

## مقایسه بین BayesC و GBLUP در برآورد ارزش‌های اصلاحی ژنومی در توزیع‌های متفاوت واریانس ژن‌های عمده اثر

مسعود شیرعلی<sup>۱\*</sup>، سید رضا میرائی آشتیانی<sup>۲</sup>، عباس پاکدل<sup>۳</sup>، کریس هیلی<sup>۴</sup> و ریکاردو پونگک-ونوگک<sup>۵</sup>  
۱، ۲، ۳، دانشجوی دکتری، استاد، دانشیار، پردیس کشاورزی و منابع طبیعی دانشگاه تهران  
۴، پژوهشگاه رازلین و مرکز سلطنتی مطالعات دامپزشکی، دانشگاه ادینبرو.  
۵، بخش ژنتیک انسانی مرکز مطالعات پزشکی، پژوهشگاه ژنتیک و درمان‌های مولکولی مرکز مطالعات پزشکی،  
دانشگاه ادینبرو.

(تاریخ دریافت: ۹۰/۹/۲۹ - تاریخ تصویب: ۹۰/۱۲/۱۰)

### چکیده

ژنومی حاوی ۱۰۰۰ چند شکلی تک نوکلئوتیدی دو آللی روی یک کروموزوم به طول ۱۰۰ سانتی مورگان شبیه‌سازی شد و توزیع‌های مختلف واریانس ژن‌های عمده اثر (یکنواخت، نرمال و گاما) و نیز تعداد ژن‌های عمده اثر متنوع (۵، ۱۰ و ۲۰) به صورت فرضیات شبیه‌سازی صفات در نظر گرفته شدند و در نتیجه ۹ صفت متفاوت ایجاد گردید. مقایسه بین ارزش‌های اصلاحی ژنومی برآورد شده به وسیله BayesC و GBLUP نشان داد که ارزش‌های اصلاحی ژنومی برآورد شده و ارزش‌های اصلاحی واقعی در جمعیت مرجع در تمامی صفات، همبستگی بسیار بالایی ( $r > 0.80$ ) با یکدیگر دارند. مقایسه صحت برآوردهای دو روش در صفات مورد مطالعه نشان داد که به جز در صفات با ۵ ژن عمده اثر که روش BayesC برتری معنی‌داری ( $P < 0.05$ ) نشان داد، در سایر صفات عملکرد هر دو روش در برآورد ارزش‌های اصلاحی یکسان بود. همچنین بررسی صحت برآوردهای انجام شده توسط BayesC نشان داد که این روش با کاهش تعداد ژن‌های عمده اثر و نیز توزیع واریانس گاما در صفت مورد مطالعه، عملکرد بهتری نشان می‌دهد. با این حال روش GBLUP در تمامی صفات شبیه‌سازی شده برآورد صحیحی ارائه نمود و تعداد ژن‌ها و توزیع واریانس ژن‌های عمده اثر، اثری بر صحت برآوردهای GBLUP نداشت که این امر می‌تواند به فرضیات هر روش در مورد مدل ژنتیکی صفت مورد مطالعه ارتباط داشته باشد.

واژه‌های کلیدی: BayesC, GBLUP، ارزش اصلاحی، توزیع واریانس ژن‌های عمده اثر

### مقدمه

یکی از اهداف مهم در هر برنامه اصلاح نژادی، برآورد دقیق و صحیح ارزش‌های اصلاحی می‌باشد. در سال‌های اخیر همگام با بهبود در فن‌آوری ژنوتیپ کردن و روش‌های انتخاب ژنومی، امکان بهبود صحت برآورد ارزش‌های اصلاحی افزایش چشمگیری داشته است.

اصطلاح انتخاب ژنومی نخستین بار توسط (1998) Visscher & Haley ارائه شد و پس از آن در سال ۲۰۰۱، روش‌ها و اصول آن توسط Meuwissen et al. (2001) ارائه شد. استفاده از نشانگرهای گسترده شده در سراسر ژنوم (Genome-Wide Markers) باعث افزایش چشمگیری در پیشرفت ژنتیکی گردید. از جمله دلایل

این افزایش می‌توان به کاهش در فاصله نسل (Meuwissen et al., 2001) و کاهش نرخ تبار آمیزی به واسطه انتخاب با استفاده از ارزش‌های اصلاحی به جای اطلاعات خانوادگی (Woolliams et al., 2002; Daetwyler et al., 2007; Dekkers, 2007) اشاره نمود.

