

شبیه‌سازی جمعیت اولیه (تاریخی) و تعیین صحت عدم تعادل پیوستگی در مطالعات شبیه‌سازی ژنومی

Simulation of Base (Historical) Population and Determining Accuracy of Linkage Disequilibrium in Genomic Simulated Studies

حسین مهربان^{۱*}، اردشیر نجاتی‌جواری^۲، سیدرضا میرایی‌آشتیانی^۳ و حسن مهربانی یگانه^۴

تاریخ پذیرش: ۹۳/۱۲/۲۶

تاریخ دریافت: ۹۲/۰۲/۳۰

چکیده

جهت بررسی صحت عدم تعادل پیوستگی در مطالعات شبیه‌سازی ژنومی، جمعیتی در دو حالت عدم تعادل پیوستگی صفر و یک در نسل اول شبیه‌سازی شد. از معیارهای اندازه مؤثر جمعیت، میزان عدم تعادل پیوستگی و واریانس بین نشانگرها جهت تعیین صحت شبیه‌سازی استفاده شد. نتایج نشان‌دهنده همبستگی بالا بین اندازه مؤثر جمعیت مشاهده شده و مورد انتظار بود. همچنین میزان عدم تعادل پیوستگی مشاهده شده و مورد انتظار بر هم منطبق شدند و چنین روندی با افزایش اندازه جمعیت مؤثر وضعیت مطلوب‌تری پیدا کرد. به علاوه میزان واریانس مشاهده شده و مورد انتظار بین نشانگرها در حالت نوترکیبی آزاد نیز بر یکدیگر منطبق شدند. با توجه به معیارهای مورد بررسی می‌توان گفت که جمعیت اولیه به درستی شبیه‌سازی شده است و می‌توان از آن برای توسعه جمعیت مرجع استفاده کرد. با توجه به نتایج به دست آمده پیشنهاد می‌شود در مطالعات شبیه‌سازی ژنومی پس از تشکیل جمعیت اولیه از معیارهای ارائه داده شده جهت بررسی صحت شبیه‌سازی استفاده شود.

واژه‌های کلیدی: صحت شبیه‌سازی، عدم تعادل پیوستگی، واریانس بین نشانگرها، اندازه مؤثر جمعیت

۱. دانشجوی دکتری ژنتیک و اصلاح نژاد دام، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج
۲. دانشیار ژنتیک و اصلاح نژاد دام، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج
۳. استاد ژنتیک و اصلاح نژاد دام، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج
۴. استادیار ژنتیک و اصلاح نژاد دام، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج

Email: hmehrban@ut.ac.ir

*: نویسنده مسوول

مقدمه

امروزه با پیشرفت‌های صورت گرفته در ژنتیک مولکولی راه برای تعیین ژنوتیپ تعداد زیادی افراد در صدها هزار جایگاه ژنی باز شده است *گودارد و هیز* (Goddard and Hayes, 2007)، که با استفاده از این اطلاعات می‌توان ارزیابی ژنومی (Genomic Evaluation) انجام داده و ارزش اصلاحی ژنومی هر فرد را در زمان تولد پیش‌بینی کرده و حیوانات مورد نظر، جهت تشکیل نسل آینده را بر این اساس انتخاب کرد که به انتخاب ژنومی (Genomic Selection) معروف است مویسن و همکاران (Meuwissen et al., 2001). جهت ارزیابی ژنومی نیاز به جامعه‌ای است که دارای اطلاعات فنوتیپی و نشانگری باشد جمعیت مرجع (Reference Population) و همچنین بین نشانگرها و ژن‌های کنترل‌کننده صفت (صفات) عدم تعادل پیوستگی وجود داشته باشد. عدم تعادل پیوستگی (Linkage Disequilibrium (LD)) به ارتباط غیرتصادفی آلل‌ها در دو جایگاه ژنی گفته می‌شود *هیز* (Hayes, 2007). بنابراین برخی از آلل‌های نشانگرها با اثر مثبت ژن‌ها، برای همه فامیل‌ها ارتباط داشته و می‌توانند جهت انتخاب بدون نیاز به وضعیت مرحله پیوستگی در هر فامیل استفاده شوند مویسن و همکاران (Meuwissen et al., 2001). عدم تعادل پیوستگی از این جنبه قابل اهمیت است که اگر در افراد انتخاب شده بین نشانگر و ژن عدم تعادل پیوستگی برقرار نباشد، انتخاب برای نشانگر منتج به بهبود ژنتیکی نخواهد شد و حتی در صورت معکوس بودن مرحله پیوستگی پاسخ ژنتیکی ممکن است منفی باشد *دی رز و همکاران* (De Roos et al., 2008). عدم تعادل پیوستگی در اثر عواملی مانند انتخاب، رانش ژنتیکی، اختلاط جمعیت‌هایی که فراوانی‌های ژنی متفاوت دارند ایجاد می‌شود *فالکنر و مکی* (Falconer and Mackay, 1996). در مطالعات شبیه‌سازی جهت ایجاد یک جمعیت در عدم تعادل پیوستگی خاص از دو روش رانش ژنتیکی برای ۵۰ نسل آمینی؛ خان سفید؛ ساعتچی؛ ویلمسن (Amini, 2009; Khansefid, 2010; Saatchi, 2010; Villumsen et al., 2009) و تعادل بین جهش و رانش ژنتیکی برای ۱۰۰۰ نسل *کالوس* و همکاران؛ مویسن و همکاران؛ *سولبرگ و همکاران* (Calus et al., 2008; Meuwissen et al., 2001; Solberg et al., 2008) استفاده می‌شود. معیارهای مختلفی جهت اندازه‌گیری عدم تعادل پیوستگی وجود دارند. دو جایگاه ژنی A و B دارای آلل‌های A₁, A₂, B₁ و B₂ را در کل یک جمعیت فرض کنید، یکی از معیارهای مهمی که جهت اندازه‌گیری عدم تعادل پیوستگی استفاده می‌شود تفاوت بین حاصل ضرب فراوانی هاپلو تیپ‌های جفتی و غیرجفتی یا D

