دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشهای تولیدات دامی

# مقایسه روش‌های چند مرحله‌ای و تک مرحله‌ای بیزی برای برآورد ارزش‌های اصلاحی ژنومی در حیوانات ژنوتیپ شده و نشده- مطالعه شبیه‌سازی

مصطفی مدد جیرنده[۱]، جلیل شجاع[۲]، صادق علیجانی[۳]، سید عباس رافت[۲] و جک سی. ام دکرز[۴]

۱- دانشجوی دکتری، گروه علوم دامی، دانشگاه تبریز، تبریز، ایران
۲- استاد، گروه علوم دامی، دانشگاه تبریز، تبریز، ایران
۳- دانشیار، گروه علوم دامی، دانشگاه تبریز، تبریز، ایران (نویسنده مسوول: sad-ali@tabrizu.ac.ir)
۴- استاد، گروه علوم دامی، دانشگاه ایالتی آیووا، ایمز، ایالات متحده آمریکا

تاریخ دریافت: ۹۸/۴/۱۱ تاریخ پذیرش: ۹۸/۶/۲۰
صفحه: ۱۲۲ تا ۱۳۱

**چکیده**

**هدف از این مطالعه، مقایسه صحت ارزیابی ژنومی روش‌های چندمرحله‌ای BayesA، BayesB، BayesC، BayesL و روش‌های تک‌مرحله‌ای بیزی SSBR-C و SSBR-A در مقادیر متفاوت π برای برآورد ارزش‌های اصلاحی ژنومی حیوانات تعیین ژنوتیپ شده و نشده بود. ژنومی حاوی ۴۰۰۰۰ نشانگر تک‌نوکلئوتیدی دوآللی پراکنده شده روی ۲۰ کروموزوم هرکدام به طول ۱۰۰ سانتی‌مورگان شبیه‌سازی شد. مقادیر بهینه π در روش BayesC به‌ترتیب ۰/۹۸۰ و ۰/۹۹۵ در دو توزیع نرمال و گامای اثرات ژنی برآورد شد و در روش SSBR-C نیز مورد استفاده قرار گرفت. صحت پیش‌بینی ژنومی در روش SSBR–C (π=0.995) نسبت به سایر روش‌ها از ۰/۰۲ تا ۰/۰۹ در توزیع گامای اثرات ژنی بیش‌تر برآورد شد. بنابراین، روش SSBR–C (π=0.995) با در نظر گرفتن توزیع مزدوج و استفاده همزمان از همه اطلاعات شجره‌ای، فنوتیپی و ژنومی توانست در حالت توزیع گامای اثرات ژنی عملکرد بهتری را از خود نشان داده و انتخاب مناسب‌تری به‌شمار رود. کلیه روش‌های تک‌مرحله‌ای و چندمرحله‌ای بیزی عملکرد تقریبا مشابهی را در حالت توزیع نرمال اثرات ژنی از خود نشان دادند. فلذا، در حالت توزیع نرمال اثرات ژنی توصیه می‌شود تا از روش SSBR–C (π=0) با توجه به ضریب تابعیت پیش‌بینی ژنومی نزدیک به یک استفاده شود. همچنین، افت صحت پیش‌بینی ژنومی با افزایش فاصله نسلی بین جمعیت مرجع و تایید برای افراد ژنوتیپ شده در مقایسه با افراد ژنوتیپ نشده از حساسیت کمتری برخوردار بود.**

**واژه‌های کلیدی: انتخاب ژنومی، بیزی، روش چندمرحله‌ای، شبیه‌سازی، صحت ژنومی**

## مقدمه

ارزیابی ژنومی با استفاده از روش‌های چند مرحله‌ای بر مبنای داده‌های فنوتیپی و ژنوتیپی صورت گرفته و اثرات نشانگرها برآورد می‌شود. در مدل آنالیز برای مرحله دوم، تنها اثرات درنظرگرفته شده، اثرات نشانگرها خواهد بود و پیش‌بینی ارزش‌های اصلاحی ژنومی[۱] برای دام‌های تعیین ژنوتیپ شده جوان با استفاده از این اثرات صورت خواهد گرفت (۳۳). به‌دلیل معایب روش‌های چندمرحله‌ای[۲] در توسعه مدل‌ها، روش تک‌مرحله‌ای[۳] پیشنهاد شد که در این روش به‌طور همزمان اطلاعات فنوتیپی، ژنومی و شجره‌ای در مدل قرار داده می‌شوند. در این روش برای حیوانات تعیین ژنوتیپ شده و نشده به‌طور همزمان ارزش اصلاحی برآورد می‌شود (۳،۱۱،۲۲).

روش‌های SSBR[۴] با تلفیق ماتریس روابط خویشاوندی ژنومی و شجره‌ای ماتریس H را به‌وجود می‌آورند. در این روش از زنجیره مارکوف مونت کارلو[۵] برای برآورد مولفه‌های واریانس استفاده شده تا بتوان از توزیع‌های اثرات ژنی متنوع‌تری بهره‌مند شد (۱۱). بنابر نتایج پژوهش انجام شده، صحت پیش‌بینی تمام روش‌های چندمرحله‌ای بیزی در توزیع نرمال اثرات ژنی به هم نزدیک بود. همچنین زمانی که توزیع اثرات ژنی به‌صورت گاما در نظر گرفته شد، صحت برآوردها با استفاده از روش‌های بیزی بالاتر از روش‌های غیر بیزی گزارش شد (۲).

توزیع نرمال یکی از مهم‌ترین توزیع‌های احتمالی پیوسته است. تابع چگالی این توزیع دارای دو پارامتر است که یکی تعیین‌کننده میانگین (μ) و دیگری تعیین‌کننده واریانس (σ) توزیع هستند. توزیع گاما نیز از توزیع‌های احتمالی پیوسته به‌شمار می‌رود و دارای دو پارامتر مقیاس β و شکل α است. به‌دلیل پارامترهای شکل و مقیاس، این توزیع انعطاف‌پذیری زیادی برای تحلیل هر مجموعه داده حقیقی مثبت را داراست (۹،۲۴).

جیانولا و همکاران (۱۲) با مقایسه روش‌های BayesA و BayesB نشان دادند که می‌توان مقادیری از پیش تعیین شده را برای اثرات نشانگرهای تک‌نوکلئوتیدی در نظر گرفت. این مقدار از پیش تعیین شده می‌تواند با احتمال π اثرات نشانگری را صفر درنظر گرفته و واریانس نشانگرها با احتمال 1-π توجیه شود. در روش BayesA مقدار π با صفر برابری می‌کند؛ به‌طوری‌که در این روش همه اثرات نشانگری در نظر گرفته می‌شود، درحالی‌که در روش‌های BayesB و BayesC مقدار π بزرگتر از صفر بوده و به‌عنوان یک عامل ناشناخته می‌تواند در ارزیابی روش‌های بیزی اعم از چندمرحله‌ای و تک‌مرحله‌ای حائز اهمیت باشد (۱۶). برخی از محققین به مروری کلی بر استفاده از روش‌های بیزی در انتخاب ژنومی پرداختند؛ اما روش‌های تک‌مرحله‌ای مبتنی بر بیزی (SSBR) را مورد مطالعه قرار نداند (۱۰،۳۰). از آنجایی‌که توالی‌یابی کل ژنوم روی تعداد بسیار زیادی ازافراد، سخت و هزینه‌بر است و از سوی دیگر نیازمند در اختیار داشتن ارزش ژنتیکی افراد است، انجام مطالعات با استفاده از شبیه‌سازی، روشی منطقی برای چنین امری خواهد بود (۳۲).