برای برآورد ارزش‌های اصلاحی در انتخاب ژنومی، دو دیدگاه ارائه شده است. در نخستین دیدگاه، فرض بر آن است که تمامی چند شکلی‌های تک نوکلئوتیدی (Single-Nucleotide Polymorphism) بر واریانس صفت مؤثرند و صفت دارای مدل ژنتیکی نامحدود می‌باشد. روش‌های بهترین پیش بینی ناریب خطی ژنومی (Genomic Best Linear Unbiased Prediction; GBLUP) از قبیل تابعیت Ridge (Meuwissen et al., 2001) بر پایه این دیدگاه طراحی شده‌اند. در این روش‌ها از ماتریس خویشاوندی نشانگری به جای ماتریس خویشاوندی شجره‌ای استفاده می‌شود (NejatiJavaremi et al., 1997; Villanueva et al., 2009; Hayes et al., 2005) در دیدگاه دوم، فرض بر آن است که تنها برخی از چند شکلی‌های تک نوکلئوتیدی بر صفت اثر دارند و صفت دارای مدل ژنتیکی ژن‌های عمده اثر می‌باشد. به عبارت دیگر فرض بر آن است که تنها تعداد محدودی از چند شکلی‌های تک نوکلئوتیدی بر صفت تأثیر دارند که از بین آنها برخی دارای اثرات عمده و سایر آنها دارای اثرات جزئی می‌باشند. بر پایه این دیدگاه، برخی روش‌های استنباط بییزی (Bayesian) از قبیل BayesC، BayesB و Lasso بنا نهاده شدند (Tibshirani, 1996; Meuwissen et al., 2001). در روش‌های استنباط بییزی (Bayesian) ابتدا پیش فرضی به منظور معرفی توزیع اثرات ژن‌های عمده اثر و نیز مقدار اولیه‌ای جهت تعداد ژن‌های عمده اثر مؤثر بر صفت در نظر گرفته می‌شود (Meuwissen et al., 2001). بر اساس گزارش Goddard (2008) اگرچه در بسیاری از صفات کمی، توزیع واقعی اثرات ژن‌های عمده اثر ناشناخته است، لیکن هنگامی که توزیع اثرات ژن‌های عمده اثر به صورت گاما ( $\alpha=1/66$ ,  $\beta=0/4$ ) در نظر گرفته می‌شود برآوردها صحت بیشتری در مقایسه با توزیع نرمال خواهند داشت. تفاوت اصلی بین دو روش BayesC و GBLUP در دیدگاه‌ها و فرضیات آنها نسبت

به مدل ژنتیکی بیان کننده تغییرات صفت است. در GBLUP، فرض بر آن است که تغییرات صفت تابع مدل ژنتیکی نامحدود است و کلیه چند شکلی‌های تک نوکلئوتیدی بر روی صفت مؤثرند. از سوی دیگر BayesC متکی بر مدل ژن‌های عمده اثر می‌باشد و بر این باور است که تعداد محدودی چند شکلی‌های تک نوکلئوتیدی بر صفت اثر دارند که تنها برخی از آنها بر صفت مورد مطالعه تأثیر عمده‌ای دارند. مطالعه انجام شده توسط Meuwissen et al. (2001) نشان داده که روش GBLUP روشی ساده، سریع و قابل فهم است و به ویژه با استفاده از ماتریس روابط خویشاوندی آلی (NejatiJavaremi et al., 2007) لذا به عنوان روشی کارا در مطالعات نشانگرهای گسترده بر ژنوم (Genome Wide Association Studies) به حساب می‌آید.

هدف اصلی از انجام این تحقیق، بررسی صحت ارزش‌های اصلاحی ژنومی برآورد شده توسط دو روش GBLUP و BayesC در صفات شبیه‌سازی شده با تعداد ۵، ۱۰ و ۲۰ ژن عمده اثر و سه توزیع یکنواخت، نرمال و گاما ( $\alpha=1/66$ ,  $\beta=0/4$ ) برای واریانس ژن‌های عمده اثر بود. همچنین به منظور یافتن اثر تعداد ژن‌های عمده اثر و همچنین نقش توزیع واریانس ژن‌های عمده اثر بر صحت برآوردهای ارزش‌های اصلاحی ژنومی، از شبیه‌سازی استفاده شد.