است. اما آماره D بیش از حد به فراوانی‌های آللی افراد وابسته بوده و بویژه جهت مقایسه عدم تعادل پیوستگی در بین جفت جایگاه‌های ژنی چندگانه (یعنی در نقاط مختلف در طول ژنوم) مفید نیست و به همین جهت آماره r^2 پیشنهاد شد *هیز* (Hayes, 2007).

تاکنون مطالعات زیادی در زمینه ارزیابی و انتخاب ژنومی بر پایه شبیه‌سازی صورت گرفته است. شبیه‌سازی به دلیل پی بردن به برخی از جنبه‌های معماری صفت و ارزیابی همزمان سناریوهای متعدد دارای مزایای بسیاری است *کلارک و همکاران؛ دتویلر و همکاران؛ زانگ و همکاران* (Clark et al., 2010; Daetwyler et al., 2010; Zhang et al., 2011). از طرفی دیگر به دلیل هزینه‌های تعیین ژنوتیپ حیوانات و در نتیجه دسترسی نداشتن به داده‌های واقعی می‌توان تحقیقات بسیار خوبی در کشور براساس این گونه مطالعات انجام داد و در صورت تأمین منابع مالی با منطبق کردن نتایج شبیه‌سازی با نتایج به دست آمده از داده‌های واقعی می‌توان از عوامل پشت صحنه صفت اطلاعات بسیار مفیدی به دست آورد. اما نکته قابل توجه این است که شبیه‌سازی باید با اصول و قوانین علمی منطبق باشد در غیر این صورت استنتاج‌هایی که بر این پایه صورت می‌گیرد معتبر نخواهد بود. به علاوه کوچکترین اشتباه در برنامه‌نویسی می‌تواند به خطاهای منطقی غیر عمدی منجر شده که می‌توانند نتایج را تحت تأثیر قرار دهند. بنابراین هدف از این تحقیق تعریف معیارهایی در فرایند شبیه‌سازی جمعیت‌ها است که می‌توانند خطاهای ناخواسته در شبیه‌سازی مطالعات ژنومی را برطرف کنند.

مواد و روش‌ها

شبیه‌سازی جمعیت

یک مدل تصادفی (Stochastic) با جمعیت مؤثر (Ne) ۱۰۰ حیوان غیرخویشاوند (۵۰ رأس نر و ۵۰ رأس ماده) با آمیزش تصادفی بین نرها و ماده‌ها (و با جلوگیری از خویش‌آمیزی) برای ۱۰۰ نسل شبیه‌سازی شد و محدودیتی برای تعداد فرزندان به ازای هر حیوان اعمال نشد، اما تعداد افراد نر و ماده متولد شده در هر نسل برابر فرض شدند.