1- Genomic Estimated Breeding Value 2- Multiple Step 3- Single Step 4- Single-Step Bayesian Regression
5- Markov chain Monte-Carlo

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]
[ DOI: 10.29252/rap.10.26.122 ]

پژوهش حاضر با هدف بررسی صحت ارزیابی ژنومی روش‌های چند مرحله‌ای بیزی BayesA، BayesB، BayesC، BayesL و روش‌های تک‌مرحله‌ای بیزی SSBR-C و SSBR-A در مقادیر متفاوت $\pi$ برای برآورد ارزش‌های اصلاحی ژنومی حیوانات ژنوتیپ شده و نشده در دو توزیع متفاوت اثرات ژنی نرمال و گاما طی سه نسل متوالی انجام شد.

## مواد و روش‌ها

### شبیه‌سازی ژنوم

جمعیت‌ها با استفاده از نرم‌افزار QMSim شبیه‌سازی شدند (۳۱). ابتدا ژنومی حاوی ۴۰۰۰۰ نشانگر تک‌نوکلئوتیدی دوآللی پراکنده شده روی ۲۰ کروموزوم هر کدام به طول ۱۰۰ سانتی‌مورگان شبیه‌سازی شد و توزیع‌های متفاوت اثرات ژنی (اعم از نرمال و گاما) به صورت فرضیات شبیه‌سازی صفت در نظر گرفته شدند و در نتیجه دو صفت با معماری ژنتیکی متفاوت ایجاد شد. در مرحله اول شبیه‌سازی، یک جمعیت پایه به اندازه موثر ۱۰۰۰ حیوان غیرخویشاوند (۵۰۰ نر و ۵۰۰ ماده) در نظر گرفته شد. این جمعیت در ۱۰۰۰ نسل اول آمیزش تصادفی داشت و اندازه آن ثابت باقی ماند، و در ۱۰۲۰ نسل بعدی (نسل های ۱۰۰۱ تا ۲۰۲۰) برای ایجاد عدم تعادل پیوستگی اندازه جمعیت به‌صورت تدریجی از ۱۰۰۰ به ۴۰۰ کاهش یافت. سپس ۱۵۰ نر و ۱۵۰ ماده از آخرین نسل از جمعیت پایه گزینش شدند. در این گام، جمعیت با استفاده از آمیزش تصادفی برای ۱۰ نسل دیگر شبیه‌سازی شد. در گام دوم شبیه‌سازی، ۴۰ نر و ۱۰۰۰ ماده به‌طور تصادفی آمیزش داده شدند و ۱۰ نسل دیگر شبیه‌سازی شد. برای تشکیل جمعیت مرجع، ۴۰۰۰ حیوان از نسل‌های ششم و هفتم انتخاب شدند، به‌طوریکه ۲۰۸۰ فرد انتخاب شده از پدرها و مادرهای نسل ششم و هفتم بودند و ۱۹۲۰ حیوان نیز به‌طور تصادفی از نسل‌های ششم و هفتم انتخاب شدند. لازم به ذکر است که جمعیت مرجع دارای هر دو رکورد فنوتیپی و ژنوتیپی بودند و افراد نسل‌های هشتم، نهم و دهم نیز به‌عنوان جمعیت کاندیدا در نظر گرفته شدند به‌طوریکه ۵۰ درصد از افراد در هر نسل از جمعیت تایید به طور تصادفی به‌عنوان حیوانات دارای ژنوتیپ و افراد باقی‌مانده به‌عنوان حیوانات ژنوتیپ نشده انتخاب شدند. خلاصه جمعیت شبیه‌سازی شده و پارامترهای به‌کار رفته در جدول ۱ نشان داده شده است.

جدول ۱- پارامترهای فرایند شبیه‌سازی

Table 1. Parameters of the simulation process

| ساختار جمعیت | جمعیت اول/دوم |
|---|---|
| جمعیت اولیه | |
| فاز اول تعداد نسل (تعداد افراد) | ۱۰۰۰(۱۰۰۰) |
| فاز دوم تعداد نسل (تعداد افراد) | ۲۰۲۰(۴۰۰) |
| جمعیت بسط داده شده | |
| تعداد نرهای از جمعیت پایه | ۱۵۰ |
| تعداد ماده‌های از جمعیت پایه | ۱۵۰ |
| تعداد نسل | ۱۰ |
| جمعیت نهایی | |
| تعداد نرهای از جمعیت بسط داده شده | ۴۰ |
| تعداد ماده‌های از جمعیت بسط داده شده | ۱۰۰۰ |
| تعداد نسل | ۱۰ |
| احتمال نر بودن نتاج | ۰/۵۰ |
| گزینش و طرح آمیزش | تصادفی |
| معیار حذف | ارزش اصلاحی برآوردی |
| ژنوم | |
| تعداد کروموزوم | ۲۰ |
| طول هرکروموزوم (سانتی مورگان) | ۱۰۰ |
| تعداد نشانگرهای تک‌نوکلئوتیدی | ۴۰۰۰۰ |
| تعدادQTL ها | ۲۰۰ |
| اثر الل هایQTL | نرمال/ گاما (با پارامترشکل=۰/۴) |
| نرخ جهش در نشانگر و QTL ها | $۲/۵ \times ۱۰^{-۵}$ |
| وراثت‌پذیری | ۰/۳۵ |

### روش‌های چندمرحله‌ای بیزی

مدل آماری در تمامی روش‌های چندمرحله‌ای مورد مطالعه، به صورت ذیل بود:

$$y_i = \mu + \sum_{i=1}^{m} X_{ij}\, b_j + e_i \qquad (۱)$$

در معادله (۱)، y بردار مشاهدات شامل افراد تعیین ژنوتیپ شده، $\mu$ عرض از مبدا، $X_{ij}$ ژنوتیپ های نشانگری، b اثر نشانگری و e اثر باقیمانده هستند. توزیع چگالی احتمال توام پیشین مجهولات رابطه فوق به صورت زیر است:

$$P(\mu,b,\sigma^2|df,s,\omega) \sim \left\{\Pi_{j=1}^{m} P\left(b_j \middle| \begin{matrix}\theta_{bj}\\ \sigma^2\end{matrix}\right) P(\theta_{bj}|\omega)\right\} X^{-2}(\sigma^2|df,s) \qquad (۲)$$

در رابطه (۲)، توزیع پیشین یکنواخت برای میانگین، $\chi^{-2}(\sigma^2|df,s)$ توزیع کای اسکور معکوس مقیاس‌دار برای

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

واریانس باقی مانده با درجه آزادی df و پارامتر مقیاس برابر با S در نظر گرفته شد. $P(b_j|\Theta_{bj},\sigma^2)$ توزیع پیشین ژامین نشانگر را نشان می‌دهد که در آن $\Theta_{bj}$ یک برداری از پارامترها است که نوع توزیع پیشین در نظر گرفته برای اثرات نشانگری را مشخص می‌کند. $P(\Theta_{bj}|\omega)$ توزیع پیشین در نظر گرفته شده برای $\Theta_{bj}$ و $\omega$ پارامتری است که این نوع توزیع را مشخص می‌کند.

**BayesA**

در روش BayesA (۲۳) واریانس اثرات نشانگری متغیر در نظر گرفته شده، و هر نشانگر واریانس خودش را داراست. برای برآورد همزمان اثرات نشانگری توزیع پیشین، اثرات نشانگری به شرط واریانس متغیر نرمال با میانگین صفر فرض شد که واریانس‌ها دارای توزیع کای مربع معکوس مقیاس‌دار هستند و در نتیجه توزیع پیشین غیرشرطی را تبدیل به یک توزیع t می‌کند (۱۲) و چون در این توزیع تمامی اثرات نشانگری در نظر گرفته می‌شوند صحت بر آوردها افزایش می‌یابد. توزیع پسین پارامترهای اثرات نشانگری و واریانس هر کدام با نمونه‌گیری گیبس بدست می‌آید.

