

بررسی کارایی روش آماری BayesC π برای پویش ژنومی پاسخ ایمنی

سلولی در مرغ

A study of the efficiency of the BayesC π statistical method for Genome Wide Association study of cellular immune response in chicken

هستی معتمد^۱، علی‌رضا احسانی^{۲*}، رسول واعظ ترشیزی^۱

۱- کارشناسی ارشد، اصلاح دام و طیور، دانشگاه تربیت مدرس

۲- به‌ترتیب استادیار، دانشیار، گروه علوم دامی، دانشگاه تربیت مدرس

Motamed H¹, Ehsani AR^{*2}, Vaez Torshizi R²

1- MSc Student, Animal Breeding, of Tarbiat Modares University

2- Assistant Professor, Associate Professor, of Tarbiat Modares University

* نویسنده مسئول مکاتبات، پست الکترونیکی: alireza.ehsani@modares.ac.ir

(تاریخ دریافت: ۹۶/۵/۱۶ - تاریخ پذیرش: ۹۶/۱۲/۱۳)

چکیده

به‌کارگیری روش آماری برای آنالیز داده‌های ژنومی می‌تواند به یافتن روابط واقعی بین ژنتیک و فنوتیپ کمک کند. در سال‌های اخیر روش‌های آماری مختلفی برای درک این روابط به کار گرفته شده‌اند و نتایج به‌دست آمده نشان‌دهنده این واقعیت است که در شرایط متفاوت و صفات مختلف روش‌های متفاوتی کارایی بهتری دارند. از جمله این روش‌ها می‌توان به روش‌های مبتنی بر حداکثر درستنمایی، روش حداقل مربعات و یا روش‌های غیر پارامتری مانند روش‌های بیزی اشاره نمود. برخلاف روش‌های پارامتری، در روش‌های بیزی اثر متغیرهای ژنتیکی می‌تواند در بیش از یک توزیع آماری مورد آنالیز قرار گیرند. با توجه به اینکه توارث صفات کمی در نتیجه تأثیر متفاوت متغیرهای ژنتیکی با اثر متفاوت بوده و توزیع‌های آماری این روش می‌تواند متفاوت از هم‌دیگر باشد لذا به‌نظر می‌رسد روش‌های بیزی بتوانند درک صحیح‌تری از نواحی ژنومی مؤثر بر صفات کمی و میزان اثر آن‌ها ارائه نمایند. به‌همین جهت در پژوهش حاضر، فنوتیپ ایمنی سلولی با استفاده از تزریق دی نیتروکلرو بنزن و ژنوتیپ پرندگان با تراکم ۶۰ هزارتایی SNP شرکت ایلومینا در ۱۹۸ پرده نسل F2 حاصل از تلاقی دوطرفه سویه آراین و مرغ بومی ارومیه با استفاده از روش بیزی BayesC π مورد مطالعه قرار گرفت. در این روش، با توجه به ماهیت صفت (وراثت‌پذیری پایین) و مشاهده نمودار همگرایی مولفه‌های واریانس و مقادیر π با استفاده از آزمون و خطا، مقدار π برابر با ۰/۱۹۹۹ و تعداد دور زنجیره ۱۰۰۰۰۰ انتخاب شد تا بتوان همگرایی و پراکندگی اثر نشانگری را به‌درستی نشان داد. تعداد ۲۸ نشانگر معنی‌دار به‌دست آمده بر روی کروموزوم‌های ۱، ۲، ۳، ۴، ۸، ۱۳، ۱۴، ۱۹، ۲۰، ۲۱ و ۲۲ قرار داشتند که این تعداد با توجه به وراثت‌پذیری پایین صفت می‌تواند قابل توجیه باشد. SNP‌های مشخص شده داخل یا نزدیک به ۲۵ ژن قرار گرفته بودند و از این میان ژن‌های NCK2، LRRTM4، MKNKI، SULF2، RBFOX1، CAMK1D و DUSP14، عملکردهای مهمی را در ایجاد پاسخ ایمنی از طریق سنتز سلول‌های T، تشکیل سیناپس ایمنونولوژیکی، دخالت در تشکیل نوتروفیل‌ها و ایجاد پاسخ به استرس‌های محیطی بر عهده داشتند. با توجه به نقش کلیدی ژن‌های ذکر شده در ایجاد پاسخ ایمنی سلولی، می‌توان کارایی روش BayesC π برای پویش ژنومی در این صفت و صفاتی مشابه را مورد تأیید قرار داد.

واژه‌های کلیدی

چندشکلی تک نوکلئوتیدی

روش بیزین

ژنوم

صفت ایمنی سلولی

مرغ بومی - سویه آراین

مقدمه

در سالیان اخیر مطالعات پویس ژنوم¹ (GWAS) و انتخاب ژنومی به طور گسترده در برنامه‌های اصلاحی دام و طیور به کار گرفته شده‌اند. هدف نهایی از مطالعات GWAS، که به منظور شناسایی وابستگی بین یک نشانگر ژنتیکی و یک صفت یا بیماری با استفاده از نشانگرهای فراوان در سطح ژنوم است، پیدا کردن چند شکلی‌های سببی می‌باشد که بر فنوتیپ یک صفت اثر می‌گذارند (Risch and Merikangas 1996). از سوی دیگر مدل‌های آماری می‌توانند اثر تعیین کننده‌ای بر یافتن نواحی ژنومی مؤثر داشته باشند تا جایی که در صفات کمی متفاوت مدل‌های متفاوتی کارایی بهتری دارند. مدل‌های آماری مورد استفاده در آنالیزهای پویس ژنومی را می‌توان به دو دسته کلی تقسیم نمود. یکی از آن‌ها روش‌های آماری کلاسیک و دیگری مدل‌های مبتنی بر تئوری بیزی است، که در سال‌های اخیر مورد توجه بیش‌تری قرار گرفته است. از جمله روش‌های مبتنی بر مدل‌های کلاسیک می‌توان به آنالیزهای رگرسیونی تک متغیره مبتنی بر روش حداکثر درستنمایی (Maximum Likelihood) و یا روش حداقل مربعات خطا (Least Squares) اشاره نمود. مطابق با یافته‌های Peters et al. (2012) در آنالیزهای مبتنی بر روش‌های کلاسیک مهم‌ترین مشکل، بالا بودن نرخ false positive و over estimation از اثر SNPها و QTLها می‌باشد.

