



## استفاده از الگوریتم ژنتیک جهت نقشه یابی همزمان اثرات متقابل و چند گانه جایگاه ژنهای صفات

### کمی (QTL,S) در ژنتیک کمی

مهدی مومن<sup>۱\*</sup> خدیجه نصیری<sup>۲</sup>

۲- دانشجویان کارشناسی ارشد ژنتیک و اصلاح نژاد دانشگاه تهران- پردیس ابوریحان\* [mmomen@ut.ac.ir](mailto:mmomen@ut.ac.ir)

#### چکیده

الگوریتم ژنتیک را می توان یک روش جستجو گر نامید که نحوه جستجو و حرکت به سمت جواب بهینه را بر اساس "اصل بقای بهترین" پایه گذاری می کند. الگوریتم های ژنتیک از اصول انتخاب طبیعی داروین برای یافتن فرمول بهینه جهت پیش بینی یا تطبیق الگو استفاده می کنند. الگوریتم های ژنتیک اغلب گزینه خوبی برای تکنیک های پیش بینی بر مبنای رگرسیون هستند. همان طوریکه به رگرسیون ساده، خطی و پارامتریک گفته می شود، به الگوریتم های ژنتیک می توان غیر پارامتریک گفت. مختصراً گفته می شود که الگوریتم ژنتیک (یا GA) یک تکنیک برنامه نویسی است که از تکامل ژنتیکی به عنوان یک الگوی حل مسئله استفاده می کند. مسئله ای که باید حل شود ورودی است و راه حلها طبق یک الگو کد گذاری می شود. متریک که تابع **fitness** هم نام دارد هر راه حل کاندید را ارزیابی می کند که اکثر آنها به صورت تصادفی انتخاب می شوند. واژگان کلیدی: الگوریتم ژنتیک، اثرات متقابل و چندگانه، جایگاه ژنهای صفات کمی

#### مقدمه

الگوریتم های ژنتیک از اصول انتخاب طبیعی داروین برای یافتن فرمول بهینه جهت پیش بینی یا تطبیق الگو استفاده می کنند. سپس از یک معیار برای پیدا کردن بهترین مجموعه ضرایب و ثابت ها جهت یافتن بهترین مدل استفاده خواهیم کرد. در این روش ۲ نکته اساسی وجود دارد. اول این روش خطی است و مسئله دوم این است که ما به جای اینکه در میان "فضای پارامترها" جستجو کنیم، پارامترهای مورد استفاده را مشخص کرده ایم. با استفاده از الگوریتم های ژنتیک ما یک ابر فرمول یا طرح تنظیم می کنیم. سپس الگوریتم ژنتیک اجرا خواهد شد که بهترین تابع و متغیرها را مورد جستجو قرار می دهد. متغیرهایی که هر فرمول داده شده را مشخص می کنند به عنوان یکسری از اعداد نشان داده شده اند که معادل دی ان ای (DNA) آن فرد را تشکیل می دهند. موتور آغاز الگوریتم ژنتیک یک جمعیت از فرمول ایجاد می شود. هر فرد در برابر مجموعه ای از داده های مورد آزمایش قرار می گیرند و مناسبترین آنها شاید ۱۰ درصد از مناسبترین ها باقی می ماند. بقیه کنار گذاشته می شوند. مناسبترین افراد با هم جفتگیری (جابجایی عناصر دی ان ای) و تغییر (تغییر تصادفی عناصر DNA) کرده اند. مشاهده می شود که با گذشت از میان تعدد ریادی از نسلها، الگوریتم ژنتیک به سمت ایجاد فرمول هایی که بیشتر دقیق هستند، میل می کنند. جذابیت زیاد الگوریتم های ژنتیک این است نتایج نهایی قابل ملاحظه ترند. در الگوریتم ژنتیک (یا GA) که یک تکنیک برنامه نویسی است که از تکامل ژنتیکی به عنوان یک الگوی حل مسئله استفاده می کند. مسئله ای که باید حل شود ورودی است و راه حلها طبق یک الگو کد گذاری می شود. متریک که تابع **fitness** نام دارد هر راه حل کاندید را ارزیابی می کند که اکثر آنها به صورت تصادفی انتخاب می شوند.

#### مواد و روشها:

شناسایی چند جایگاه ژنهای صفات کمی خیلی پیوسته با استفاده از الگوریتم ژنتیک وجود جایگاه صفات کمی معمولاً با استفاده از احتمالاتی از صفات کمی مبنی بر داده های خصوصیات فنوتیپی بعلاوه نرخ نوترکیبی بین QTL و مارکر های مجاور تست می شود. با انجام این عمل احتمالات برای همه موقعیت ها (Locations) روی نقشه پیوستگی محاسبه