شبیه‌سازی ژنوم

ژنومی به طول ۱۰ مورگان با ۱۰ کروموزوم با طول یکسان (۱۰۰ سانتی‌مورگان) شبیه‌سازی شد. نشانگرها به صورت دوآلی (در این مطالعه از SNPها (Single Nucleotide Polymorphism)) به عنوان نشانگر استفاده شد و با تراکم ۰/۱

رابطه فوق را براساس معیار r^2 می توان بدین صورت نوشت:
(۵)

$$r_t^2 = \frac{D_t^2}{\text{freq}(A_1) * \text{freq}(A_2) * \text{freq}(B_1) * \text{freq}(B_2)}$$

$$= \frac{D_0^2(1-c)^{2t}}{\text{freq}(A_1) * \text{freq}(A_2) * \text{freq}(B_1) * \text{freq}(B_2)}$$

$$r_t^2 = r_0^2(1-c)^{2t}$$

رابطه (۵) مخصوص جمعیت های بزرگ با آمیزش تصادفی است که در آن فراوانی آلی از نسلی به نسل دیگر ثابت است. حال هر چه تعداد جمعیت مؤثر کمتر باشد میزان مشاهده شده عدم تعادل پیوستگی از مقدار مورد انتظار فاصله می گیرد.
۲- بررسی اندازه خانواده ها و جمعیت مؤثر در هر نسل چون در این شرایط اندازه خانواده ها دارای توزیع پواسن هستند بنابراین باید میانگین اندازه خانواده ها با واریانس آن برابر باشند فالکنر و مکی (Falconer and Mackay, 1996). زمانی که واریانس اندازه خانواده ها بین نرها و ماده ها متفاوت باشد جمعیت مؤثر مورد انتظار را می توان در هر نسل به صورت زیر محاسبه کرد فالکنر و مکی (Falconer and Mackay, 1996):

$$N_e = \frac{8N-4}{V_{km}+V_{kf}+4} \quad (6)$$

که N_e جمعیت مؤثر، N تعداد کل افراد جمعیت، V_{km} واریانس اندازه خانواده در ماده ها می باشد.

۳- بررسی واریانس فراوانی جایگاه های نشانگری در هر نسل با توجه به اینکه هر جایگاه نشانگری را می توان به عنوان یک جمعیت در نظر گرفت که از یک جمعیت بزرگ مشتق شده اند می توان واریانس فراوانی آلی (q) بین زیر جمعیت ها را با استفاده از فرمول زیر محاسبه کرد (Falconer & Mackay, 1996):

$$\sigma_{q_t}^2 = p_0 q_0 (1 - (1 - \Delta F)^t), \Delta F = 1/2N_e \quad (7)$$

در معادله فوق p_0 و q_0 فراوانی اولیه جمعیت، ΔF میزان همخونی در هر نسل و $\sigma_{q_t}^2$ واریانس آلی در نسل t است. کدهای برنامه در محیط برنامه نویسی ++C نوشته شد و با مترجم Visual Studio 2010 کامپایل شدند.

نتایج و بحث

جهت ارزیابی های ژنومی ابتدا باید از صحت شبیه سازی جمعیت اولیه اطمینان حاصل کرد. با توجه به تعداد محدود افراد، در اثر رانش ژنتیکی برخی از نشانگرها تثبیت شدند (شکل ۱).

سانتی مورگان (۱۰۰۰ عدد نشانگر در هر کروموزوم) بر روی ژنوم به صورت تصادفی در نظر گرفته شدند. فراوانی های اولیه آلی برای نشانگرها یکسان (۰/۵) در نظر گرفته شدند. میزان نوترکیبی براساس تابع نقشه هالدین (Haldane, 1919) منظور و معیار عدم تعادل پیوستگی در این تحقیق r^2 در نظر گرفته شد که به صورت زیر محاسبه می شود هیل و رابرتسون (Hill and Robertson, 1968):

$$r^2 = \frac{D^2}{\text{freq}(A_1) * \text{freq}(A_2) * \text{freq}(B_1) * \text{freq}(B_2)} \quad (1)$$

که $\text{freq}(A_1)$, $\text{freq}(A_2)$, $\text{freq}(B_1)$ و $\text{freq}(B_2)$ به ترتیب فراوانی های آلی A_1 , A_2 , B_1 و B_2 در جمعیت می باشند. دامنه r^2 بین صفر برای جفت جایگاه های ژنی در تعادل پیوستگی تا یک برای جفت جایگاه های ژنی در عدم تعادل پیوستگی کامل، متغیر است و D عبارت است از هیل (Hill, 1981):

$$D = \text{freq}(A_1/B_1) * \text{freq}(A_2/B_2) - \text{freq}(A_1/B_2) * \text{freq}(A_2/B_1)$$

که $\text{freq}(A_1/B_1)$ فراوانی هاپلوتیپ A_1B_1 , $\text{freq}(A_1/B_2)$ فراوانی هاپلوتیپ A_1B_2 , $\text{freq}(A_2/B_1)$ فراوانی هاپلوتیپ A_2B_1 و $\text{freq}(A_2/B_2)$ فراوانی هاپلوتیپ A_2B_2 در جمعیت می باشند. در این تحقیق جهت سنجش صحت شبیه سازی از سه روش استفاده شد:

۱- بررسی روند عدم تعادل پیوستگی در هر نسل:

برای این منظور از دو آزمون به شرح زیر استفاده شد:

۱-۱ آزمون عدم تعادل پیوستگی زمانی که میزان r^2 در نسل اول نزدیک به صفر است:

مقدار مورد انتظار عدم تعادل پیوستگی در هر نسل براساس فرمول زیر محاسبه شد (Sved, 1971):

$$E(r_n^2) = \frac{1}{1+4N_e c \left[\frac{(1-c)^n}{(1-c)^2} \right]} \left\{ 1 - \left[\left(1 - \frac{1}{2N_e} \right) (1-c)^2 \right]^n \right\} \quad (3)$$

در رابطه فوق $E(r_n^2)$ میزان مورد انتظار عدم تعادل پیوستگی در نسل n ام، N_e تعداد افراد مؤثر جمعیت و c میزان نرخ نوترکیبی براساس مورگان هستند.

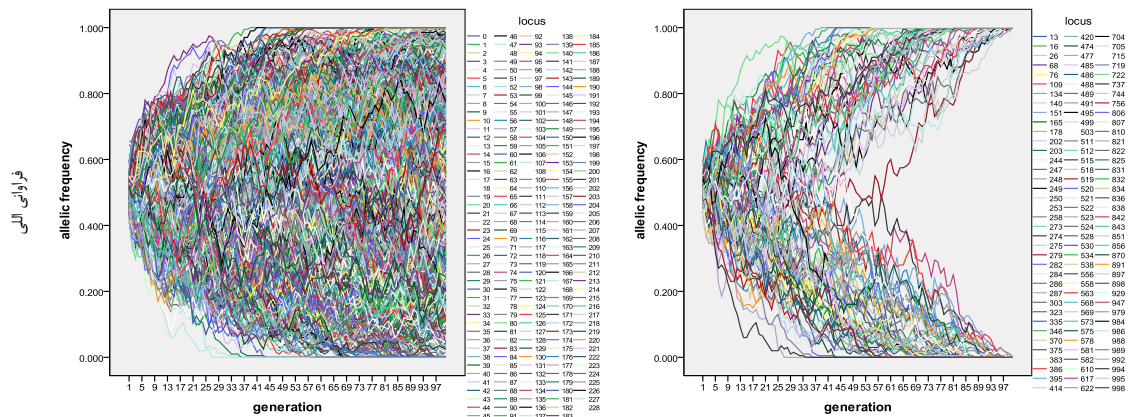
۱-۲ آزمون عدم تعادل پیوستگی زمانی که میزان r^2 در نسل اول یک است:

براساس معیار D عدم تعادل پیوستگی در نسل t و با نرخ نوترکیبی c برابر است با فالکنر و مکی (Falconer and Mackay, 1996):

$$D_t = D_0(1-c)^t \quad (4)$$

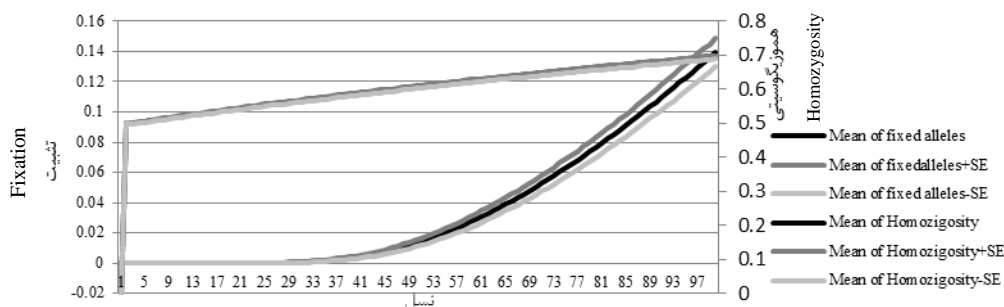
(شکل ۲). ساعتچی و همکاران (Saatchi et al., 2010) میزان هموزیگوسیتی در نسل ۵۰ را ۶۱٪ گزارش کردند. میزان هموزیگوسیتی در انتخاب ژنومی از این دیدگاه دارای اهمیت است که می‌تواند جایگزین ضریب همخونی شود ویگل (Weigel, 2011). با در دسترس داشتن ژنوتیپ حیوانات می‌توان جفت‌گیری بین آن‌ها را براساس ارزش‌های اصلاحی ژنومی آن‌ها و حداقل کردن میزان هموزیگوسیتی صورت داد. البته با توسعه مدل‌های آماری سهم غالبیت نیز در این آمیزش-ها در نظر گرفته می‌شود و با توجه به سهم واریانس غالبیت در صفت موردنظر پیشرفت ژنتیکی افزایش می‌یابد ترو و ورنه (Toro and Varona, 2010).