از طرف دیگر این روش‌ها براساس تئوری توزیع نرمال از آثار متغیرهای ژنتیکی استوار می‌باشند، که ممکن است تئوری صحیحی برای تفسیر اثر متغیرهای ژنتیکی در همه صفات کمی نباشد. زیرا فرض بر این است که در توزیع نرمال همه‌ی متغیرهای ژنتیکی از یک توزیع نرمال پیروی می‌کنند که آشکارا با یافته‌های دانشمندان در مورد اثرات جایگاه‌های ژنتیکی موثر بر صفات کمی مابینت دارد. بنابراین یک جایگزین مناسب برای حل این مشکلات، می‌تواند استفاده از روش‌های بیزین در آنالیزهای پویس ژنومی باشد که در آن می‌توان به دلخواه توزیع‌های متفاوتی برای دسته‌های مختلف از متغیرهای ژنتیکی در نظر گرفت.

Hayes and Goddard (2001) نشان دادند که استفاده از روش‌های بیزین دقت انتخاب ژنومی را بالا برده است، زیرا با

¹ Genome wide association study

استفاده از این مدل‌ها نشانگرهای بیشتری می‌توانند به طور همزمان آنالیز شوند، و در واقع نرخ over estimation و false positive کاهش پیدا کند. از سوی دیگر آماره‌ی مورد استفاده برای بررسی اثر نشانگرها در مدل‌های تک نشانگری، p-value می‌باشد که در طی پژوهش‌های انجام شده همواره تفاوت‌هایی بین نتایج حاصل از آنالیزهای بیزین و کلاسیک وجود داشته است، به گونه‌ای که آماره p-value فقط درصدی از اطلاعات را برای تصمیم‌گیری فراهم می‌کند و مشاهده شده‌است حتی در صورت رد فرض صفر، این فرض ممکن است تا ۲۰ درصد درست باشد (Lin and Yin 2015). استنباط بیزی بر اساس یک رابطه‌ی احتمال شرطی شکل می‌گیرد، که مشروط بر مشاهدات است.

مطالعه (Sahana et al. 2010) نشان داد که نتایج QTL mapping با روش بیزی، صحت و قدرت بالاتری در یافتن جایگاه QTLها در مقایسه با مدل رگرسیون تک نشانگری، مدل‌های مبتنی بر هاپلو تیپ یا مدل‌های میکس دارد. همچنین در مطالعه‌ای که با استفاده از داده‌های شبیه‌سازی شده از جمعیت گاو شیری انجام شد، دو روش Bayes C و Bayes Cπ مورد مطالعه قرار گرفتند که در نهایت روش Bayes Cπ کارایی بهتری برای یافتن اثر SNPها و همگرایی آنالیز نسبت به روش BayesC داشت (Van den Berg et al. 2013). در پژوهشی که به منظور مقایسه‌ی روش‌های بیزی و کلاسیک برای یافتن نشانگرهای تک نشانگری درگیرنده‌ی لپتین در خوک انجام شد، مشخص شد که روش بیزی نسبت به روش کلاسیک قوی‌تر عمل کرده است (Varona 2010). مطالعات نشان می‌دهند که در بسیاری از موارد نتایج حاصل از آنالیزهای کلاسیک و بیزین باهم متفاوت هستند (Johnson 2013; Varona 2010). از طرف دیگر در آنالیزهای بیزی زمان محاسبات و آنالیز داده‌ها مناسب‌تر بوده (Hayes and Goddard 2001) و تجزیه و تحلیل داده‌ها دارای چارچوب طبیعی‌تر و منطقی‌تری است (Sham and Purcell 2014). همچنین در این آنالیزها عدم تعادل پیوستگی بین SNP و جایگاه صفات کمی به طور دقیق‌تری نشان داده شده و به طور کلی این روش نسبت به روش‌های رگرسیونی و حداقل مربعات در یافتن جایگاه‌های موثر صفات کمی بهتر عمل می‌کند (Habier et al. 2011). علاوه بر این در روش‌های بیزی Over estimation کاهش یافته و مقادیر اثر

ضخامت پوست با استفاده از کولیس الکترونیکی در سه تکرار پیش از تیمار پرندگان و ۲۴ ساعت پس از تیمار مورد استفاده قرار گرفت (Verma et al. 2004).

تعداد ۱۹۸ نمونه خون از نسل F₂ با همکاری دانشگاه آرهوس دانمارک به‌وسیله‌ی پانل تجاری Illumina 60k SNP chip تعیین ژنوتیپ شدند که تعداد ۵۵۳۲۹ SNP برای جمعیت مورد مطالعه ژنوتایپ شد. کنترل کیفیت داده‌ها در نرم‌افزار PLINK نسخه‌ی ۱/۹ انجام شد. نمونه‌های با نرخ ژنوتیپی (call rate) کمتر از ۹۰ درصد، SNP‌های با MAF^۵ کمتر از پنج درصد و geno بیش‌تر از ۱۰ درصد و همچنین بخشی از SNP‌هایی که فراوانی آن‌ها در تعادل هاردی-وینبرگ نبود از تجزیه و تحلیل حذف شدند. پس از ویرایش‌های مربوطه در نهایت تعداد ۴۴۲۲۸ SNP برای آنالیزهای GWAS استفاده شد.

