12: 1-15.

Armion M, Bihamta MR, Maali Amiri R, Khodarahmi M, Isobe S (2020) Discovery and genotyping of SNP markers and grouping of advanced bread wheat lines by ddRAD-Seq. *Iranian Journal of Field Crop Science* 51:103-114. (In Farsi).

Battenfield SD, Guzmán C, Gaynor RC, Singh RP, Peña RJ, Dreisigacker S, Fritz AK, Poland JA (2016) Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. *Plant Genome* 9.

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Wilkinson DB (ed). *Proceedings of the 49th Annual Corn and Sorghum Research Conference*. Washington, DC: American Seed Trade Association 250-65.

Bennewitz J, Solberg T., Meuwissen THE (2009) Genomic breeding value estimation using nonparametric additive regression models. *Genetics Selection Evolution* 41:20.

Bernardo AN, Ma H, Zhang D, Bai G (2012) Single nucleotide polymorphism in wheat chromosome region harboring Fhb1 for Fusarium head blight resistance. *Molecular Breeding* 29:477-488.

Bernardo R, Yu J (2007) Prospects for genome wide selection for quantitative trait in maize. *Crop Science* 47:1082-1090.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for

association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633-2635

Breiman L (1995) Better subset regression using the nonnegative garrote. *Technometrics* 37:373-384.

Carter BP, Morris CF, Anderson JA (1999) Optimizing the SDS sedimentation test for end-use quality selection in a soft white and club wheat-breeding program. *Cereal Chemistry* 76:907-911.

Combs E., Bernardo R. (2013) Accuracy of genome wide selection for different traits with constant population size, heritability, and number of markers. *Plant Genome* 6:1-7.

Comstock RE, Robinson H F (1952) *Heterosis*, chapter 30, Iowa State College Press. Ames, Iowa.

Comstock RE, Moll RH (1963) *Genotype-Environment Interactions*. In: Hanson, W. D. and Robinson, H. F., Eds., *Statistical Genetics and Plant Breeding*, NAS-NRC, Washington DC, 164-196.

Crossa J, Campos G, de L Pérez P (2010) Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186:713-724.

DaCosta JM, Sorenson MD (2014) Amplification Biases and Consistent Recovery of Loci in a Double-Digest RAD-seq Protocol. *PLoS ONE* 9:1-14.

D'Agostino RB, Belanger A, D'Agostino RBJr (1990) A suggestion for using powerful and informative tests of normality. *American Statistician* 44:316-321.

Desheva G (2016) Correlation and path-coefficient analysis of quantitative characters in winter bread wheat varieties. *Trakia Journal of Sciences* 14:24-29.

Desheva G (2016) Effects of Genotype, Environment and their Interaction on Quality Characteristics of Winter Bread Wheat. *Journal of Basic Sciences and Applied Research* 2:2-13.

Drezner G, Gunjaca J, Novoselovic D, Horvat D (2010) Interpretation of GEI effect analysis for some agronomic and quality traits in ten winter wheat (*Triticum aestivum* L.) cultivars. *Cereal Research Communications* 38:260-266