**BayesB**

در روش BayesB (۲۳) اثرات نشانگری با احتمال $\pi$ برابر با صفر و با احتمال $(1-\pi)$ از توزیع t به روشی که در BayesA شرح داده شده به دست می آیند. به این صورت که:

(۳)

$$b_j|df_b,S_b,\pi \sim IID\begin{cases}0 \text{ with probablity } \pi\\ t(b_i|df_b,S_b) \text{ with probablity } 1\end{cases}$$

در رابطه (۳)، با صفر گرفتن اثر تعداد زیادی از نشانگرها صحت برآورد بالا رفته و چون از تمامی نشانگرها در مدل استفاده نمی‌شود به علت آزمون‌ها چندگانه همزمان مسئله بیش برآورد نیز حل می‌شود و اثرات سایر نشانگرها را بر اساس توزیع پیشین منقبض می‌کند.

**BayesC**

BayesC شکلی از BayesB است (۱۶) و در برخی از مقالات به عنوان بیز $C\pi$ شناخته می‌شود. فقط بجای اینکه از توزیع پیشین t برای اثرات نشانگری استفاده کند از توزیع نرمال استفاده می‌کند و لذا توزیع پسین هم مزدوج و نرمال می‌باشد. BayesC برای تصحیح بعضی از نقض‌های BayesB پیشنهاد شده است و برعکس روش‌های فراوانی‌گرا که واریانس معین فرض می‌شود، این واریانس برآورد می‌گردد.

**BayesL**

در روش BayesL که بیز LASSO[1] هم نامیده می‌شود (۲۸)، تابع چگالی احتمال پیشین نمایی دوگانه (Double-Exponential) به اثرات نشانگری اختصاص داده شده و شکل آن به صورت زیر است:

(۴)

$$DE(b_j|\lambda,\sigma_e^2) = \int N(b_j|0,\sigma_e^2\tau_j^2)EXP(\tau_j^2|\lambda^2)\tau_j^2$$

در رابطه (۴)، $\lambda$ ناشناخته فرض شده و به آن توزیع گاما با پارامتر شکل $\lambda$ و مقیاس $\tau$ اختصاص داده می‌شود.

## روش‌های تک‌مرحله‌ای بیزی در انتخاب ژنومی

مدل روش‌های تک‌مرحله‌ای بیزی به‌صورت زیر بود:

$$y = X\beta + ZM\alpha + Z_n\varepsilon + e \qquad (۵)$$

در رابطه (۵)، y بردار مشاهدات برای حیوانات تعیین ژنوتیپ شده و نشده ، $X=\begin{bmatrix}1 & -Z_nA_{ng}A_{gg}^{-1}1\\ 1 & -Z_g1\end{bmatrix}$ ، $\beta=[\begin{matrix}\mu\\ \mu_g\end{matrix}]$، $\mu$: میانگین کل و $\mu_g$: نشان‌دهنده تفاوت در ارزش‌های اصلاحی بین حیوانات ژنوتیپ شده و نشده می‌باشد. علاوه براین، Z: ماتریس طرح، $M=[\begin{matrix}M_n\\ M_g\end{matrix}]$ که $M_g$: ماتریس ژنومی برای حیوانات ژنوتیپ شده و $M_n=A_{ng}A_{gg}^{-1}M_g$، نشان‌دهنده ماتریس ژنومی پیش‌بینی شده برای حیوانات ژنوتیپ نشده است. همچنین، $\varepsilon$: عوامل باقیمانده $Z_n$ و $Z_g$: ماتریس‌های طرح هستند که مشاهدات را به ارزش‌های اصلاحی حیوانات ژنوتیپ شده و نشده تخصیص می‌دهند. برای روش تک‌مرحله‌ای مبتنی بر BayesA(SSBR-A) توزیع پیشین مربوط به اثرات نشانگری توزیع t در نظر گرفته شد و همه نشانگرهای تک‌نوکلئوتیدی دارای واریانس بودند. در روش تک‌مرحله‌ای مبتنی بر BayesC(SSBR-C) اثرات نشانگری به‌صورت زیر در نظر گرفته شدند:

(۶)

$$\alpha_j|\pi,\sigma_{\alpha j}^2 \sim IID\begin{cases}0 \text{ with probablity } \pi\\ N(0,\sigma_{\alpha j}^2) \text{ with probablity } 1-\pi\end{cases}$$

که در رابطه (۶)، $\pi$: بیانگر نسبتی از نشانگرهایی است که اثرت آنها صفر در نظر گرفته شده است. مقادیر $\pi$ که بالاترین صحت پیش‌بینی را در جمعیت تایید دارا بود برای روش‌های BayesB، BayesC و SSBR–C در نظر گرفته شده است. صحت پیش‌بینی در روش‌های BayesB و BayesC با استفاده از مقادیر مختلف $\pi$ (۰/۹۹۹۹، ۰/۹۹۹، ۰/۹۹۵، ۰/۹۹۰، ۰/۹۸۰، ۰/۹۵۰، ۰/۹۰۰، ۰/۸۵۰، ۰/۸۰۰، ۰/۷۵۰، ۰/۷۰۰، ۰/۶۵۰ و ۰/۶۰۰) مورد مقایسه قرار گرفتند. به منظور پیاده‌سازی روش تک‌مرحله‌ای بیزی از بسته نرم‌افزاری JWAS در زبان برنامه‌نویسی JULIA استفاده گردید (۶). علاوه بر این، از بسته نرم‌افزاری BGLR به منظور پیاده‌سازی روش‌های چندمرحله‌ای بیزی استفاده شد (۲۹). برای تجزیه و تحلیل داده‌های شبیه‌سازی شده از زنجیره مارکوف مونت کارلو با زنجیره‌ای به طول ۵۰۰۰۰ نمونه شامل ۵۰۰۰ نمونه برای گرم کردن و ۴۵۰۰۰ نمونه بعدی برای استنباط‌های توزیع پسین استفاده شد. صحت برآورد ارزش‌های اصلاحی ژنومی از همبستگی بین ارزش‌های اصلاحی واقعی و ارزش‌های اصلاحی ژنومی برآورده شده بدست آمد. تغییرات صحت GEBV برای نسل مرجع و سه نسل گروه تایید بررسی شد. هر سناریوی شبیه‌سازی که در برگیرنده توزیع اثرات QTL بوده ۱۰ بار تکرار شد و میانگین تکرارها همراه با خطای استاندارد میانگین‌ها محاسبه شد.