بر اساس قانون بیز احتمال یافتن متغیرهای موثر بر یک صفت به کمک پیش‌برآوردها و تولید هزاران توزیع پشت سرهم سلسله‌ای از پارامترهای مربوطه محاسبه می‌شود. در صورتی‌که پیش‌برآوردها از توزیع‌های آماری با احتمال وقوع متفاوت پارامترها در فضای نمونه انتخاب شوند احتمال یافتن پارامترها خارج از فضای نمونه کمتر شده و اصطلاحاً به آن پیش‌برآورد غنی از اطلاعات گفته می‌شود و برآورد نهایی عموماً متکی به پیش‌برآوردها بوده و نقش داده‌ها در برآوردها کمتر می‌شود. اما در صورتی‌که پیش‌برآوردهای انتخاب شده از توزیع‌های آماری با فضای نمونه‌ی باز و احتمال وقوع هر پارامتر به‌طور مساوی باشند به این پیش‌برآوردها اصطلاحاً پیش‌برآورد فاقد اطلاعات گفته می‌شود و در نتیجه برآورد نهایی بیش‌تر متکی به داده‌های موجود بوده و می‌تواند اطلاعات بیشتری از داده‌های موجود در اختیار قرار دهد. معادله‌ی بیزی به‌صورت زیر محاسبه می‌شود.

$$P(y | \theta) = \frac{p(\theta)p(\theta | y)}{p(y)}$$

در معادله فوق $P(y|\theta)$ معرف توزیع پسین از داده‌ها (y) مشروط به آماره (θ) می‌باشد. $P(\theta)$ معرف پیش‌برآورد آماره θ ، $p(\theta | y)$ بیانگر توزیع آماری آماره θ (likelihood) مشروط به داده‌های موجود و $p(y)$ توزیع آماری داده‌ها می‌باشد. آنالیزهای آماری برای روش بیزی با استفاده از Bayes Cπ انجام شد (Gianola et

Costa et al. 2015) از این رو دقت GEBV^۱های برآورد شده در روش‌های بیزین بالاتر است (Gianola et al. 2009). از سوی دیگر مطالعات نشان داده است که روش‌های بیزین در صفات با وراثت پذیری پایین ممکن است کارایی بهتری داشته باشند (Baneh 2015). بنابراین از آنجا که صفات مرتبط با ایمنی عموماً دارای وراثت‌پذیری پایینی هستند.

در پژوهش حاضر، پویس ژنومی در صفت پاسخ ایمنی سلولی به دی‌نیتروکلرو بنزن (DNCB) به‌عنوان محرکی برای فعالیت ایمنی سلولی، با استفاده از روش BayesCπ مورد مطالعه قرار گرفت. اجزای سیستم ایمنی سلولی^۲ که شامل نوتروفیل‌ها و سلول‌های T سیتوتوکسیک^۳ می‌باشند، باعث تخریب آنتی‌ژن‌های خارجی، حذف کل سلول آسیب‌دیده و در نتیجه کمک به حیوانات برای ایجاد مقاومت در برابر آنتی‌ژن‌های خارجی و سرطان می‌شود (Raeesi et al. 2017). بنابراین اهمیت این سلول‌ها در حفظ ایمنی همواره حائز اهمیت می‌باشد.

مواد و روش‌ها

در تحقیق حاضر از یک جمعیت F₂ حاصل از تلاقی ضربداری لاین گوشتی پرتولید آراین و توده بومی ایرانی که مشهور به مرغ بومی آذربایجان است استفاده شد. والدین این جوجه‌ها در بهمن ماه سال ۱۳۹۰ به دانشکده کشاورزی دانشگاه تربیت مدرس انتقال داده شدند و به صورت ضربداری تلاقی داده شدند (Maghsodi 2014). از این تلاقی‌ها جمعیت F₁ و از آمیزش F₁ها جمعیت جوجه‌های F₂ متشکل از تعداد ۷ خانواده ناتی و ۳۴ خانواده تنی در جمعیت F₂ مورد مطالعه قرار گرفت. جهت بررسی عملکرد سیستم ایمنی سلولی، ضخامت سطح مقطع قسمت بدون پر زیر بال راست به مساحت ۷ سانتی‌متر مربع در سن ۱۰ هفتگی پیش و پس از مالش محلول ۰/۱ میلی‌لیتر دی‌نیتروکلروبنزن^۴ (DNCB) به‌عنوان شاخص ایمنی سلولی استفاده شد. میانگین تفاوت اندازه

¹ Genomic estimated breeding values

² Cell-mediated immunity

³ Cytotoxic T-cells

⁴ Dy-initro chloro benzene

⁵ Minor allele frequency

BF مخفف کلمه Bayes Factor، P فراوانی احتمال پیشین از SNP‌های فاقد اثر و π احتمال پسین SNP‌هایی فاقد اثر می‌باشد. $1-p$ و $1-\pi$ به ترتیب احتمال پیشین و پسین SNP‌های دارای اثر می‌باشد.

سه دامنه برای بیز فاکتور و بررسی معنی‌داری تعیین شده‌است (Kass and Raftery 1995) که شامل:

(۱) بیز فاکتور بین ۳ تا ۲۰: ارتباط ضعیف با صفت دارد.

(۲) بیز فاکتور بین ۲۰ تا ۱۵۰: ارتباط قوی با صفت دارد.

(۳) بیز فاکتور بزرگتر از ۱۵۰: ارتباط قطعی با صفت دارد.

در این پژوهش مبنای ارزشیابی بیز فاکتور بین ۲۰ و ۱۵۰ و بزرگ‌تر از ۱۵۰ در نظر گرفته شده‌است.