روند تثبیت شدن آلل‌ها حاکی از این امر بود که از حدود نسل ۴۰ میزان تثبیت شدن به‌طور چشمگیری افزایش یافت که می‌تواند نشان‌دهنده رسیدن به حدود آستانه‌ای تثبیت شدن از نسل ۴۰ به بعد باشد (شکل ۲). میزان تثبیت شدن آلل‌ها در نسل ۵۰ برابر ۱/۳۵٪ بود. ویلامسن و همکاران (Villumsen et al., 2009) میزان تثبیت شدن آلل‌ها را ۱/۶٪ گزارش کردند. کدهای برنامه به نحوی نوشته شد که در شروع شبیه‌سازی تمام جایگاه‌های ژنی، هر دو آلل را دارا بودند، به عبارت دیگر میزان هموزیگوسیتی در نسل پایه صفر بود. در نسل اول به‌علت اینکه فراوانی اولیه آلل‌ها ۰/۵ بود میزان هموزیگوسیتی به حدود ۰/۵٪ رسید و به تدریج به دلیل اندازه کوچک جمعیت یک روند خطی صعودی را طی کرد و در نسل ۵۰ به ۰/۶۰۵٪ رسید



شکل ۱: پراکندگی فراوانی‌های آلی، شکل a کل جایگاه‌های ژنی و شکل b جایگاه‌های ژنی تثبیت شده است (کروموزم یک از تکرار اول)

Fig. 1: Dispersion of allelic frequency, whole genome (a), fixed locus (b) (First chromosome from first replication)



شکل ۲: میزان تثبیت و هموزیگوسیتی نشانگرها (میانگین ۱۰۰ تکرار)

Fig. 2: Amount of fixed and homozygosis of Markers

مقدار عدم تعادل پیوستگی در طول زمان به دلیل رانش ژنتیکی افزایش یافت، اما دارای یک رشد غیرخطی از نسل ۴۰ به بعد بود (شکل ۳) که می‌تواند به دلیل افزایش میزان تثبیت شدن آلل‌ها باشد (شکل ۲).

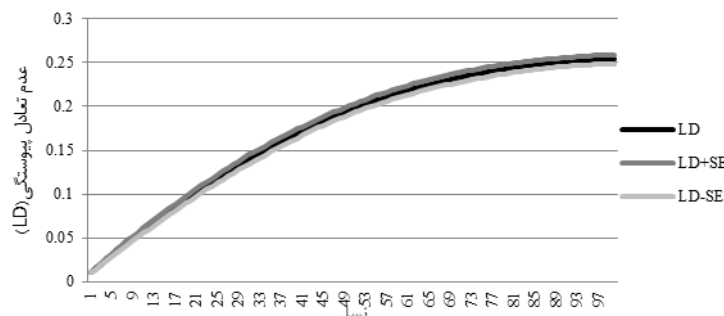
بررسی صحت شبیه‌سازی

۱- آزمون عدم تعادل پیوستگی

۱-۱- آزمون عدم تعادل پیوستگی زمانی که میزان r^2 در نسل اول نزدیک به صفر است:

همکاران (2009) این پارامتر را در تراکم ۰/۱ سانتی مورگان به- ترتیب ۰/۲ و ۰/۱۹، کالوس و همکاران (Calus et al., 2008) در فاصله متوسط بین دو نشانگر در ۰/۱۲۸ سانتی مورگان ۰/۲۱۱، خان سفید (2009) در تراکم‌های ۱ و ۰/۱ سانتی- مورگان به ترتیب ۰/۱۳۳ و ۰/۲۰۹، امینی (2009) و ساعتچی و همکاران (2010) در تراکم ۱ سانتی مورگان به ترتیب ۰/۱۸ و ۰/۱۷ گزارش کردند.