- Elliot L., Heffner EL, Sorrells ME, Jannink JL (2009) Genomic Selection for Crop Improvement. *Crop Science* 49:1-12.
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES et al. (2011) A robust, simple genotyping - by-sequencing (GBS) approach for high diversity species. *PLoS ONE*. 6:e19379.
- Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome* 4:250-255.
- Friedman J, Hastie T, Tibshirani R (2009) glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models. R package version 1.1-4, URL <http://CRAN.R-project.org/package=glmnet>.
- Gianola D, van Kaam JBCH M (2008) Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178:2289-2303.
- Gonzalez-Recio O, Forni S (2011) Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. *Genetics Selection Evolution* 43:7.
- Haile TA (2018) Genomic Selection, Quantitative Trait Loci and Genome-Wide Association Mapping for Spring Bread Wheat (*Triticum aestivum* L.) Improvement. Dissertation, University of Saskatchewan, Saskatoon, Canada.
- Hayashi T, Iwata H (2010) EM algorithm for Bayesian estimation of genomic breeding values. *BMC Genetics* 11:1-9.
- Hayes BJ, Panozzo J, Walker CK et al. (2017) Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. *Theoretical and Applied Genetics* doi: 10.1007/s00122-017-2972-7
- He S., Schulthess AW, Mirdita V, Zhao Y, Korzun V, Bothe R, Ebmeyer E, Reif JC, Jiang Y (2016) Genomic selection in a commercial winter wheat population. *Theoretical and Applied Genetics* 129:641-651.
- Heffner EL, Jannink JL, Sorrells ME 2011a. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome* 4:65-75.
- Henderson C (1984) Application of linear models in animal breeding. University of Guelph, Ontario, Canada
- International Wheat Genome Sequencing Consortium (2018) Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science* 361.
- Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. *Briefings in Functional Genomics* 9:166-177.
- Jarque CM, Bera AK (1980) "Efficient tests for normality, homoscedasticity and serial independence of regression residuals". *Economics Letters* 6:255-259.
- Kristensen PS, Jahoor A, Andersen JR, Cericola F, Orabi J, Janss LL, Jensen J (2018) Genome-Wide Association Studies and Comparison of Models and Cross-Validation Strategies for Genomic Prediction of Quality Traits in Advanced.
- Long N, Gianola D, Rosa GJM, et al. (2007) Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. *Journal of Animal Breeding and Genetics* 124:377-89.
- Lorenz AJ, Chao S, Asoro FG, Heffner EL, Hayashi T, Iwata H, Smith KP, Sorrells ME, Jannink JL (2011) Genomic selection in plant breeding: knowledge and prospects. In: D.L. Sparks, editor, *Advances in agronomy*. Elsevier. San Diego 110:77-123.
- Matzinger DF (1963) Experimental estimation of genetic parameters and their applications in self-fertilizing plants. *Statistical Genetics and Plant Breeding Symp.* Raleigh, 1961: 252-276.
- Meuwissen THE, Hayes BJ, Goddard, ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-29.
- Mladenov N, Przulj N, Hristov N, Duric V, Milovanovic M (2001) Cultivar-by-environment interactions for wheat quality traits in Semi arid conditions. *Cereal Chemistry* 78:363-367.
- Moser G, Tier B, Crump RE, Khatkar MS, Raadsma H W (2009) A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. *Genetics Selection Evolution* 41:56.
- Mut Z, Aydin N, Bayramoglu HO, Ozcan H (2010) Stability of some quality traits in bread wheat (*Triticum aestivum*) genotypes. *Journal of Environmental Biology* 31:489-495.
- Nejati-Javaremi A, Smith C, Gibson, JP (1997) Effect of total allelic relationship on accuracy of evaluation and response to selection. *Animal Science Journal* 7:1738-1745.
- Peña RJ, Trethowan R, Pfeiffer WH, van Ginkel M (2002) Quality (end-use) improvement in wheat: compositional, genetic, and environmental factors. *Journal of Environmental Biology* 5:1-37.
- Perez P, de los Campos G, Crossa J, Gianola D (2010) Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian linear regression package in R. *Plant Genetic Resources* 3:106-116
- Peterson CJ, Graybosch RA, Shelton DR, Baenziger PS (1998) Baking quality of hard red winter wheat: Response of cultivars to environments in the Great Plains. *Euphytica* 100:157-162.
- Shirasawa K, Hirakawa H, Isobe S (2016) Analytical workflow of double-digest restriction site-associated DNA sequencing based on empirical and in silico optimization in tomato. *DNA Research* 23:145-153.
- Šimićet G, Horvat D, Jurković Z, Drezner G, Novoselović D, Dvojković K (2006) The genotype effect on the ratio of wet gluten content to total wheat grain protein. *Journal of Central European Agriculture* 7:13-18.
- Surma M, Adamski T, Banaszak Z, Kaczmarek Z, Kuczyńska H, Majcher M, Ługowska B, Obuchowski W, Salmanowicz B, Krystkowiak K (2012) Effect of genotype, environment and their interaction on quality parameters of wheat breeding lines of diverse grain hardness. *Plant Production Science* 15:192-203.
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*. 58:267-288.

Van Raden PM (2008) Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* 91: 4414-4423.

Vázquez D, Bergera AG, Cuniberti M, Bainotti C, Zavariz de MM, Luiz SP, Jobet C, Zúñiga J, Cabrera G, Verges R, Peña RJ (2012) Influence of cultivar and environment on quality of Latin American wheat. *Journal of Cereal Science* 56:196-203.

Wei Z, Wang K, Qu HQ et al. (2009) From disease association to risk assessment: an optimistic view from genome-wide association studies on type 1 diabetes. *PLOS Genetics* 5:e1000678.

Whittaker JC, Thompson R, Denham, MC (2000) Marker-assisted selection using ridge regression. *Genetics Research* 75:249-252.

Williams RM, O'Brien L, Eagles HA, Solah VA, Jayasena V (2008) The influences of genotype, environment, and genotype-environment interaction on wheat quality. *Australian Journal of Agricultural Research* 59: 95-111.

Zecevic V, Knežević D, Bošković J, Madić M (2009) Effect of genotype and environment on wheat quality. *Genetika* 41: 247 -253.

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B* 67: 301-320