بعد از انجام آنالیز نشانگرهایی که اثر معنی‌دار بر فنوتیپ مربوطه داشتند برای آنالیز بیوانفورماتیکی استفاده شدند. با استفاده از پایگاه داده NCBI^۱ و جستجو در ژنوم مرغ، ژن‌هایی که ۵۰۰۰۰۰ جفت‌باز پیرامون هر SNP معنی‌دار بر روی کروموزم مربوطه قرار گرفته بودند شناسایی شدند (Wolfsberg 2010). فهرست این ژن‌ها به پایگاه داده (www.uniprot.org) Uniprot^۲ منتقل شدند و بعد از ثبت فهرست ژن‌ها، مسیرهای بیوشیمیایی که از لحاظ آماری معنی‌دار بودند شناسایی شدند و به آن‌ها پرداخته شد.

نتایج و بحث

فیلترهای کنترل کیفیت داده باعث حذف ۲۵۶ نشانگر به علت تصحیح خطای ژنوتیپ، ۱۸۹۶ نشانگر به علت تعادل هاردی واینبرگ و ۸۶۴۹ نشانگر به علت MAF شد و در نهایت ۴۴۲۲۸ نشانگر SNP وارد آنالیز شد. بیش‌تر از ۹۰ درصد نشانگرها برای نمونه‌های موردنظر ژنوتیپ شده بودند و هیچ‌کدام از آنالیز حذف نشدند. مقادیر اولیه واریانس ژنتیکی و باقی‌مانده محاسبه شده با استفاده از نرم‌افزار DMU به ترتیب ۰/۰۱۱۵۴۰۹ و ۰/۴۵۸۶۵۷۴ به دست آمد. واریانس محاسبه شده برای SNP‌های دارای اثر، با استفاده از واریانس ژنتیکی ۰/۰۱۳۳۱ محاسبه شد. وراثت‌پذیری محاسبه شده برابر ۰/۰۲ بود. با توجه به اینکه در این نرم‌افزار، وراثت‌پذیری محاسبه شده با استفاده از ساختار شجره انجام شد و

در این روش نشانگرها به دو دسته مجزا شامل دسته‌ی π که فاقد هرگونه اثر بر روی صفت مورد نظر هستند و دسته‌ی $1-\pi$ که دارای اثر معنی‌دار بر روی صفت مورد نظر هستند تقسیم می‌شوند. واریانس پیش برآورد (priori) گروه اخیر از یک توزیع کای اسکوتر معکوس وزنی با درجه‌ی آزادی (v) و پارامتر وزنی (s^2) پیروی می‌کنند و اثرات SNP‌ها از یک توزیع t استیودنت تک متغیره است (Costa et al. 2015). در این مطالعه مقدار درجه‌ی آزادی $4/2$ و تعیین میزان s^2_p بر اساس واریانس ژنتیکی افزایشی (Habier et al. 2011) به دست آمد و همچنین طول زنجیره MCMC برابر با ۱۰۰۰۰۰ و مقدار $1-\pi$ برابر با ۰/۰۰۰۵ تعیین شد. ارزش‌های اولیه برای واریانس ژنتیکی و واریانس باقی‌مانده با استفاده از یک مدل آماری خطی مختلط در نرم‌افزار DMU به صورت زیر محاسبه شد:

$$y = Xb + Zu + e$$

در این مدل، y, b, u, e به ترتیب بردارهای مربوط به فنوتیپ، اثر ثابت همچون جنس و دوره جوجه‌کشی، اثر پلی‌ژنیک و مقادیر باقی‌مانده می‌باشد. Z و X به ترتیب ماتریس طرح ربط دهنده‌ی اثر ثابت و آثار پلی‌ژنیک به فنوتیپ‌های مشاهده شده می‌باشد. با استفاده از فرمول $Vg(\pi 2 \sum p_i q_i)$ مقادیر واریانس ژنتیکی برای SNP‌های موثر بر صفت محاسبه شد.

مدل آماری مورد استفاده در آنالیز بیزین برای یافتن جایگاه‌های موثر بر صفات مورد مطالعه (معادله ۲) تنظیم شد:

$$y = Xb + Zu + Wg + e$$

که در این مدل، y, b, u, g, e به ترتیب بردارهای مربوط به فنوتیپ، اثر ثابت همچون جنس و دوره جوجه‌کشی، اثرات پلی‌ژنیک، آثار SNP‌ها و مقادیر باقی‌مانده می‌باشد. X, Z, W به ترتیب ماتریس‌های طرح ربط دهنده‌ی اثرات عوامل، آثار پلی‌ژنیک و آثار SNP‌ها به فنوتیپ‌های مشاهده شده می‌باشد.

عناصر داخل ماتریس W شامل عددهای ۰، ۱ و ۲ می‌باشد که نشان‌دهنده‌ی تعداد ال‌های جهش‌یافته در هر جایگاه از نشانگرهاست. مقدار فاکتور بیز که در واقع مبنای بررسی اثر نشانگرها می‌باشد با استفاده از فرمول زیر محاسبه می‌شود:

$$BF = \frac{P}{\frac{1-p}{\pi} - 1 - \pi}$$

¹ National center for biotechnology information (NCBI)

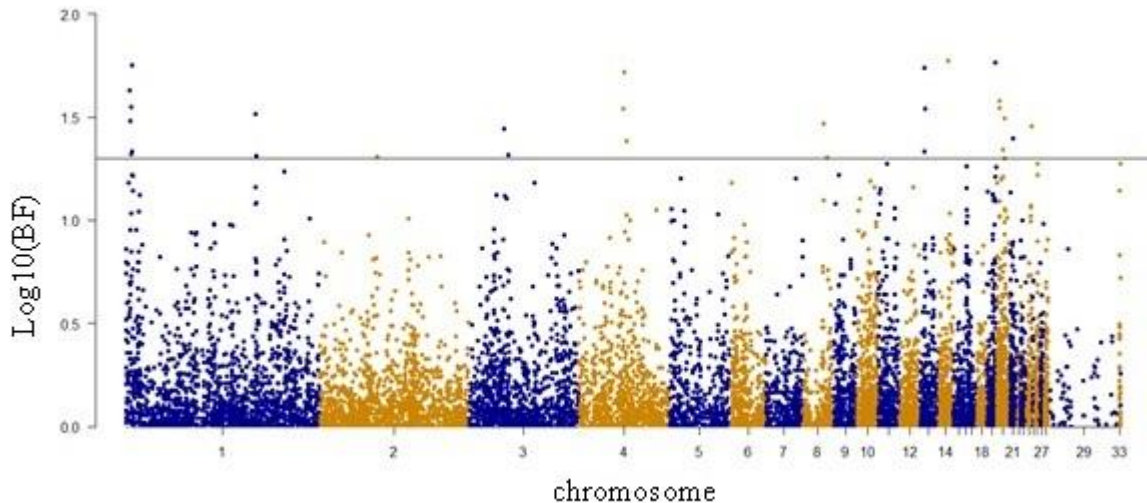
² Universal protein resource (UniProt)