با توجه به معادله (۱) و (۲) زمانی که جایگاهی تثبیت شود میزان عدم تعادل پیوستگی این جایگاه با جایگاه قبلی و بعدی آن صفر می‌شود. از طرفی رانش تصادفی همچنان باعث افزایش عدم تعادل پیوستگی شده اما به دلیل افزایش صعودی تثبیت شدن آلل‌ها (شکل ۲) این روند از نسل ۴۰ به بعد غیرخطی بود. مقدار عدم تعادل پیوستگی در نسل ۵۰ در تراکم‌های ۱ و ۰/۱ سانتی مورگان به ترتیب ۰/۱۴۵ و ۰/۱۹۵ به دست آمد. سولبرگ و همکاران (Solberg et al., 2008) و ویلامسن و

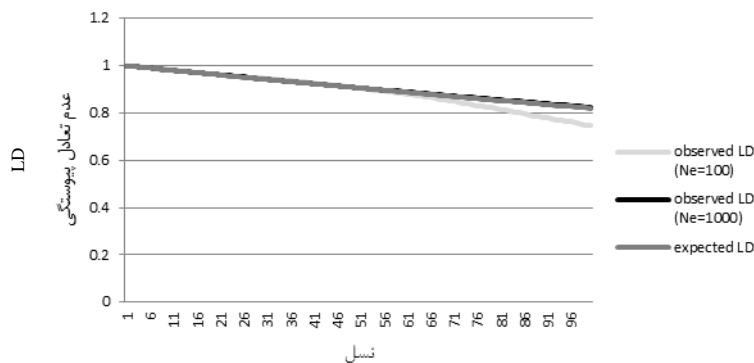


شکل ۳: میزان عدم تعادل پیوستگی ($r^2 = LD$) نشانگرها (میانگین ۱۰۰ تکرار (نرخ r^2 در شروع شبیه‌سازی صفر بود (SE: خطای استاندارد))

Fig. 3: Amount of Markers (r^2) LD (mean of 100 replications (Rate of r^2 was zero in beginning of the simulation))

در این حالت براساس معادله (۵) میزان نوترکیبی (c) برابر با ۰/۰۰۱ مورگان (۰/۱ سانتی مورگان) و تعداد افراد جمعیت مؤثر دارای دو وضعیت ۱۰۰ و ۱۰۰۰ هستند. نتایج نشان داد عدم- تعادل پیوستگی مشاهده شده و مورد انتظار بویژه زمانی که جمعیت مؤثر افزایش یافت بر یکدیگر منطبق شدند (شکل ۴) و همچنین با توجه به همبستگی بین عدم تعادل پیوستگی مشاهده شده ($Ne=100$ و $Ne=1000$) و مورد انتظار (۰/۹۸۵ و ۰/۹۹۹) می‌توان بیان کرد که جمعیت به درستی شبیه‌سازی شده است.

مقدار مورد انتظار عدم تعادل پیوستگی در نسل ۵۰ براساس معادله (۳) در تراکم‌های ۱ و ۰/۱ سانتی مورگان به ترتیب برابر با ۰/۱۳۹ و ۰/۲۱۱ است که بسیار نزدیک به مقادیر به دست آمده در این مطالعه بود (در تراکم‌های ۱ و ۰/۱ سانتی مورگان به ترتیب ۰/۱۴۵ و ۰/۱۹۵). براساس این نتایج به نظر می‌رسد مقدار عدم تعادل پیوستگی در مطالعات امینی (2009) و ساعتچی و همکاران (2010) بیش‌برآورد شده است. ۱-۲- آزمون عدم تعادل پیوستگی زمانی که میزان r^2 در نسل اول یک است:



شکل ۴: میزان عدم تعادل پیوستگی (r^2) مشاهده شده و مورد انتظار نشانگرها (میانگین ۱۰۰ تکرار (نرخ r^2 در شروع شبیه‌سازی یک بود))

Fig. 4: Amount of observed and expected LD of Markers (mean of 100 replications (Rate of r^2 was one in beginning of the simulation))