## مواد و روش‌ها

### شبیه سازی داده‌ها

در این مطالعه، ژنومی شامل یک کروموزوم به طول یک مورگان در نظر گرفته شد. بر روی این ژنوم هزار چند شکلی تک نوکلئوتیدی دو آلی با فاصله یکسان گسترده شد. برای تعداد ژن‌های عمده اثر مؤثر بر صفت سه حالت شامل ۵، ۱۰ و ۲۰ ژن عمده اثر دو آلی در نظر گرفته شد که به صورت یکنواخت بر کروموزوم پراکنده شده‌اند. برای ایجاد جمعیت پایه (G0) ۱۰۰ رأس دام شامل ۵۰ رأس نر و ۵۰ رأس ماده در نظر گرفته شد. فراوانی آلی اولیه برای چند شکلی‌های تک نوکلئوتیدی و ژن‌های عمده اثر در نسل پایه، ۰/۵ لحاظ گردید. برای ایجاد نسل اول (G1)، والدین از بین نرها و ماده‌های نسل G0 به طور تصادفی انتخاب شدند.

گامت‌های والدی بر اساس فرض عدم تعادل پیوستگی و با استفاده از روش تابع مکان یابی (Haldane, Haldane 1919) شبیه‌سازی شدند و سپس گامت‌های نوترکیب ایجاد شده به طور تصادفی انتخاب و آمیخته شدند تا فرد جدیدی از نسل G1 ایجاد شود. روش استفاده شده تا ایجاد نسل ۵۰ام پیگیری شد و در نتیجه جمعیت عدم تعادل پیوستگی (Linkage Disequilibrium population) ایجاد شد. در بررسی نوترکیبی‌های ایجاد شده، مشاهده گردید که این نوترکیبی‌ها دارای توزیع پواسن می‌باشند. در کلیه نسل‌ها، میزان عدم تعادل

پیوستگی با استفاده از آماره  $F_2$  اندازه‌گیری شد. این آماره در واقع میانگین عدم تعادل پیوستگی در بین تمام چند شکلی‌های تک نوکلئوتیدی می‌باشد. پس از ایجاد جمعیت عدم تعادل پیوستگی، ۴ نسل دیگر (نسل‌های ۵۱ تا ۵۴) با اندازه جمعیت ۵۰۰ رأس دام در هر نسل ایجاد و در این نسل‌ها ۲۵۰ رأس نر و ۲۵۰ رأس ماده در نظر گرفته شد. همانطور که در جدول ۱ نشان داده شده است علاوه بر آماره‌های فوق سه حالت تعداد ژن‌های عمده اثر (۵، ۱۰ و ۲۰) در نظر گرفته شد.

جدول ۱- ساختار جمعیت استفاده شده در شبیه‌سازی صفات

ارزش	متغیر
۱	تعداد کروموزوم
۱۰۰۰	تعداد چند شکلی‌های تک نوکلئوتیدی در هر کروموزوم
۱۰۰ سانتی مورگان	طول ژنوم
۰/۰۱	فاصله بین نشانگرها به سانتی مورگان
۵، ۱۰ و ۲۰	تعداد ژن‌های عمده اثر (QTL)
یکنواخت، نرمال و گاما ( $\beta=0/4$ , $\alpha=1/66$ )	توزیع‌های واریانس ژن‌های عمده اثر
تابع مکان یابی Haldane	نوترکیبی
۵۴	تعداد نسل
۵۰	تعداد نسل در جمعیت عدم تعادل پیوستگی
۱۰۰ رأس دام (۵۰ رأس نر و ۵۰ رأس ماده)	اندازه جمعیت عدم تعادل پیوستگی در هر نسل
۴ نسل (۵۱ تا ۵۴)	تعداد نسل جمعیت
۵۰۰ رأس دام (۲۵۰ رأس نر و ۲۵۰ رأس ماده)	اندازه جمعیت
همه افراد نسل ۵۱ (۵۰۰ دام)	جمعیت مرجع
همه افراد نسل‌های ۵۲ تا ۵۴ (۱۵۰۰ دام)	جمعیت تأیید
۰/۳	وراثت پذیری
۱۰	تعداد تکرار

برای شبیه‌سازی واریانس ژن‌های عمده اثر سه توزیع مختلف شامل توزیع یکنواخت، نرمال و گاما برای شبیه‌سازی (جدول ۲). بدین ترتیب ۹ صفت (سناریو) متفاوت شبیه‌سازی گردید و برای افزایش صحت نتایج حاصل از شبیه‌سازی ۱۰ تکرار

برای شبیه‌سازی واریانس ژن‌های عمده اثر سه توزیع مختلف شامل توزیع یکنواخت، نرمال و گاما

برای هر صفت انجام شد و ارزش‌های اصلاحی ژنومی در تمام ۹۰ مورد با دو روش GBLUP و BayesC برآورد شد.