که تقریباً با مناطق مشخص شده در پژوهش حاضر همخوانی قابل توجهی را نشان داد. مناطقی که همپوشانی را نشان دادند شامل کروموزوم‌های ۳، ۴، ۷، ۱۳، ۱۹، ۲۱، ۲۲، ۲۴ و ۲۶ بودند (Biscarini et al. 2010). به‌گونه‌ای که اسنپ‌های قابل توجه بر روی کروموزوم شماره‌ی ۴ در فاصله‌ی ۲/۷-۵۳/۴ Mb قرار گرفته و نشانگرهای مشخص شده بر روی کروموزوم ۴ در این پژوهش هم هرچند در دامنه‌ی فاصله‌ی کمتری قرار دارد (Mb ۸۱-۴۷) ولی با این وجود SNP‌های قابل توجهی در این مناطق برای این کروموزوم به‌دست آورده شده‌است.

به‌دلیل اینکه در شجره‌ی جمعیت مورد استفاده فنوتیپ‌ها فقط برای نسل F2 ثبت شده بود و به عبارتی شجره دارای عمق کافی نبود لذا علت اصلی پایین بودن برآورد واریانس ژنتیکی و در نهایت وراثت‌پذیری محاسبه شده را می‌توان به این امر مرتبط دانست. با قرار دادن پیش برآوردهای به‌دست آمده از آنالیز داده‌ها به‌وسیله نرم‌افزار DMU در آنالیز رگرسیونی بیزی BayesCπ نقاط ژنومی مؤثر بر روی صفت مورد نظر شناسایی شد. در سطح بیز فاکتور بالاتر از ۲۰، تعداد ۲۸ SNP به‌دست آمد (شکل ۱) در مطالعه‌ی پویش ژنومی که برای پاسخ ایمنی ذاتی و اکتسابی بر روی مرغان تخم‌گذار انجام شد ۵۹ نشانگر معنی‌دار مشخص شد

جدول ۱- اسامی نشانگر SNPها و نزدیک‌ترین ژن‌ها به آن‌ها برای شاخص ایمنی سلولی

SNP ID	Chromosome	Position	BF	Nearest gene	distance	Function
Gga_rs13940125	1	135642332	20.54	NCK2	193397 ^D	Direct effect on immune response
Gga_rs14715931	1	6644324	56.34	CAMK1D	28067 ^D	
Gga_rs15190100	1	6085128	20.93	CAMK1D	304108 ^U	
Gga_rs314887051	1	6443169	21.59	CAMK1D	inside	Indirect effect on immune response
Gga_rs14650175	8	22057133	29.43	MKNK1	165159 ^D	
Gga_rs312320267	19	8549169	57.98	DUSP14	inside	
Gga_rs15173871	20	6293769	22.05	SULF2	44538 ^U	
Gga_rs315982277	22	3994230	22.77	LRRTM4	inside	
Gga_rs16690735	1	3972881	42.74	SFMBT2	inside	
Gga_rs14790910	1	5712839	35.50	MIR1626	65653 ^U	
Gga_rs317722455	1	4957770	30.41	GOLGB1	inside	
Gga_rs14189267	2	59976935	20.29	ID4	43357 ^D	
Gga_rs13723786	3	36068925	27.81	OPN3	127227 ^D	
Gga_rs80598487	3	38873725	20.62	KCNK1	33436 ^D	
Gga_rs16400395	4	46791780	52.43	MIR1804	69858 ^D	
Gga_rs314689872	4	46602965	34.64	FGF5	161641 ^U	
Gga_rs16405303	4	49847226	24.17	ERII	3037 ^D	
Gga_rs14653191	8	24881496	20.03	CPT2	163040 ^D	
Gga_rs15012011	14	11096818	59.38	RBFOX1	inside	
Gga_rs14270598	20	2973501	34.90	PTPRT	inside	
Gga_rs312910067	20	3202775	37.78	PTPRT	inside	
Gga_rs16169293	20	8207933	31.13	ADRM1	185676 ^D	
Gga_rs318064141	21	3241843	24.91	SLC2A5	inside	
Gga_rs15224353	24	5255843	28.64	ZPRI	152008 ^D	
Gga_rs10725395	1	134720474	32.60	LOC101750912	90424 ^U	Unknown function
Gga_rs14053468	13	5188463	21.46	LOC101750661	230069 ^D	
Gga_rs317718030	13	5838653	34.67	LOC107054490	inside	
Gga_rs314739725	13	5310319	54.96	LOC107054491	inside	



شکل ۱- نمودار منهن برای پاسخ صفت ایمنی سلولی به تزریق دی نیتروکلرو بنزن با روش Bayes Cπ. محور y مقادیر $\log_{10}(BF)$ و محور x نشان‌دهنده توزیع SNPها در سطح ژنوم است که با شماره کروموزمها مشخص شده است. آستانه معنی داری در سطح BF برابر با ۲۰ قرار داده شده است.