1- Bayesian Least Absolute Shrinkage and Selection Operator

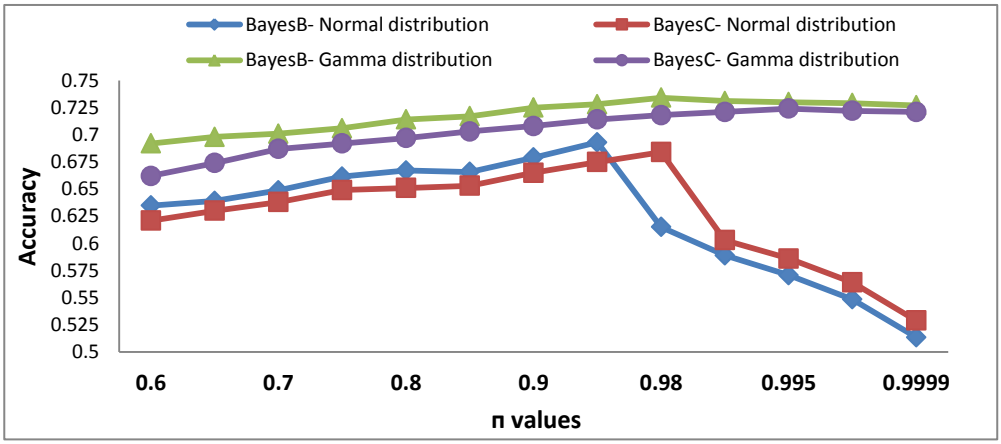
[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

## نتایج و بحث

### بررسی تاثیر مقادیر متفاوت π در پیش‌بینی ژنومی

صحت پیش‌بینی ارزش‌های اصلاحی ژنومی برای دو توزیع متفاوت اثرات ژنی مربوط به BayesB و BayesC در مقادیر مختلف π در شکل ۱ نشان داده شده است. مقادیر بهینه π که بالاترین صحت پیش‌بینی ژنومی را با استفاده از روش BayesB دارا بودند به‌ترتیب ۰/۹۵۰ و ۰/۹۸۰ در دو توزیع نرمال و گامای اثرات ژنی تخمین زده شد. مقادیر بهینه در نظرگرفته شده π در روش BayesC به‌ترتیب ۰/۹۸۰ و ۰/۹۹۵ در دو توزیع نرمال و گامای اثرات ژنی برآورد گردید و در روش SSBR-C نیز مورد استفاده قرار گرفت. برای توزیع نرمال اثرات ژنی، صحت‌های پیش‌بینی شده با روش‌های BayesB و BayesC مشابه بودند، اما صحت پیش‌بینی شده با افزایش مقادیر π از مقدار بهینه به بعد کاهش یافت. در توزیع گامای اثرات ژنی، الگوی متفاوتی مشاهده گردید، به‌طوری‌که صحت‌های پیش‌بینی شده با افزایش مقادیر π تا مقدار بهینه افزایش و سپس روندی تقریبا ثابت داشت و همواره صحت‌های ژنومی با استفاده از BayesB بالاتر از BayesC تخمین زده شد. نتایج نشان داد که صفات با توریع گامای اثرات ژنی تحت تاثیر تعداد کمی QTL قرار دارند، به‌طوریکه می‌توان بخش قابل توجهی از واریانس ژنتیکی را به وسیله آنها توجیه کرد. همچنین، روش BayesB در مقایسه با BayesC، QTLهای بیشتری با اثرات کوچک را به صفر منقبض کرد، که در نتیجه آن موجب بدام انداختن QTLهای با اثرات بزرگتر و افزایش صحت‌های پیش‌بینی ژنومی را در توزیع گامای اثرات ژنی به همراه داشت (۳۴). صحت‌های پیش‌بینی شده مشابه روش‌های BayesB و BayesC در مقادیر π بالا نشان داد که این دو روش در این شرایط می‌توانند در هر دو توزیع اثرات ژنی، تعداد QTL های کم با اثرات بزرگ را توجیه کنند.




شکل ۱- میانگین صحت‌های پیش‌بینی ژنومی با استفاده از مقادیر مختلف π در توزیع‌های متفاوت اثرات ژنی نرمال و گاما

Figure 1. The mean prediction accuracies using different π values in the normal and gamma distributions of QTL

### مقایسه روش‌های چند مرحله‌ای و تک مرحله‌ای بیزی

میانگین صحت پیش‌بینی ارزیابی ژنومی حیوانات ژنوتیپ شده روش‌های مورد مطالعه در جدول ۲ و ۳ نشان داده شد. میانگین صحت پیش‌بینی ژنومی روش‌های چند مرحله‌ای بیزی شامل BayesA، BayesB (π=0.98)، BayesB (π Estimation)، BayesC (π=0.995)، BayesC (π=0)، BayesC (π Estimation) و BayesL به‌ترتیب ۰/۷۱۲، ۰/۷۳۴، ۰/۷۲۳، ۰/۷۲۴، ۰/۷۲۱، ۰/۷۱۶و ۰/۶۸۷ در حالت توزیع گامای اثرات ژنی و در نسل اول برآورد شد. به‌طوریکه میانگین پیش‌بینی ژنومی روش‌های تک‌مرحله‌ای بیزی شامل SSBR-A، SSBR–C (π=0.995)، SSBR–C (π=0) و SSBR–C (π Estimation) در شرایطی مشابه به ترتیب ۰/۷۱۸، ۰/۷۷۳، ۰/۷۰۶ و ۰/۷۵۶ به‌دست آمد. صحت پیش‌بینی ژنومی در روش چندمرحله‌ای BayesB (π=0.98) نسبت به سایر روش‌های چندمرحله‌ای بیزی (BayesA، BayesB (π Estimation)، BayesC (π=0.995)، BayesC (π=0)، BayesC (π Estimation)، BayesL) از ۰/۰۱ تا ۰/۰۵ در حالت توزیع گامای اثرات ژنی بیشتر برآورد شد.

1- Bayesian Least Absolute Shrinkage and Selection Operator

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

جدول ۲- میانگین صحت (خطای استاندارد) پیش‌بینی ژنومی برای حیوانات ژنوتیپ شده در توزیع گامای اثرات ژن‌ها طی سه نسل متوالی
Table 2. Meanof genomic prediction accuracies and standard errors (in parenthesis) for the genotyped animals in the gamma distribution of QTL effects during three consecutive generations

| روش‌ها | نسل ۸ | نسل ۹ | نسل ۱۰ |
|---|---|---|---|
| BayesA | ۰/۷۱۲ (۰/۰۰۸) | ۰/۶۶۴ (۰/۰۰۹) | ۰/۶۲۸ (۰/۰۰۹) |
| BayesB (π=0.98) | ۰/۷۳۴ (۰/۰۰۹) | ۰/۶۹۱ (۰/۰۱) | ۰/۶۶۴ (۰/۰۱) |
| BayesB (π Estimation) | ۰/۷۲۳ (۰/۰۰۶) | ۰/۶۸۱ (۰/۰۰۷) | ۰/۶۵۷ (۰/۰۰۷) |
| BayesC (π=0.995) | ۰/۷۲۴ (۰/۰۰۹) | ۰/۶۷۶ (۰/۰۱) | ۰/۶۴۵ (۰/۰۱) |
| BayesC (π=0) | ۰/۷۲۱ (۰/۰۰۹) | ۰/۶۶۳ (۰/۰۲) | ۰/۶۳۲ (۰/۰۲) |
| BayesC (π Estimation) | ۰/۷۱۶ (۰/۰۰۶) | ۰/۶۴۹ (۰/۰۰۷) | ۰/۶۲۵ (۰/۰۰۸) |
| BayesL | ۰/۶۸۷ (۰/۰۰۸) | ۰/۶۳۵ (۰/۰۱) | ۰/۵۹۳ (۰/۰۱) |
| SSBR-A | ۰/۷۱۸ (۰/۰۰۸) | ۰/۶۷۴ (۰/۰۰۹) | ۰/۶۰۳ (۰/۰۲) |
| SSBR–C (π=0.995) | ۰/۷۷۳ (۰/۰۰۶) | ۰/۷۱۸ (۰/۰۱) | ۰/۷۰۱ (۰/۰۱) |
| SSBR–C (π=0) | ۰/۷۰۶ (۰/۰۰۶) | ۰/۶۴۷ (۰/۰۱) | ۰/۶۱۶ (۰/۰۱) |
| SSBR–C (π Estimation) | ۰/۷۵۶ (۰/۰۰۸) | ۰/۶۹۱ (۰/۰۲) | ۰/۶۷۸ (۰/۰۲) |