جدول ۲- صفات با توزیع واریانس متفاوت و تعداد ژن‌های عمده اثر

نام صفت	تعداد ژن‌های عمده اثر	توزیع واریانس ژن‌های عمده اثر
NV5	۵	نرمال
GV5	۵	گاما
FV5	۵	یکنواخت
NV10	۱۰	نرمال
GV10	۱۰	گاما
FV10	۱۰	یکنواخت
NV20	۲۰	نرمال
GV20	۲۰	گاما
FV20	۲۰	یکنواخت

### محاسبه ارزش‌های اصلاحی

در این مطالعه همه ۵۰۰ رأس دام نسل ۵۱ به عنوان جمعیت مرجع به کار گرفته شدند و از فنوتیپ دام‌ها در این نسل برای پیش بینی ارزش‌های اصلاحی ژنومی در جمعیت تأیید (نسل‌های ۵۲ تا ۵۴) استفاده گردید. برای پیش بینی ارزش‌های اصلاحی ژنومی (GEBV) از مدل زیر استفاده گردید:

$$GEBV = X \cdot \hat{g}$$

که در آن  $X$  ماتریس طرح می‌باشد که به ژنوتیپ دام‌ها ارتباط دارد و  $\hat{g}$  بردار اثرات برآورد شده برای چند شکلی‌های تک نوکلئوتیدی می‌باشد.

### روش GBLUP

روش GBLUP به کار رفته در این مطالعه بر اساس یک مدل مختلط ساده طراحی شد که فرض گردید که در آن کلیه چند شکلی‌های تک نوکلئوتیدی بر صفت مورد مطالعه مؤثر می‌باشند. برای به کار بردن این روش از نرم افزار ASReml (Gilmour et al., 1995) و مدل زیر استفاده شد:

$$y = \mu 1_n + X_i g_i + e$$

که در آن  $y$  بردار ارزش‌های فنوتیپی،  $\mu$  میانگین فنوتیپی،  $n$  تعداد رکوردها،  $1_n$  بردار یک‌ها به ابعاد  $1 \times n$ ،  $g_i$  ارائه کننده اثر ژنتیک افزایشی  $i$  امین چند شکلی تک نوکلئوتیدی،  $X_i$  ماتریس طرح برای  $i$  امین چند شکلی تک نوکلئوتیدی و  $e$  انحراف خطا با توزیع نرمال می‌باشد.

اثرات ژنتیک افزایشی چند شکلی‌های تک نوکلئوتیدی ( $g$ ) دارای توزیع نرمال  $N(0, \sigma_g)$  فرض شد که  $g$  ماتریس روابط خویشاوندی برای همه چند شکلی‌های تک نوکلئوتیدی می‌باشد. باید توجه داشت که بر اساس احتمالات identical-by-state بین هر دو دام محاسبه شده است و برای تمام دام‌های دارای ژنوتیپ (دام‌های جمعیت مرجع و جمعیت تأیید) تعمیم یافته است.

برای محاسبه مجموع روابط آلی بین هر دو دام از روش Nejatijavaremi et al. (1997) استفاده شد. برای بدست آوردن اثرات افزایشی چند شکلی‌های تک نوکلئوتیدی از مدل مختلط زیر استفاده شد:

$$\begin{bmatrix} 1_n' 1_n & 1_n' X \\ X' 1_n & X' X + I\alpha \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 1_n' y \\ X' y \end{bmatrix}$$

که در آن  $\alpha = \frac{\sigma_e^2}{\sigma_g^2}$  و  $I$  ماتریس همانی می‌باشد.

### روش BayesC

در روش BayesC فرض این است که تنها بخشی از چند شکلی‌های تک نوکلئوتیدی ( $\pi$ ) بر صفت اثر دارند. برای به کار بردن این روش از مدل زیر استفاده گردید:

$$y = \mu 1_n + \sum_{i=1}^m X_i g_i + e$$

### نتایج

میانگین عدم تعادل پیوستگی ( $r^2$ ) افراد نسل آخر جمعیت عدم تعادل پیوستگی (نسل ۵۰ ام)  $\pm 0/015$   $0/184$  بدست آمد.

میانگین صحت ارزش‌های اصلاحی ژنومی حاصل از دو روش BayesC و GBLUP در ۹ صفت مورد مطالعه در جدول ۳ ارائه شده است. در صفات دارای ۵ ژن عمده اثر، ارزش‌های اصلاحی ژنومی برآورد شده توسط BayesC صحت بالاتری ( $P < 0/05$ ) را نسبت GBLUP نشان دادند. به علاوه در صفات با ۱۰ ژن عمده اثر، تنها در صفت دارای ۱۰ ژن عمده اثر با توزیع واریانس گاما (GV10) صحت بیشتری در BayesC در مقایسه با GBLUP مشاهده شد و در دو صفت دیگر تفاوت معنی داری مشاهده نشد. در صفات با ۲۰ ژن عمده اثر نیز تفاوت معنی داری بین دو روش مشاهده نشد.