نمایش اولیه از اتصال آنتی ژن خاص از سلولهای T تشکیل ارتباط خاصی را به نام سیناپس ایمنولوژیکی می‌دهد، سیناپس ایمنی بر اساس یک الگوی خاصی از گیرنده‌های تفکیک شده با گیرنده‌های آنتی ژن سلولهای T است که توسط مولکولهای خانواده‌ی اینترگرین احاطه شده است (Monks et al. 1998). تشکیل سیناپس ایمنولوژیکی به سلولهای T اجازه می‌دهد به صورت بالقوه لیگاندهای آنتی ژنیک را تشخیص دهند. نقشی که ژنهای *NCK2* و *LRRTM4* در ایجاد ایمنی سلولی دارند بسیار مهم و قابل توجه است چراکه مهم‌ترین گام در ایجاد ایمنی اتصال سلولهای T به آنتی ژن و به این طریق از بین بردن آن می‌باشد. ژن *CAMK1D* در ارتباط با سه SNP معنی دار بر روی کروموزوم یک قرار دارد که تفرق و تکامل نئوتروفیلها را تنظیم می‌کند (Gaines et al. 2008). مهار این ژن در انسان باعث کاهش فاگوسیتوز، مهار تولید ¹ROS، اختلال در مهاجرت نئوتروفیلها و پیوندهای بین مولکولی حاصل از اینترگرین بتا ۲ شده است (Verploegen et al. 2005). نئوتروفیلهای فعال شده، سیتوکینها و کموکینهایی را تولید می‌کنند که این سلولها به نوبه‌ی خود پاسخ التهابی سایر سلولها مثل ماکروفاژها و سایر نئوتروفیلها را تنظیم می‌کنند.

(Dorshorst et al. 2011). در مطالعه بر روی یک جمعیت F2 از لگورن، ۱۱ جایگاه موثر بر پاسخ آنتی‌بادی علیه SRBC که بر روی کروموزومهای ۱، ۲، ۳، ۵، ۷، ۱۱، ۱۶ و ۲۴ قرار داشتند را شناسایی کردند. در این پژوهش نشانگرهای SNP قابل توجهی را برای پاسخ آنتی‌بادی بر روی کروموزوم شماره‌ی یک در فاصله‌ی ۴ Mb/۱۳۶-۴/۸ به دست آوردند. همپوشانی نتایج حاصل از این مطالعه و پژوهش حاضر برای کروموزوم یک که در فاصله‌ی ۷ Mb/۱۳۵-۳ نشانگرهای قابل توجهی را نشان داد و همچنین همپوشانی در دیگر جایگاه‌های معرفی شده (کروموزوم ۲، ۷ و ۲۴) می‌تواند بر وجود یک QTL در منطقه‌ی Mb/۱۳۶-۴/۳ بر کروموزوم یک و همچنین نقش دیگر کروموزومها در کنترل صفات ایمنی را نشان بدهد. اسامی نشانگرهای SNP معنی دار در روش بی‌زین همراه با نزدیک‌ترین ژنهای آنها در جدول ۱ گزارش شده است.

همان‌طور که قبلاً گفته شد اهمیت اجزای سیستم ایمنی سلولی از جمله سلولهای T در ایجاد پاسخ ایمنی مناسب با استفاده از هدایت سیگنال و در نتیجه مقاوم کردن سلولها در برابر آنتی‌ژنهای خارجی بسیار مهم می‌باشد. عملکرد ژنهای *NCK2* و *LRRTM4* که به ترتیب بر روی کروموزوم ۱ و ۲۲ قرار دارند، تشکیل سیناپس ایمنولوژیکی، هدایت سیگنال و فعال‌سازی سلولهای T می‌باشد (Siddiqui et al. 2013; Gil et al. 2002)

¹ Reactive Oxygen Species (ROS)

است. این دسته از RNAها جزئی از سیستم ایمنی ژنتیکی هستند و در مقابل سدهای نوکلوتیک و ویروسی مقابله می‌کند (Thomas et al. 2014). ژن *CPT2* که روی کروموزوم ۸ قرار دارد در بتا اکسیداسیون اسیدهای چرب که بخشی از متابولیسم اسیدهای چرب است، دخالت دارد و همچنین پروتیین‌هایی که توسط این ژن کد می‌شوند تماما در بتا اکسیداسیون اسیدهای چرب و متابولیسم لیپید و همچنین مسیر کارنتین شاتل دخالت دارد. از طرف دیگر سنتز اسیدهای چرب در طی فرآیندهای التهابی باعث تولید سیتوکین‌ها و در نتیجه پاسخ ایمنی و التهاب می‌شود و اکسیداسیون این اسیدهای چرب در هنگام مواجهه مجدد عامل بیماری‌زا با سلول‌های T خاطره موجب سرکوب شدن پاسخ ایمنی و افزایش طول عمر سلول می‌شود (O'Neill et al. 2016). بیان ژن *RBFOX1* که روی کروموزوم ۱۴ قرار دارد، باعث تنظیم قطعات جایگزین از آگزون‌های بافت‌های خاصی به هم می‌شود. نشان داده شده‌است که این ژن بیش‌تر در طول فرآیند اریتروپویزیس (تولید گلبول قرمز خون) فعالیت دارد (Ponthier et al. 2006). گویچه سرخ، گلبول قرمز یا اریتروسیت متداول‌ترین نوع یاخته‌های خونی است و در بدن مهره‌داران، نقش اصلی رساندن اکسیژن را، از شش‌ها یا آبشش‌ها به بافت‌ها، ایفا می‌کند. در پژوهشی مشخص شد که سطح گلبول‌های قرمز خون را ایمنوگلوبولین‌ها پوشش می‌دهند و دیده شده‌است که بعد از تزریق خون به بیمارانی که بیماری آنمی دارند سطح پلاکت در خون آن‌ها بالا رفته است (Salama et al. 1983).