جدول ۳- میانگین صحت (خطای استاندارد) پیش‌بینی ژنومی برای حیوانات ژنوتیپ شده در توزیع نرمال اثرات ژن‌ها طی سه نسل متوالی
Table 3. Meanof genomic prediction accuracies and standard errors (in parenthesis) for the genotyped animals in the normal distribution of QTL effects during three consecutive generations

| روش‌ها | نسل ۸ | نسل ۹ | نسل ۱۰ |
|---|---|---|---|
| BayesA | ۰/۶۹۱ (۰/۰۰۶) | ۰/۶۲۹ (۰/۰۰۷) | ۰/۶۰۳ (۰/۰۰۹) |
| BayesB (π=0.95) | ۰/۶۹۳ (۰/۰۰۶) | ۰/۶۳۷ (۰/۰۰۸) | ۰/۶۲۱ (۰/۰۰۸) |
| BayesB (π Estimation) | ۰/۶۸۶ (۰/۰۰۸) | ۰/۶۳۴ (۰/۰۰۹) | ۰/۶۱۹ (۰/۰۰۹) |
| BayesC (π=0.98) | ۰/۶۸۴ (۰/۰۰۷) | ۰/۶۲۵ (۰/۰۰۸) | ۰/۶۰۷ (۰/۰۰۸) |
| BayesC (π=0) | ۰/۶۸۹ (۰/۰۰۳) | ۰/۶۳۱ (۰/۰۰۷) | ۰/۵۹۸ (۰/۰۰۹) |
| BayesC (π Estimation) | ۰/۶۷۶ (۰/۰۰۶) | ۰/۶۲۶ (۰/۰۰۹) | ۰/۵۸۲ (۰/۰۱) |
| BayesL | ۰/۶۷۶ (۰/۰۰۶) | ۰/۶۳۲ (۰/۰۰۷) | ۰/۶۱۴ (۰/۰۰۹) |
| SSBR-A | ۰/۷۰۴ (۰/۰۰۸) | ۰/۶۳۳ (۰/۰۰۹) | ۰/۶۰۷ (۰/۰۱) |
| SSBR–C (π=0.98) | ۰/۷۰۶ (۰/۰۰۹) | ۰/۶۶۴ (۰/۰۱) | ۰/۶۳۵ (۰/۰۱) |
| SSBR–C (π=0) | ۰/۶۹۸ (۰/۰۰۶) | ۰/۶۵۱ (۰/۰۰۸) | ۰/۶۲۶ (۰/۰۰۸) |
| SSBR-C (π Estimation) | ۰/۷۰۲ (۰/۰۰۸) | ۰/۶۵۷ (۰/۰۱) | ۰/۶۲۴ (۰/۰۱) |

گودارد (۱۳) گزارش کرد که توزیع پیش فرض روش‌های مبتنی بر BayesB با توزیع گامای ژن‌های عمده اثر هماهنگی دارد که باعث برآورد صحیح‌تر اثرات چندشکلی‌های تک‌نوکلئوتیدی می‌شود، به‌طوریکه توزیع پیش فرض گاما در روش BayesB برآورد صحیح‌تری از این اثرات ایجاد می‌کند. علاوه براین، با افزایش فاصله نسل بین جمعیت مرجع و جمعیت تایید صحت ارزش‌های اصلاحی ژنومی به‌شدت کاهش یافت (۱۷،۴). صحت‌های پیش‌بینی با استفاده از روش SSBR–C (π=0.995) در حالت توزیع گامای اثرات ژنی بالاتر از روش‌های چندمرحله‌ای بیزی و سایر روش‌های تک‌مرحله‌ای بیزی تخمین زده شد، به‌طوریکه مقدار این برتری از ۰/۰۲ تا ۰/۰۹ برآورد گردید. برتری روش تک‌مرحله‌ای SSBR–C (π=0.995) بیانگر این واقعیت است که در حالت توزیع گامای اثرات ژنی، تعداد کمی از *QTL*ها بخش عمده‌ای از واریانس ژنتیکی را به خود اختصاص می‌دهد. بنابراین، روش‌های دارای توزیع مزدوج، می‌توانند عملکرد مناسب‌تری را در این شرایط از خود نشان دهند. دلیل دیگر این امر، این است که در روش SSBR–C(π=0.995) از رکورد حیوانات تعیین ژنوتیپ شده و نشده به‌طور همزمان استفاده شد، بنابراین، در این روش تعداد مشاهدات بیشتری نسبت به روش‌های چندمرحله‌ای بیزی استفاده می‌شود (۲۱). نادری (۲۶) نیز نشان داد که صحت پیش‌بینی با افزایش درصد حیوانات رکوردارمی‌تواند روند صعودی داشته باشد.

همانطورکه در جداول ۲ و ۳ مشاهده می‌شود، روش SSBR–C (π=0.995) در حالت توزیع گامای اثرات ژنی (۰/۷۷۳) عملکرد مناسب‌تری نسبت به روش SSBR–C (π=0.98) درحالت توزیع نرمال (۰/۷۰۶) از خود نشان داد. هنگام در نظر گرفتن توزیع گاما برای اثرات ژنی، برخی از ژن‌ها دارای اثر عمده و درصد بالایی از ژن‌ها اثرات نزیک به صفر دارند، لذا روش‌هایی که دارای توزیع مزدوج باشند، در حالت توزیع گامای اثرات ژنی، عملکرد بهتری خواهند داشت (۱۱،۲۱). مشابه بودن نتایج حاصل از روش‌های چندمرحله‌ای در پیش‌بینی ارزش‌های اصلاحی ژنومی را می‌توان به معماری ژنتیکی پلی‌ژنی و توزیع نرمال اثرات ژنی مرتبط دانست (۱،۷). در این مطالعه نیز به‌دلیل عدم تفاوت بین صحت روش‌های چندمرحله‌ای و تک‌مرحله‌ای بیزی در

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]
[ DOI: 10.29252/rap.10.26.122 ]

حالت توزیع نرمال اثرات ژنی (جدول ۳) و با توجه به توانایی روش‌های تک‌مرحله‌ای بیزی در پیش‌بینی ارزش‌های اصلاحی حیوانات تعیین ژنوتیپ نشده، به‌کارگیری روش‌های تک‌مرحله‌ای برای ارزیابی ژنومی در این معماری ژنتیکی نیز ارجعیت دارد. از طرف دیگر، نتایج بدست آمده در هر دو حالت توزیع گاما و نرمال اثرات ژنی نشان می‌دهد که با افزایش فاصله نسل بین جمعیت مرجع و تایید، وضعیت LD بین QTLها و نشانگرها به هم خورده و اثرات QTLها نمی‌توانند به‌وسیله نشانگرها به خوبی توجیه شوند. این نتایج با یافته‌های پیشین نیز تطابق دارد (۱۴،۱۸). نکته دیگر این است که با افزایش فاصله بین نسل مرجع و نسل هشتم و نهم، خطای استاندارد صحت‌های پیش‌بینی ارزش‌های اصلاحی نیز افزایش یافته است. مویر (۲۵) نیز نشان داد که با گذشت چند نسل پس از برآورد اثر نشانگرها، صحت پیش‌بینی ارزش‌های اصلاحی ژنومی کاهش می‌یابد و از کارایی روش‌های برآورد اثرات نشانگری کاسته شده و لازم است اثرات نشانگرها مجددا برآورد شود.