بالاترین صحت برآوردهای انجام شده به وسیله BayesC در صفت دارای ۵ ژن عمده اثر با توزیع واریانس گاما (GV5) بدست آمد. از سوی دیگر بالاترین صحت بدست آمده توسط GBLUP در صفت دارای ۲۰ ژن عمده اثر با واریانس یکسان (FV20) مشاهده گردید. اثر تعداد ژن‌های عمده اثر بر برآوردها در هر روش با استفاده از آزمون t مورد مطالعه قرار گرفت. نتایج حاصل از مقایسات نشان داد که در صفات دارای توزیع مشابه و تعداد ژن‌های عمده اثر متفاوت، بین برآوردهای GBLUP هیچ گونه تفاوت معنی داری وجود ندارد. همین مقایسات در برآوردهای BayesC نشان داد که صفات دارای ۵ ژن عمده اثر برآوردهای بالاتری نسبت سایر صفات دارند. در مقایسه بین صفات دارای ۱۰ و ۲۰ ژن عمده اثر نیز مشخص شد که تنها در صفات دارای توزیع واریانس گاما، تفاوت معنی داری در این خصوص وجود داشت و صفت دارای ۱۰ ژن عمده اثر (GV10) نسبت به صفت دارای ۲۰ ژن عمده اثر (GV20) برآوردهای صحیح‌تری ( $P < 0/01$ ) نشان داد. اگر چه تفاوت معنی داری بین میانگین مربعات پیش بینی برآورد ارزش‌های اصلاحی ژنومی در BayesC و GBLUP در هیچ کدام از صفات مشاهده نشد. اما با افزایش تعداد ژن‌های عمده اثر مؤثر بر صفت، میانگین

که در آن  $\gamma$  بردار ارزش‌های فنوتیپی،  $\mu$  میانگین فنوتیپی،  $n$  تعداد رکوردها،  $1_n$  بردار یکا به ابعاد  $1 \times n$ ،  $X_i$  بردار ژنوتیپ‌ها (۱، ۲ و ۳ به ترتیب برای ژنوتیپ‌های 00، 10/01 و 11) برای  $i$  امین چند شکلی تک نوکلئوتیدی،  $g_i$  اثر جایگزینی آلی  $i$  امین چند شکلی تک نوکلئوتیدی و  $e$  بردار اثرات باقیمانده است که دارای توزیع نرمال  $N(0, \sigma_e^2)$  می‌باشد.

در این مدل فرض شده است که چند شکلی‌های تک نوکلئوتیدی مؤثر بر صفت ( $\pi$ ) دارای توزیع نرمال  $N(0, \sigma_{snps}^2)$  هستند و سایر چند شکلی‌های تک نوکلئوتیدی تأثیری بر صفت ندارند. در اجرای روش BayesC از روش نمونه‌برداری Gibbs استفاده شد. برای برآورد ارزش‌های اصلاحی و همچنین آماره‌های  $\pi$ ،  $\sigma_{snps}^2$  و  $\sigma_e^2$  از مقادیر پیش فرض اولیه استفاده گردید. برای هر تحلیل یک زنجیره Markov chain Monte Carlo به طول ۲۱۰ هزار چرخه شامل ۱۰ هزار چرخه اولیه جهت گرم کردن و ۲۰۰ هزار چرخه اصلی به کار رفت که از هر ۵ دور یک بار نتایج ذخیره شد و در پایان نتایج ۴۰ هزار چرخه ذخیره گردید.

### مقایسه روش‌ها در برآورد ارزش‌های اصلاحی

برای مقایسه دو روش از دو معیار استفاده شد: معیار اول مقایسه، صحت برآوردهای ارزش‌های اصلاحی ژنومی حاصل از هر یک از دو روش می‌باشد و در معیار دوم مقایسه، میانگین مربعات خطای پیش بینی (Mean Squared Prediction Error) ارزش‌های اصلاحی برآورد شده از ارزش‌های اصلاحی واقعی در هر یک از دو روش، به کار گرفته شد.