نتیجه‌گیری

در این پژوهش ۲۸ نشانگر با استفاده از روش بی‌زین، که در سطح بی‌فاکتور بین ۲۰ و ۱۵۰ قرار داشتند، که این حالت نشان‌دهنده‌ی ارتباط قوی با صفت ایمنی است، معنی‌دار بودند. اگرچه وراثت‌پذیری پایین صفات ایمنی تاحدی روش Bayes π را تحت تاثیر قرار داد، اما این روش حتی با این وجود پراکندگی اثرات نشانگر SNPها را به‌خوبی نشان داد در نتیجه می‌توان این روش را برای پویش ژنومی در صفات ایمنی و صفاتی مشابه (به لحاظ وراثت‌پذیری) مورد تایید قرار داد. تعداد نشانگرهای معنی‌دار مشخص شده ۱۱ جایگاه را برای کنترل صفت ایمنی معرفی می‌کنند. با توجه به نقش کلیدی ژن‌های ذکر شده در ایجاد

ژن‌های *MKNKI* و *SULF2* که به‌ترتیب بر روی کروموزوم‌های ۸ و ۲۰ قرار دارند به‌عنوان گیرنده‌های قوی برای فاکتورهای رشد متصل به هپارین و همچنین سیتوکین‌ها می‌باشند و باعث رسیدن و دریافت سیگنال توسط سلول‌ها و همچنین ایجاد پاسخ به استرس‌های محیطی می‌شود (Dai et al. 2005; Fukunaga and Hunter 1997). جهش در ژن *DUSP14* که بر روی کروموزوم ۱۹ قرار دارد، باعث کاهش ایمنی از طریق کاهش سنتز سلول‌های T و سیتوکینین، که سلول‌های اصلی درگیر در ایمنی هستند و همچنین غیر فعال شدن مولکول‌های سیگنالی در سلول می‌شود (Hijikata et al. 2016) از میان انواع مختلف سلول‌های T؛ سلول‌های T کشنده و سلول‌های T کمکی اهمیت بیشتری دارند. سلول‌های کشنده T، سلول‌های عفونی را از بین می‌برند و سلول‌های کمکی T، سایر سلول‌های ایمنی را فعال می‌کنند همچنین عملکرد ژن‌هایی که به صورت غیر مستقیم در کنترل صفات ایمنی دخالت دارند، بسیار مهم و قابل توجه می‌باشد. ژن *SFMBT* که بر روی کروموزوم ۱ قرار دارد، میکرو آر ان ای *miR-466a-3p* را کد می‌کند و ارتباط نزدیکی با اغلب miRNAها که در پاسخ به محرک‌های محیطی بیان می‌شوند دارد. این ژن ترجمه miRNAهایی را سبب می‌شود که در ایجاد پاسخ هایپرتونیک در مقابل استرس سلولی فعال می‌شوند و باعث حفظ و ایمنی سلول می‌شوند (Luo et al. 2014) miRNAها قطعاً ۲۰-۲۴ نوکلئوتیدی از مناطق غیر کدشونده‌ی RNA هستند.

ژن *OPN3* که بر روی کروموزوم ۳ نقشی را در پاسخ به محرک نوری در مغز به‌عهده داشته و همچنین در تشخیص محرک نور و تنظیم ریتم شبانه‌روزی دخالت دارد. اهمیت بیان ریتم شبانه‌روزی در واقع در اجزای ایمنی ذاتی است که باعث شروع التهاب می‌شود (Silver et al. 2012; Keller et al. 2009). پارامترهای کلیدی در سیستم ایمنی در ریتم شبانه‌روزی شامل سلول‌های خونساز، هومون‌ها و سیتوکین‌ها هستند (Haus and Smolensky 1999). سیتوکین‌ها دسته‌ای از پروتیین‌های محلول در آب هستند که توسط ژن‌های سلول‌های دستگاه ایمنی بیان می‌شوند. این سلول‌ها پاسخ التهابی سایر سلول‌ها مثل ماکروفاژها، سلول‌های T و سایر نئوتروفیل‌ها را تنظیم می‌کنند. ژن *ERII* که بر روی کروموزوم ۴ قرار دارد، دارای نقش تنظیم‌کنندگی در RNAi

محمودی در طراحی آزمایش و جمع‌آوری داده‌ها و همچنین از آقای مهندس علی‌اکبری به‌جهت همکاری در تجزیه و تحلیل داده‌ها و آقای اسدالهی، خانم عسگری، فرهوش، شریعتی تشکر می‌شود. از پروفسور یوست یونس از دانشگاه آرهوس دانمارک به‌جهت همکاری در تعیین ژنوتیپ‌ها و از کنسر سوم کابونترس به‌جهت در اختیار قرار دادن چپ تجاری قدردانی می‌شود.

پاسخ ایمنی سلولی، می‌توان کارایی روش BayesCπ برای پیش‌ژنومی صفاتی با وراثت‌پذیری پایین را نیز مورد تایید قرار داد.

تشکر و قدردانی

از آقای دکتر مسعودی، خانم دکتر امامقلی و آقایان کیانی، مقصودی، جوان‌روح، عمرانی، دیده‌بان، حیدری، رئیسی و