**مقایسه روش‌های تک‌مرحله‌ای بیزی**

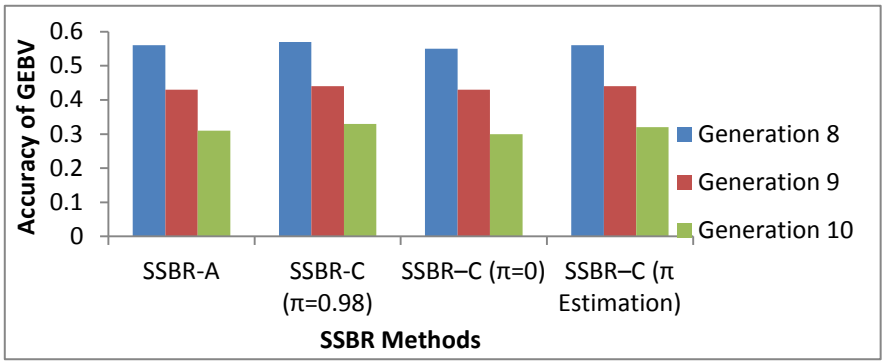
صحت پیش‌بینی ارزیابی ژنومی چهار روش تک‌مرحله‌ای بیزی SSBR-A، SSBR–C ($\pi$=0.995)، SSBR–C ($\pi$=0) و SSBR–C ($\pi$ Estimation) در حالت توزیع گاما و نرمال اثرات ژنی در جداول ۲ و ۳ نشان داده شده است. در حالت توزیع گامای اثرات ژنی، روش‌های SSBR–C ($\pi$ Estimation) و SSBR–C ($\pi$=0.995) عملکرد مناسب‌تری را نسبت به SSBR-A از خود نشان دادند به‌طوری‌که این برتری به‌ترتیب ۰/۰۴ و ۰/۰۶ تخمین زده شد. برتری SSBR-C نسبت به SSBR-A در این مطالعه ناشی از توزیع مزدوجی است که در روش SSBR-C مورد استفاده قرار می‌گیرد و باعث تغییر ساختار نشانگری شده به‌طوری‌که در این ساختار اثرات نشانگری به QTL ها نزدیک‌تر و دامنه اثرات نشانگری بیشتر تخمین زده می‌شود (۳۵). همچنین، برتری ۰/۰۷ صحت پیش‌بینی ارزش‌های اصلاحی ژنومی توسط روش SSBR–C ($\pi$=0.995) نسبت به روش SSBR–C ($\pi$=0) در حالت توزیع گامای اثرات ژنی نشان می‌دهد که در نظر گرفتن توزیع مزدوج در این شرایط می‌تواند قدرت مدل را در توجیه واریانس ژنتیکی افزایشی و به دام انداختن مقدار بیشتر LD بین نشانگرها و QTLها بهبود بخشد. علاوه بر این، با توجه به درنظر گرفتن تعداد ۲۰۰ QTL در سناریوهای شبیه‌سازی شده و وجود معماری ژنتیکی پلی‌ژنی در حالت توزیع نرمال اثرات ژنی، روش‌های تک‌مرحله‌ای بیزی عملکرد مشابهی داشتند. کارمن و همکاران (۲۰) و کانگ و همکاران (۱۹) نشان دادند که بازدهی روش‌های تک‌مرحله‌ای مبتنی بر BLUP و بیزی تحت توزیع نرمال اثرات ژنی مشابه بود که تاییدی بر نتایج تحقیق حاضر است.

**تاثیر روابط خویشاوندی شجره‌ای و ژنومی بین جمعیت مرجع و تایید**

صحت پیش‌بینی ارزش‌های اصلاحی ژنومی برای حیوانات ژنوتیپ نشده در نسل‌های ۸، ۹ و ۱۰ با استفاده از روش‌های تک‌مرحله‌ای بیزی در دو حالت توزیع اثرات ژنی گاما و نرمال در شکل‌های ۲ و ۳ نشان داده شده است. با افزایش فاصله بین جمعیت مرجع و تایید، صحت پیش‌بینی در هر دو حالت توزیع اثرات ژنی کاهش یافت. این روند کاهشی برای روش‌های SSBR-C نسبت به روش SSBR-A تا حدودی بیشتر مشاهده شد. این روند کاهشی صحت‌های پیش‌بینی شده با مطالعات متعددی تطابق داشت (۱۹،۳۴). این نتایج نشان می‌دهد که هر چه روابط خویشاوندی ژنومی بین افراد نزدیک‌تر باشد، صحت پیش‌بینی ژنومی به مقدار کمتری کاهش می‌یابد (۲۷). هابیر و همکاران (۱۵) و کانگ و همکاران (۱۹) نشان دادند که به موازات کاهش ارتباط ژنتیکی بین افراد، فاصله بین صحت‌های پیش‌بینی شده افزایش پیدا می‌کند. دتویلر و همکاران (۸) در جمعیت گوسفند آمیخته، نتیجه گرفتند که بخش قابل توجهی از صحت پیش‌بینی ارزش‌های اصلاحی ژنومی به‌جای LD بین نشانگرها و QTLها تحت تاثیر ساختار جمعیت و روابط خویشاوندی شجره‌ای بین افراد قرار می‌گیرد. با مقایسه جدول ۲ و شکل ۳ می‌توان نتیجه گرفت که افت صحت پیش‌بینی ژنومی با افزایش فاصله نسلی برای افراد ژنوتیپ شده در مقایسه با افراد ژنوتیپ نشده از حساسیت کمتری برخوردار بود. این بدان معنی است که بخش قابل توجهی از این کاهش صحت پیش‌بینی را می‌توان به روابط خویشاوندی شجره‌ای نسبت داد (۳۵). هابیر و همکاران (۱۴) نشان دادند که صحت‌های پیش‌بینی شده با استفاده از روش‌های GBLUP وRR-BLUP نسبت به روش‌های بیزی بیشتر تحت تاثیر روابط خویشاوندی ژنتیکی اعم از شجره‌ای و ژنومی قرار می‌گیرند. رابطه $M_n=A_{ng}A_{gg}^{-1}M_g$ نشان می‌دهد که روابط خویشاوندی شجره‌ای در پیش‌بینی روابط خویشاوندی ژنومی حیوانات ژنوتیپ نشده موثر است به‌طوریکه عنصر موثر $A_{ng}$ می‌تواند $M_n$ را بیشتر از $M_g$ تحت تاثیر قرار دهد ($M_n$، ماتریس روابط خیوشاوندی ژنومی پیش‌بینی شده برای حیوانات ژنوتیپ نشده، $M_g$ ماتریس روابط خویشاوندی ژنومی برای حیوانات ژنوتیپ شده). چن و همکاران (۵) گزارش کردند که افراد دارای روابط خویشاوندی شجره‌ای نزدیک به جمعیت مرجع در مقایسه با افراد دارای روابط خویشاوندی ضعیف‌تر، می‌توانند از صحت ژنومی پیش‌بینی شده بالاتری برخوردار باشند. نتایج این پژوهش نشان می‌دهد که روابط خویشاوندی ژنتیکی شجره‌ای بین جمعیت مرجع و تایید می‌تواند تاثیر به‌سزایی در تخمین صحت‌های ژنومی پیش‌بینی شده در حیوانات ژنوتیپ نشده داشته باشد (جدول ۳ و شکل ۲).