برای اندازه‌گیری صحت برآوردهای ارزش‌های اصلاحی ژنومی از همبستگی بین ارزش‌های اصلاحی ژنومی واقعی و برآورد شده استفاده شد. برای محاسبه این دو آماره، برآورد ارزش‌های اصلاحی ژنومی در جمعیت مرجع مورد استفاده قرار گرفت (Coster et al., 2010).

همچنین این آماره‌ها در نسل‌های تأیید نیز محاسبه گردید. برای مقایسه بین برآوردهای GBLUP و BayesC نیز از آزمون t استفاده شد.

مربعات پیش بینی برآورد ارزش‌های اصلاحی ژنومی در روش GBLUP افزایش یافت (جدول ۳).

جدول ۳- میانگین (اشتباه معیار) صحت ارزش‌های اصلاحی ژنومی برآورد شده توسط BayesC و GBLUP

GBLUP	BayesC	نسل	صفت	GBLUP	BayesC	نسل	صفت	GBLUP	BayesC	نسل	صفت
۰/۸۲(۰/۰۱)	۰/۹۱(۰/۰۱)	۵۱	GV5	۰/۸۱(۰/۰۱)	۰/۸۷(۰/۰۱)	۵۱	GV10	۰/۸۱(۰/۰۱)	۰/۸۲(۰/۰۱)	۵۱	GV20
۰/۷۳(۰/۰۲)	۰/۸۹(۰/۰۲)	۵۴		۰/۷۰(۰/۰۲)	۰/۸۱(۰/۰۲)	۵۴		۰/۷۱(۰/۰۲)	۰/۷۳(۰/۰۲)	۵۴	
۰/۸۲(۰/۰۲)	۰/۸۷(۰/۰۲)	۵۱	NV5	۰/۸۱(۰/۰۲)	۰/۸۳(۰/۰۲)	۵۱	NV10	۰/۸۲(۰/۰۱)	۰/۸۴(۰/۰۱)	۵۱	NV20
۰/۷۳(۰/۰۳)	۰/۸۲(۰/۰۲)	۵۴		۰/۶۸(۰/۰۳)	۰/۷۳(۰/۰۳)	۵۴		۰/۷۳(۰/۰۲)	۰/۷۵(۰/۰۲)	۵۴	
۰/۸۲(۰/۰۲)	۰/۸۵(۰/۰۱)	۵۱	FV5	۰/۸۰(۰/۰۲)	۰/۸۲(۰/۰۲)	۵۱	FV10	۰/۸۲(۰/۰۱)	۰/۸۳(۰/۰۱)	۵۱	FV20
۰/۷۴(۰/۰۲)	۰/۷۹(۰/۰۲)	۵۴		۰/۶۸(۰/۰۲)	۰/۷۰(۰/۰۲)	۵۴		۰/۷۳(۰/۰۲)	۰/۷۳(۰/۰۲)	۵۴	

۱۰ ژن عمده اثر، صفات با توزیع گاما برآوردهای خیلی صحیح‌تر ( $P < 0.01$ ) نسبت به توزیع یکنواخت و برآوردهای صحیح‌تری ( $P < 0.05$ ) نسبت به توزیع نرمال ایجاد کرد.

برآوردها کاهش می‌یابد، هماهنگ است. این مطالعه نشان داد که صحت برآوردهای BayesC با افزایش تعداد ژن‌های عمده اثر مؤثر بر صفت، کاهش چشمگیری نشان می‌دهد تا جایی که در کلیه صفات با ۲۰ ژن عمده اثر برتری خود را نسبت به GBLUP از دست می‌دهد. همچنین برآوردهای BayesC نشان داد که صفاتی با ۵ ژن عمده اثر و صفات با ۱۰ ژن عمده اثر با توزیع واریانس گاما (GV10) صحت‌های بالاتری نسبت به سایر صفات نشان می‌دهند.

این نتایج نشان داد که حساسیت BayesC نسبت به تعداد ژن‌های عمده اثر، بیشتر از حساسیت آن به تغییر نوع توزیع واریانس ژن‌های عمده اثر می‌باشد. همچنین نتایج حاصل از این تحقیق نشان می‌دهد که توزیع اثرات گاما و تعداد کم ژن‌های عمده اثر باعث افزایش صحت برآوردها می‌شوند که برای نمونه می‌توان از صفت ۱۰ ژن عمده اثر با توزیع گاما (GV10) نام برد. در اکثر صفات شبیه سازی شده، مشاهده گردید که GBLUP برآوردهایی مشابه با BayesC ارائه می‌کنند.