منابع

- Baneh H (2015) Comparison of accuracy of some Bayesian methods in different strategies of genomic evaluation of threshold traits. Dissertation, University of Sari Agricultural Sciences and Natural Resources. (In Farsi).
- Biscarini F, Bovenhuis H, Van Arendonk J, Parmentier H, Jungerius A, Van Der Poel J (2010) Across-line SNP association study of innate and adaptive immune response in laying hens. *Animal Genetics* 41:26-38.
- Dorshorst B, Siegel P, Ashwell C (2011). Genomic regions associated with antibody response to sheep red blood cells in the chicken. *Animal Genetics* 42:300-08
- Costa RB, Camargo GM, Diaz ID, Irano N, Dias MM, Carvalheiro R, Boligon AA, Baldi F, Oliveira HN, Tonhati H (2015) Genome-wide association study of reproductive traits in Nellore heifers using Bayesian inference. *Genetics Selection Evolution* 47:1.
- Dai Y, Yang Y, MacLeod V, Yue X, Rapraeger AC, Shriver Z, Venkataraman G, Sasisekharan R, Sanderson RD (2005) HSulf-1 and HSulf-2 are potent inhibitors of myeloma tumor growth in vivo. *Journal of Biological Chemistry* 280:40066-73.
- Fukunaga R, Hunter T (1997) MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *The EMBO Journal* 16:1921-33.
- Gaines P, Lamoureux J, Marisettey A, Chi J, Berliner N (2008) A cascade of Ca²⁺/calmodulin-dependent protein kinases regulates the differentiation and functional activation of murine neutrophils. *Experimental Hematology* 36:832-44
- Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando R (2009) Additive genetic variability and the Bayesian alphabet. *Genetics* 183:347-63.
- Gil D, Schamel WW, Montoya MA, Sánchez-Madrid F, Alarcón B (2002) Recruitment of Nck by CD3ε reveals a ligand-induced conformational change essential for T cell receptor signaling and synapse formation. *Cell* 109:901-12.
- Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011) Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* 12:1.
- Haus E, Smolensky MH (1999) Biologic rhythms in the immune system. *Chronobiology International* 16:581-622.
- Hayes B, Goddard M (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-29.
- Hijikata M, Matsushita I, Le Hang N, Thuong P, Tam D, Maeda S, Sakurada S, Cuong V, Lien L, Keicho N (2016) Influence of the polymorphism of the DUSP14 gene on the expression of immune-related genes and development of pulmonary tuberculosis. *Genes and Immunity* 17: 207-12.
- Johnson VE (2013) Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences* 110:19313-17.
- Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association* 90:773-95.
- Keller M, Mazuch J, Abraham U, Eom GD, Herzog ED, Volk HD, Kramer A, Maier B (2009). A circadian clock in macrophages controls inflammatory immune responses. *Proceedings of the National Academy of Sciences* 106:21407-12
- Lin R, Yin G (2015) Bayes factor and posterior probability: Complementary statistical evidence to p-value. *contemporary clinical trials* 44:33-35.
- Luo Y, Liu Y, Liu M, Wei J, Zhang Y, Hou J, Huang W, Wang T, Li X, He Y (2014) Sfm2 10th intron-hosted miR-466 (a/e)-3p are important epigenetic regulators of Nfat5 signaling, osmoregulation and urine concentration in mice. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1839:97-106
- Maghsoudi A (2014) Study of some immune characteristics of Iranian naïve and high productive fowls and their crosses. Dissertation, University of Taebat Modares (In Farsi).
- Monks CR, Freiberg BA, Kupfer H, Sciaky N, Kupfer A (1998). Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395: 82-86.
- O'Neill LA, Kishton RJ, Rathmell J (2016) A guide to immunometabolism for immunologists. *Nature Reviews Immunology* 16:553-565
- Peters S, Kizilkaya K, Garrick D, Fernando R, Reecy J, Weaber R, Silver G, Thomas M (2012) Bayesian genome-wide association analysis of growth and yearling

ultrasound measures of carcass traits in Brangus heifers. *Journal of Animal Science* 90:3398-409.

Ponthier JL, Schluopen C, Chen W, Lersch RA, Gee SL, Hou VC, Lo AJ, Short SA, Chasis JA, Winkelmann JC (2006) Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1 R alternative exon 16. *Journal of Biological Chemistry* 281:12468-74.

Raeesi V, Ehsani A, Torshizi R, Sargolzaei M, Masoudi A and Dideban R (2017) Genome-wide association study of cell-mediated immune response in chicken. *Journal of Animal Breeding and Genetics* doi: 10.1111. 405-411.

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273:1516-17.

Sahana G, Gulbrandtsen B, Janss L, Lund MS (2010) Comparison of association mapping methods in a complex pedigreed population. *Genetic Epidemiology* 34:455-62.

Salama A, Mueller-Eckhardt C, Kiefel V (1983) Effect of intravenous immunoglobulin in immune thrombocytopenia: Competitive inhibition of reticuloendothelial system function by sequestration of autologous red blood cells. *The Lancet* 322:193-95.

Sham PC and Purcell SM (2014) Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics* 15:335-46.

Siddiqui TJ, Tari PK, Connor SA, Zhang P, Dobie FA, She K, Kawabe H, Wang YT, Brose N, Craig AM (2013) An LRRTM4-HSPG complex mediates excitatory synapse

development on dentate gyrus granule cells. *Neuron* 79:680-95.

Silver AC, Arjona A, Walker WE, Fikrig E (2012) The circadian clock controls toll-like receptor 9-mediated innate and adaptive immunity. *Immunity* 36:251-61.

Thomas MF, Noelle D, Ansel KM (2014) Eri1: a conserved enzyme at the crossroads of multiple RNA-processing pathways. *Trends in Genetics* 30:298-307.

Van den Berg I, Fritz S, Boichard D (2013) QTL fine mapping with Bayes C (π): a simulation study. *Genetics Selection Evolution* 45:1.

Varona L (2010) Understanding the use of Bayes factor for testing candidate genes. *Journal of Animal Breeding and Genetics* 127:16-25.

Verma J, Johri T, Swain B, Ameena S (2004) Effect of graded levels of aflatoxin, ochratoxin and their combinations on the performance and immune response of broilers. *British Poultry Science* 45:512-18.

Verploegen S, Ulfman L, van Deutekom HW, van Aalst C, Honing H, Lammers J-WJ, Koenderman L, Coffey PJ (2005) Characterization of the role of CaMKI-like kinase (CKLiK) in human granulocyte function. *Blood* 106:1076-83.

Wolfsberg TG (2010) Using the NCBI map viewer to browse genomic sequence data. *Current Protocols in Bioinformatics* 1.5. 1-1.5. 25.