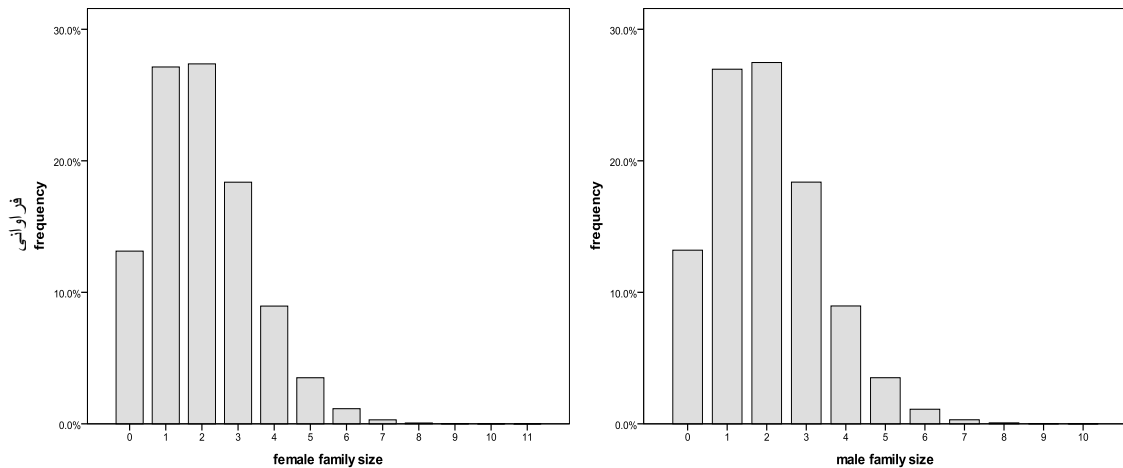
به صورت مشاهده اختلاف قابل توجه بین میزان عدم تعادل پیوستگی مورد انتظار و مشاهده شده براساس معادلات (۳) و (۵)، این تفاوت را می‌توان به میزان نوترکیبی کمتر در اثر اشتباه در برنامه‌نویسی نسبت داد. این اشتباه معمولاً در روند برنامه‌نویسی در قسمت تشکیل گامت از والدین برای تولید نسل جدید اتفاق می‌افتد.

۲- بررسی اندازه خانواده‌ها و جمعیت مؤثر در هر نسل وضعیت فراوانی اندازه خانواده‌ها بیان‌کننده توزیع پوواسن برای آنها بود (شکل ۵). در توزیع پوواسن میانگین و واریانس برابر هستند. به دلیل اینکه تعداد افراد تشکیل‌دهنده نسل‌های بعدی به نسبت مساوی از نر و ماده می‌باشد، میانگین برای هر جنس ۲ می‌باشد. می‌توان مشاهده کرد که واریانس اندازه خانواده‌ها تقریباً در هر دو جنس نزدیک به ۲ است. چنین نتایجی برای کل نسل‌ها و نسل ۵۰ نیز به دست آمد (جدول ۱). حال با توجه

به اندازه خانواده‌ها در هر نسل می‌توان جمعیت مؤثر در هر نسل را محاسبه و وضعیت صحت شبیه‌سازی را بررسی کرد. براساس معادله (۶) روند جمعیت مؤثر در نسل‌های مختلف بسیار نزدیک به جمعیت مؤثر شروع شبیه‌سازی یعنی ۱۰۰ فرد می‌باشد، که گویای درستی شبیه‌سازی ۱۰۰ نسل اولیه است (شکل ۶). این آزمون از این دیدگاه دارای اهمیت است که اگر در روند برنامه‌نویسی خطایی در آمیزش بین افراد صورت گیرد واریانس اندازه خانواده در نر و یا ماده دیگر برابر با میانگین آن نخواهد بود و اختلاف بین جمعیت مؤثر مشاهده شده و مورد انتظار افزایش می‌یابد. در چنین وضعیتی اگر جمعیت مؤثر کمتر (بیشتر) از میزان مورد انتظار آن باشد براساس معادله (۳) مقدار عدم تعادل پیوستگی در نسل ۵۰ بیشتر (کمتر) از مقدار مورد انتظار آن خواهد بود که منجر به خطاهای منطقی دیگر در ادامه برنامه می‌شود.

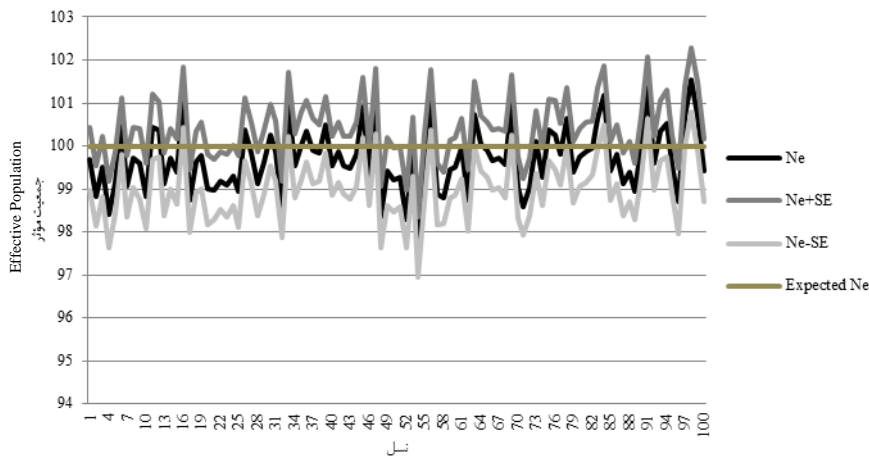
۲- بررسی اندازه خانواده‌ها و جمعیت مؤثر در هر نسل