اثر توزیع واریانس بر صحت برآوردهای انجام شده توسط GBLUP نشان داد که هیچ تفاوت معنی داری در بین صفات با توزیع‌های متفاوت وجود نداشت. نتایج حاصل از روش BayesC نشان داد که در صفات با ۵ و

### بحث

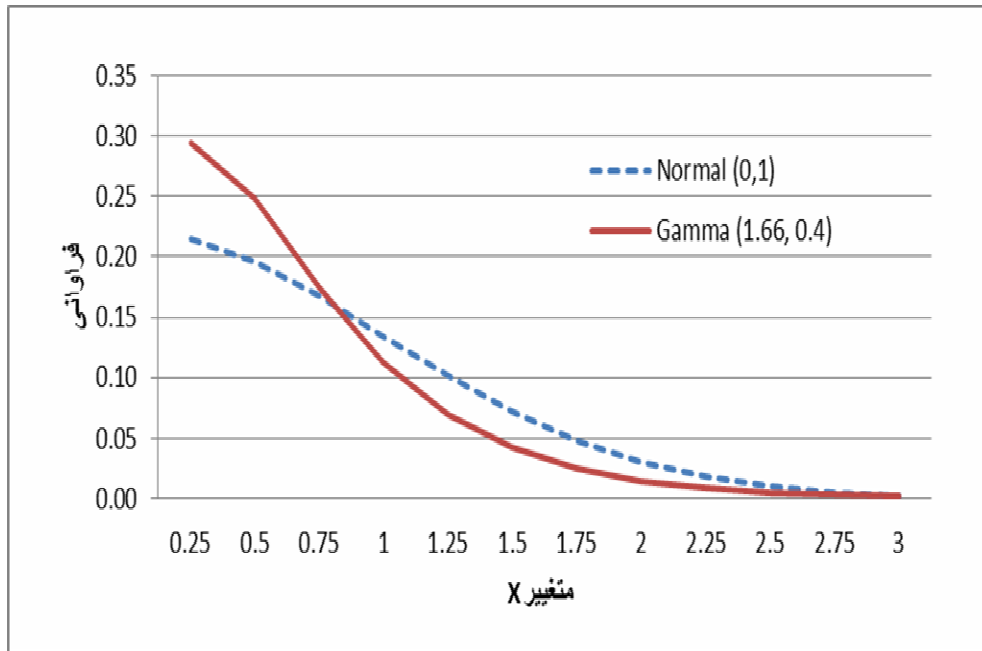
در تحقیق حاضر، میانگین  $r^2$  در نسل آخر جمعیت-های عدم تعادل پیوستگی ( $G_{50}$ )  $0.12 \pm 0.185$  محاسبه شد. این در حالی است که مقدار مورد انتظار برای  $r^2$  توسط فرمول (Sved 1971)، ۰/۲۱۱ برآورد گردید.

مقایسه صحت برآوردهای GBLUP و BayesC در صفات مختلف از نظر ساختار ژنتیکی و تعداد ژن‌های عمده اثر انجام گردید. نتایج حاصل نشان داد که روش GBLUP همواره و در همه صفات و تحت هر شرایطی، برآوردهای صحیحی ارائه می‌کند. در BayesC مشاهده شد که بالاترین صحت در صفاتی با تعداد کم ژن‌های عمده اثر دیده می‌شود. این مطالعه نشان داد که BayesC در صفاتی با تعداد کم ژن‌های عمده اثر نسبت به GBLUP صحت برآوردهای بالاتری دارد.

این نتایج با یافته‌های Daetwyler et al. (2010) که گزارش کرده‌اند با افزایش تعداد ژن‌های عمده اثر، صحت

برآوردها در روش BayesC گردید.

در این مطالعه، توزیع گاما در اثرات و واریانس ژن های عمده اثر، باعث افزایش صحت



شکل ۱- تابع چگالی احتمال توزیع های گاما و نرمال

Wong (2010) و Daetwyler et al. (2010) مطابقت دارد.

### نتیجه گیری کلی

ارزش های اصلاحی ژنومی برآورد شده توسط روش های BayesC و GBLUP در تمامی صفات شبیه سازی شده همبستگی بسیار بالایی با ارزش های اصلاحی واقعی نشان دادند. برآورد ارزش های اصلاحی ژنومی توسط دو روش مذکور نشان داد که هر دو روش عملکرد مشابهی دارند و تنها در صفت دارای ۵ ژن عمده اثر با توزیع واریانس گاما، BayesC برآوردهای صحیح تری ارائه نمود. از سوی دیگر در بررسی صحت برآوردهای انجام شده توسط BayesC مشاهده شد که کاهش تعداد ژن های عمده اثر و نیز توزیع واریانس گاما در صفت باعث بهبود صحت برآوردهای BayesC می گردد. این در حالی است که در تمام صفات شبیه سازی شده، GBLUP برآوردهای همواره صحیح و قابل اعتمادی ارائه نمود و تعداد ژن ها و توزیع واریانس ژن