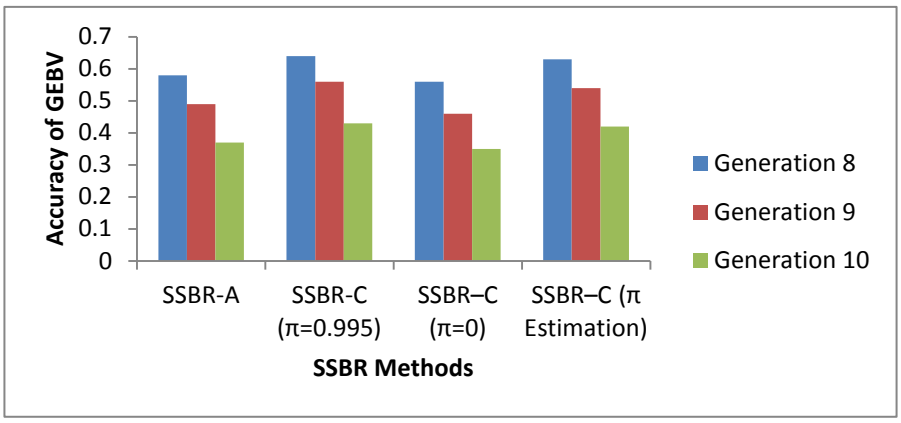
[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]




شکل ۲- میانگین صحت پیش‌بینی ارزش‌های اصلاحی ژنومی با استفاده از روش‌های تک‌مرحله‌ای بیزی برای حیوانات ژنوتیپ نشده مربوط به هر نسل از جمعیت تایید در توزیع نرمال اثرات ژن‌ها

Figure 2. Mean of genomic prediction accuracies using single step Bayesian regression for the nongenotyped animals for each validation set in the normal distribution




شکل ۳- میانگین صحت پیش‌بینی ارزش‌های اصلاحی ژنومی با استفاده از روش‌های تک‌مرحله‌ای بیزی برای حیوانات ژنوتیپ نشده مربوط به هر نسل از جمعیت تایید در توزیع گامای اثرات ژن‌ها

Figure 3. Mean of genomic prediction accuracies using single step Bayesian regression for the nongenotyped animals for each validation set in the gamma distribution

میانگین ضریب تابعیت ارزش‌های اصلاحی واقعی به ارزش‌های اصلاحی تخمین زده شده، در جدول ۴ نشان داده شده است. ضریب تابعیت پیش‌بینی ژنومی به‌ترتیب برای روش‌های BayesA و SSBR–C (π= chosen) ۱/۰۲ و ۰/۹۹ در حالت توزیع گامای اثرات ژنی و ۱/۰۴ و ۰/۹۵ در حالت توزیع نرمال اثرات ژنی تخمین زده شد. ضرایب تابعیت پیش‌بینی ژنومی کمتر از یک نشان می‌دهد که واریانس ارزش‌های اصلاحی ژنومی بیش از حد برآورد شده است. دلیل احتمالی این مسئله این است که نشانگرها در LD کامل با ژن‌های عامل نبوده و لذا واریانس ژنتیکی کل به‌طور کامل محاسبه نمی‌شود. با توجه به مقادیر برآورد شده، می‌توان گفت که ارزش اصلاحی ژنومی به‌دست آمده توسط روش SSBR–C (π=0) نسبت به سایر روش‌ها نااریب‌تر بوده است؛ زیرا که ضریب تابعیت آن در حالت توزیع گاما و نرمال اثرات ژنی به‌ترتیب مساوی و نزدیک به یک تخمین زده شد.

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

جدول ۴- ضریب تابعیت (خطای استاندارد) ارزش‌های اصلاحی واقعی به ارزش‌های اصلاحی تخمین زده شده در توزیع اثرات ژنی نرمال و گاما

Table 4. Regression and standard error (in parenthesis) of true on estimated breeding value in the distribution of normal and gamma QTL effects

| روش‌ها | توزیع اثرات ژنی | |
|---|---|---|
| | گاما | نرمال |
| BayesA | ۱/۰۲ (۰/۰۳) | ۱/۰۴ (۰/۰۲) |
| BayesB (π=chosen)[a] | ۱/۰۲ (۰/۰۲) | ۱/۰۳ (۰/۰۲) |
| BayesB (π Estimation) | ۱/۰۳ (۰/۰۲) | ۱/۰۵ (۰/۰۲) |
| BayesC (π=chosen)[b] | ۱/۰۳ (۰/۰۳) | ۱/۰۴ (۰/۰۳) |
| BayesC (π=0) | ۱/۰۰ (۰/۰۳) | ۱/۰۳ (۰/۰۲) |
| BayesC (π Estimation) | ۱/۰۲ (۰/۰۱) | ۱/۰۵ (۰/۰۱) |
| BayesL | ۱/۰۳ (۰/۰۳) | ۱/۰۵ (۰/۰۲) |
| SSBR-A | ۱/۰۶ (۰/۰۱) | ۱/۰۳ (۰/۰۲) |
| SSBR–C (π=chosen) | ۰/۹۹ (۰/۰۲) | ۰/۹۵ (۰/۰۳) |
| SSBR–C (π=0) | ۱/۰۰ (۰/۰۲) | ۰/۹۹ (۰/۰۲) |
| SSBR-C (π Estimation) | ۰/۹۶ (۰/۰۲) | ۰/۹۸ (۰/۰۱) |

a: مقدار π انتخاب شده برای روش BayesB به‌ترتیب ۰/۹۵۰ و ۰/۹۸۰ در دو توزیع نرمال و گامای اثرات ژنی در نظر گرفته شد.

b: مقدار π انتخاب شده برای روش‌های BayesC و SSBR-C به‌ترتیب ۰/۹۸۰ و ۰/۹۹۵ در دو توزیع نرمال و گامای اثرات ژنی در نظر گرفته شد.

روش‌های تک‌مرحله‌ای بیزی که از مزیت استفاده همزمان از اطلاعات فنوتیپی، ژنوتیپی و شجره بهره‌مند هستند، صحت‌های مشابه و یا بالاتری را نسبت به روش‌های چند مرحله‌ای بیزی از خود نشان دادند. در بین روش‌های چند مرحله‌ای، روش BayesB (π=0.98) بهترین عملکرد را در حالت توزیع گامای اثر ژن‌ها از خود نشان داد. علاوه بر این، استفاده از روش‌های SSBR–C (π=0.98) و SSBR–C (π=0) در حالت توزیع نرمال اثرات ژنی دارای برتری نبود. با توجه به نتایج حاضر، هنگامی که معماری صفات مورد بررسی از مدل تعداد زیاد جایگاه‌های ژنی تبعیت نکند، روش SSBR–C (π=0.995) و BayesB (π=0.98) نسبت به روش‌های دیگر ارجحیت دارد. همچنین، ضریب تابعیت در توزیع اثرات ژنی گاما و نرمال در روش SSBR–C (π=0) نزدیک به یک تخمین زده شد، بنابراین، نسبت به سایر روش‌ها نااریب‌تر بوده و پیشنهاد می‌شود که در توزیع نرمال اثرات ژنی از این روش استفاده شود. علاوه بر این، نتایج این پژوهش نشان داد که افت صحت پیش‌بینی ژنومی با افزایش فاصله نسلی برای افراد ژنوتیپ شده در مقایسه با افراد ژنوتیپ نشده از حساسیت کمتری برخوردار بود.

## تشکر و قدردانی

از جناب آقای دکتر مهدی ساعتچی به‌دلیل اظهار نظرهای ارزشمند ایشان سپاسگزاری می‌شود.