وضعیت فراوانی اندازه خانواده‌ها بیان‌کننده توزیع پوواسن برای آنها بود (شکل ۵). در توزیع پوواسن میانگین و واریانس برابر هستند. به دلیل اینکه تعداد افراد تشکیل‌دهنده نسل‌های بعدی به نسبت مساوی از نر و ماده می‌باشد، میانگین برای هر جنس ۲ می‌باشد. می‌توان مشاهده کرد که واریانس اندازه خانواده‌ها تقریباً در هر دو جنس نزدیک به ۲ است. چنین نتایجی برای کل نسل‌ها و نسل ۵۰ نیز به دست آمد (جدول ۱). حال با توجه



شکل ۵: هیستوگرام اندازه خانواده در نرها و ماده‌ها (میانگین ۱۰۰ تکرار)

Fig. 5: Family size histogram in males and females (mean of 100 replications)



شکل ۶: روند جمعیت مؤثر (میانگین ۱۰۰ تکرار)

Fig. 6: Trend of Ne (mean of 100 replications)

جدول ۱: میانگین و واریانس اندازه خانواده (میانگین ۱۰۰ تکرار)

Table 1: Mean and variance of family size (mean of 100 replications)

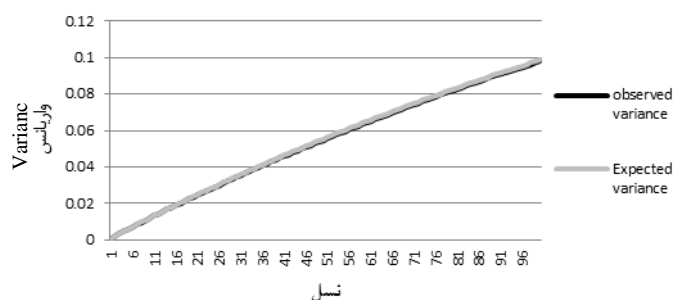
بیشینه (Maximum)	کمینه (Minimum)	واریانس (Variance)	میانگین (Mean)	
10	0	2.011(0.043)*	2	کل نسل‌ها (نرها) Total generation (Males)
11	0	2.014 (0.040)	2	کل نسل‌ها (ماده‌ها) Total generation (Females)
9	0	2.025(0.375)	2	نسل ۵۰ (نرها) 50 th generation (Males)
9	0	2.045(0.456)	2	نسل ۵۰ (ماده‌ها) 50 th generation (Females)

*Standard Error (SE)

* اعداد داخل پرانتز اشتباه معیار هستند.

واریانس فراوانی جایگاه‌های نشانگری و میزان مورد انتظار آن منطبق بر هم هستند (شکل ۷).
با توجه به نتایج به دست آمده پیشنهاد می‌شود در مطالعات شبیه‌سازی ژنومی پس از تشکیل جمعیت اولیه از معیارهای ارائه داده شده جهت بررسی صحت شبیه‌سازی استفاده شود.

۳- بررسی واریانس فراوانی جایگاه‌های نشانگری در هر نسل جهت انجام این آزمون فاصله بین هر دو نشانگر بیشتر از ۵۰ سانتی‌مورگان در نظر گرفته شد به نحوی که بین آن‌ها نوترکیبی آزاد رخ دهد چون معادله (۷) در جمعیت‌هایی است که از یکدیگر مستقل هستند. نتایج نشان دادند که میزان



شکل ۷: میزان واریانس بین نشانگرها (میانگین ۱۰۰ تکرار)

Fig. 7: Amount of within markers variance (mean of 100 replications)

مراحل اولیه شبیه‌سازی، از آن برای توسعه جمعیت مرجع استفاده کرد.

پیشنهادات

با توجه به نتایج به دست آمده پیشنهاد می‌شود در مطالعات شبیه‌سازی ژنومی پس از تشکیل جمعیت اولیه از معیارهای ارائه داده شده جهت بررسی صحت شبیه‌سازی استفاده شود.

نتیجه‌گیری

استفاده از برخی آزمون‌ها جهت بررسی صحت شبیه‌سازی می‌تواند باعث جلوگیری از برخی خطاهای ناخواسته شود. استفاده از معیارهای اندازه مؤثر جمعیت، عدم تعادل پیوستگی و واریانس بین جایگاه‌های نشانگری نشان‌دهنده صحت زیاد شبیه‌سازی جمعیت اولیه بود. بر این اساس می‌توان با اطمینان

منابع:

جهت مطالعه منابع به صفحه‌های ۱۱-۱۲ متن انگلیسی مراجعه شود.