دلیل این امر دو عامل می تواند باشد: نخست آنکه توزیع پیش فرض BayesC با توزیع گامای اثرات و واریانس ژن های عمده اثر هماهنگی دارند که باعث برآورد صحیح تر اثرات چند شکلی های تک نوکلئوتیدی می گردد. همانطور که Goddard (2008) گزارش کرده است توزیع پیش فرض گاما در Bayes برآوردهای صحیح تری ایجاد می کند که با نتایج این مطالعه همخوانی دارد. دلیل دوم این است که توزیع گاما دارای تعداد بیشتری ژن های عمده اثر با اثرات بزرگتر نسبت به توزیع نرمال است (شکل ۱). در نتیجه برآوردهای ارزش های اصلاحی ژنومی در صفاتی با تعدادی ژن های عمده اثر با اثرات بزرگتر، صحیح تر است. ارزش های اصلاحی ژنومی واقعی و ارزش های اصلاحی ژنومی برآورد شده به وسیله BayesC در همه صفات همبستگی بسیار بالایی ( $r > 0.80$ ) را نشان می دادند که با گزارش های Solberg et al. (2008)، Nadaf & Pong-

**سپاسگزاری**

نویسندگان از همکاری و حمایت‌های بی دریغ دکتر Pau Navarro سپاس فراوان دارند .

های عمده اثر تأثیری بر صحت برآوردهای GBLUP نداشت. لذا پیشنهاد می‌گردد در تحلیل داده‌های ژنومی با نشانگرهای چگال (Dense Markers)، استفاده از هرکدام از دو روش مناسب و صحیح می‌باشد.

**REFERENCES**

1. Coster, A., Bastiaansen, J.W.M., Calus, M.P.L., van Arendonk, J.A.M., Bovenhuis, H., 2010. Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. *Genetics Selection Evolution* 42, 9.
2. Daetwyler, H.D., Pong-Wong, R., Villanueva, B., Woolliams, J.A., 2010. The Impact of genetic architecture on genome-wide evaluation methods. *Genetics* 185, 1021–1031.
3. Daetwyler, H.D., Villanueva, B., Bijma, P., Woolliams, J.A., 2007. Inbreeding in genome-wide selection. *Journal of Animal Breeding and Genetics* 124, 369–376.
4. Dekkers, J.C.M., 2007. Prediction of response from marker-assisted and genomic selection using selection index theory. *Journal of Animal Breeding and Genetics* 124, 331–341.
5. Gilmour, A.R., Thompson, R., Cullis, B.R., 1995. Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 51, 1440–1450.
6. Goddard, M., 2008. Genomic selection prediction of accuracy and maximisation of long term response. *Genetica* 136, 245–257.
7. Haldane, J.B.S., 1919. The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics* 8, 299–309.
8. Hayes, B.J., Visscher, P.M., Goddard, M.E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research* 91, 47–60.
9. Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423–447.
10. Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 321–322.
11. Nadaf, J., Pong-Wong, R., 2011. Applying different genomic evaluation approaches on QTLMAS2010 dataset. *BMC Proceedings* 5(Suppl 3), S9.
12. NejatiJavaremi, A., Smith, C., Gibson, P.J., 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. *Journal of Animal Science* 75, 1738–1745.
13. Solberg, T.R., Sonesson, A.K., Woolliams, J.A., Meuwissen, T.H.E., 2008. Genomic selection using different marker types and densities. *Journal of Animal Science* 86, 2447–2454.
14. Sved, J.A., 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theoretical Population Biology* 2, 125–141.
15. Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society Series B: Methodological* 58, 267–288.
16. Villanueva, B., Pong-Wong, R., Fernandez, J., Toro, M.A., 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. *Journal of Animal Science* 83, 1747–1752.
17. Visscher, P.M., Haley, C.S., 1998. Strategies for marker-assisted selection in pig breeding programmes. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production WCGALP, Armidale, Australia*, pp. 503–510.
18. Woolliams, J.A., Pong-Wong, R., Villanueva, B., 2002. Strategic optimisation of short- and long-term gain and inbreeding in MAS and non-MAS schemes. Communication No. 32-02 on CD-ROM. Proc. The 7th World Congress of Genetics Applied to Livestock Production, Montpellier, France.