## منابع


1. Abdollahi-Arpanahi, R. 2013. The Impact of allelic architecture of complex traits on genetic evaluations and evolutionary genetics. Ph.D. Dissertation University of Tehran, Tehran.
2. Abdollahi-Arpanahi, R., A. Pakdel, A. Nejati-Javaremi and M.M. Shahrbabak. Comparison ofdifferent methods of genomic evolution in traits with different genetic architecture. Journal of Animal Production, 15: 65-77 (In Persian).
3. Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra and S. Tsuruta. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science, 93: 743-752.
4. Calus, M.P. 2009. Genomic breeding value prediction: methods and procedures. Animal, 4(2): 157-164.
5. Chen, L, C. Li, M. Sargolzaei and F. Schenkel. 2014. Impact of genotype imputation on the performance of GBLUP and Bayesian methods for genomic prediction. PLoS ONE, 9(7): e101544.
6. Cheng, H., D.J. Garrick and R.L. Fernando. 2016. JWAS: Julia implementation of whole-genome analyses software using univariate and multivariate Bayesian mixed effects model. Retrieved June 8, 2019 from http://QTL.rocks.
7. Colombani, C., A. Legarra, S. Fritz, F. Guillaume, P. Croiseau, V. Ducrocq and C. Robert-Granié. 2012. Application of Bayesian least absolute shrinkage and selection operator (LASSO) and BayesCp methods for genomic selection in French Holstein and Montbéliarde breeds. Journal of Dairy Science, 96: 575-91.
8. Daetwyler, H.D., K.E. Kemper, J.H. Vander Werf and B.J. Hayes. 2012. Components of the accuracy of genomic prediction in a multi-breed sheep population, Journal of Animal Science, 90(10): 3375-3384.
9. De los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra, E. Manfredi, K. Weigel and J.M. Cotes. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics, 182(1): 375-385.
10. De los Campos, G., A.I. Vazquez, R. Fernando, Y.C. Klimentidis and D. Sorensen. 2013. Prediction of complex human traits using the genomic best linear unbiased predictor. PLoS Genetics, 9: e1003608.


[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

11. Fernando, R.L., J.C. Dekkers and D.J. Garrick. 2014. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. Genetics Selection Evolution, 46: 50.
12. Gianola, D., G. DeLos Campos, W.G. Hill, E. Manfredi and R. Fernando. 2009. Additive genetic variability and the Bayesian alphabet. Genetics, 183: 347-363.
13. Goddard, M. 2009. Genomic selection: prediction of accuracy and maximization of long term response. Genetica, 136: 245-257.
14. Habier, D., R.L. Fernando, J.C. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics, 177(4): 2389-2397.
15. Habier, D., J. Tetens, F.R. Seefried, P. Lichtner and G. Thaller. 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genetics Selection Evolution, 42(1): 5.
16. Habier, D., R.L. Fernando, K. Kizilkaya and D.J. Garrick. 2011. Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics, 2: 186-194.
17. Hayes, B.J., P.J. Bowman, A.J Chamberlain and M.E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of DairyScience, 92: 433-443.
18. Hickey, J., S. Dreisigacker, J. Crossa, S. Hearne, R. Babu, B. Prasanna, M. Grondona, A. Zambelli, V. Windhausen, K. Mathews and G. Gorjanc. 2014. Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. Crop Science, 54: 1476-1488.
19. Kang, H., L. Zhou, R. Mrode, Q. Zhang and J.F. Liu. 2016. Incorporating single-step strategy into random regression model to enhance genomic prediction of longitudinal trait. Heredity, 119: 459-467.
20. Karaman, E., H. Cheng, M.Z. Firat, D.J. Garrick and R.L. Fernando. 2016. An upper bound for accuracy of prediction using GBLUP. PLoS ONE, 11(8): e0161054.
21. Lee, J., H. Cheng, D. Garrick, B. Golden, J.C. Dekkers, K. Park, D. Lee and R. Fernando. 2017. Comparison of alternative approaches to single‐trait genomic prediction using genotyped and non‐genotyped Hanwoo beef cattle. Genetiscs Selection Evolution, 49: 2.
22. Legarra, A., I. Aguilar and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science, 92: 4656-63.
23. Meuwissen, T., B. Hayes and M. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157: 1819-29.
24. Moschopoulos, P.G. 1985. The distribution of the sum of independent gamma random variables. Annals of the Institute of Statistical Mathematics, 37(1): 541-544.
25. Muir, W.M. 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. Journal of Animal Breeding and Genetics, 124: 342-355.
26. Naderi, Y. 2018. The importance of genetic relationships and phenotypic record on genomic accuracy of simulated imputation data via animal models in presence of genotype × environment interactions. Research on Animal Production, 22: 119-130 (In Persian).
27. Nejati-Javaremi, A., C. Smith and J. Gibson. 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. Journal of animal science, 75: 1738-45.
28. Park, T. and G. Casella. 2008. The Bayesian lasso. Journal of the American Statistical Association, 103: 681-6.
29. Pérez, P. and G. Delos Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics, 198: 483-495.
30. Saheb Alam, H., M. Gholizadeh, H. Hafezian and A. Farhadi. 2017. Comparison of Bayesian methods in the genomic evaluation with different genetic architecture. Research on Animal Production, 18: 177-186 (In Persian).
31. Sargolzaei, M. and F.S. Schenkel. 2009. QMSim: a large-scale genome simulator for livestock. Bioinformatics, 25: 680.
32. Toghiani, S., S.E. Aggrey and R. Rekaya. 2016. Multi-generational imputation of single nucleotide polymorphism marker genotypesand accuracy of genomic selection. Animal, 10: 1077-85.
33. VanRaden, P.M., C.P. Van Tassell, G.R. Wiggans, T.S. Sonstegard, R.D. Schnabel, J.F. Taylor and F.S. Schenkel. 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. Journal of Dairy Science, 92: 16-24.
34. Wolc, A., J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan and J.C. Dekkers. 2016. Mixture models detect large effect QTL better than GBLUP and result in more accurate and persistent predictions. Journal of Animal Science, 7: 7.
35. Zhou, L., R. Mrode, S. Zhang, Q. Zhang, B. Li and J. Liu. 2018. Factors affecting GEBV accuracy with single-step Bayesian models. Heredity, 120: 100-109.

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]

# Comparison of Single and Multi-Step Bayesian Methods for Predicting Genomic Breeding Values in Genotyped and Non-Genotyped Animals-A Simulation Study

**Mostafa Madad Jirandeh[1], Jalil Shodja[2], SadeghAlijani[3], Seyed Abbas Rafat[2] and Jack C.M. Dekkers[4]**

1- Ph.D. Student, Department of Animal Sciences, University of Tabriz, Tabriz, Iran
2- Professor, Department of Animal Sciences, University of Tabriz, Tabriz, Iran.
3- Associate Professor, Department of Animal Sciences, University of Tabriz, Tabriz, Iran
(Corresponding author: sad-ali@tabrizu.ac.ir)
4- Professor, Department of Animal Sciences, Iowa state university, Ames, United States of America
Received: July 2, 2019 Accepted: September 11, 2019

## Abstract

The purpose of this study was to compare the accuracy of genomic evaluation for Bayes A, Bayes B, Bayes C and Bayes L multi-step methods and SSBR-C and SSBR-A single-step methods in the different values of π for predicting genomic breeding values of the genotyped and non-genotyped animals. A genome with 40000 SNPs on the 20 chromosom was simulated with the same distance (100cM). The π values that maximized the prediction accuracies in BayesC were 0.980 and 0.995 for the normal and gamma distributions of QTL, respectively, and were also used in SSBR-C method. Genomic prediction accuracy in the SSBR-C (π = 0.99) method was higher than multi step methods from 0.02 to 0.09 for gamma distribution. Results showed that considering mixture distribution and use of phenotype, genotype and pedigree information simultaneously, the SSBR-C (π = 0.99) method had higher accuracy than other methods and is considered a better choice in this scenario. Moreover, both single and multi-step methods showed similar prediction accuracy when the genetic architecture appeared to approach the normal distribution. Furthermore, SSBR-C (π = 0) method appeared to be more reliable choice that was due to regressions of true breeding value on estimated breeding value close to one in normal distribution. Generally, GEBV accuracy decreased as the distance increased between validations and training set, which was more sensitive for non-genotyped individuals compared to genotyped individuals.

**Keywords:** Bayesian, Genomic Accuracy, Genomic Selection, Multi-Step Methods, Simulation

[ Downloaded from rap.sanru.ac.ir on 2024-10-12 ]

[ DOI: 10.29252/rap.10.26.122 ]