می شود. هنگامیکه چندین QTL مظنون و نزدیک به هم می باشند، متفاوت از عملی است که احتمالات را برای همه ترکیبات ممکن از تعداد و جایگاه های QTL محاسبه کنیم. به طور کلی GA می تواند مناسب ترین ها را توسط بهبود ژنوتیپ با تغییرات حاصل از recombination و mutation جستجو کند. ژنوتایپ ها بر مبنای تعداد و جایگاه های QTL هستند. شایستگی یک تابع بر مبنای احتمالات بعلاوه (Akaikes' Information Criterion) AIC می باشد که باعث می شود از نتایج مثبت دورغین دور شویم. نتایج نشان دادند الگوریتم ژنتیک به طور قابل اعتمادی چندین QTL قرار گرفته در فاصله تک مارکری را شناسایی می کند. در این روش ژنوتایپ ها تعداد و موقعیت QTL ها هستند و M تعداد QTL را مشخص می کند. عبارت (Cx, Lx) موقعیت xth QTL را مشخص می کند. Cx کروموزوم روی Xth, QTL موجود و Lx جایگاه اندازه گیری شده بر اساس سانتی مورگان روی Cx کروموزوم است. در تابع انتخاب و نسل های متوالی از انتخاب tournament (رقابتی) استفاده شده است. شبیه سازی (Simulation experiments): عملکرد الگوریتم ژنتیک را روی مجموعه ای از آزمایشات عددی با استفاده از جمعیت F2 اجرا شد. مدل Holdans به کار برده شد تا داده ها در مورد ژنوتایپ مارکرها را ایجاد کنیم با فرض اینکه هیچ فاصله ای بین مارکرها نباشد. واریانس باقیمانده یک و مقادیر ثابت نیز یک گذارده شد. مقادیر داده های فنوتیپی ایجاد شده به دنبال یک توزیع عددی با واریانس یک و میانگین تعیین شده توسط جمع اثرات ژنتیکی QTL و اندازه نمونه جمعیت F2, 500 نفر بود. هریک از افراد F2 یک یا سه جفت کروموزوم به طول 100CM داشتند و 11 مارکر هر 10 CM روی کروموزوم قرار گرفته بود. در این آزمایش تعداد افراد GA, 100 تا بودند و نرخ جهش 0.05 بود. در این آزمایش 3QTL روی کروموزوم های یکسانی در موقعیت های 17, 43, 85 CM از کروموزوم قرار گرفته بودند. هر QTL اثر افزایشی 1 و اثر غلبه 0 داشت. جدول (۱) همگرایی الگوریتم ژنتیک توسط 20th نسل نشان می دهد. برآوردها از 3QTL جایگاه های 16, 84, 43 CM, اثرات افزایشی 1.424, 1.358, و 1.084 داشتند که همه این رقم های پیش بینی شده نزدیک به مقادیر حقیقی بود. اثرات غلبه به طور صحیح تر و واریانس باقیمانده (residual) به طور قابل اعتمادی تخمین زده شد (1 vs. 0.977). این آزمایشات 100 بار تکرار شد. سه QTL در 96 بار به واسطه 100 شبیه سازی و چهار QTL که یکی مثبت دورغین بود در بقیه موارد شناسایی شد.

جدول ۱- شبیه سازی همگرایی الگوریتم ژنتیک در 20 نسل متوالی

	1st generation			10th generation			20th generation		
	Location (cM)	a	d	Location (cM)	a	d	Location (cM)	a	d
QTL 1	20	-0.441	0.669	15	1.602	-0.022	16	1.424	-0.014
QTL 2	98	1.182	-0.097	87	1.243	-0.036	84	1.358	-0.056
QTL 3	51	1.059	-0.109	50	1.045	-0.100	43	1.084	-0.099
QTL 4	22	1.922	-0.737	—	—	—	—	—	—
$\mu$		0.147			0.090			0.101	
$\sigma^2$		1.082			1.100			0.977	
Fitness		-218.1			-234.2			-242.8	

$\mu$ , constant value;  $\sigma^2$ , residual variance; a, additive effect; d, dominance effect.

شبیه سازی جایگاههای ژنی کاملاً به هم پیوسته (Closely Linked QTL): جدول ۲ عملکرد GA در موقعیت هایی از QTL خیلی پیوسته که مورد بررسی قرار گرفت را نشان می دهد. در این موقعیت دو QTL روی یک کروموزوم در 43 و 47 CM تعیین محل شدند بطوریکه مارکرها هر 10 سانتی مورگان قرار گرفته بودند. هیچ مارکری بین دو QTL نبود و این دو QTL اثرات ژنتیکی عکسی روی همدیگر داشتند. اثرات افزایشی و غلبه از QTL در 43 CM از 0.4 و 0.2 به 1 و 2, QTL که در 47CM بود از 0.04 و 0.02 به 2- و 1- تغییر یافت و شبیه سازی 100 بار تکرار شد. جدول ۲ نشان می دهد که برای مثال وقتی اثرات افزایشی و غلبه از QTL در 43CM, 1.6 و 0.8 بود برای QTL در 47 CM, 1.6 و 0.8- می باشد، دو QTL با اثرات ژنتیکی معکوس در 87 بار بواسطه 100 بار Trail در فاصله



مارکری 40-50 CM شناسایی شدند. در 10 بار یک QTL و در 3 بار هیچ QTL شناسایی نشد. تخمین ها از اثرات ژنتیکی کمتر از مقادیر حقیقی بود. اربب حاصل از اثرات ژنتیکی تخمین زده شده به علت همبستگی بالای بین دو QTL بدست آمد.

جدول ۲- شبیه سازی جایگاههای ژنی کاملا به هم پیوسته با الگوریتم ژنتیک

**Ability of GA to locate two very close QTL with opposite effects (100 replicates for each combination of additive and dominant effect)**

	Two QTL detected: First chromosome		One QTL detected: First chromosome		False-positive QTL
	First QTL	Second QTL	First QTL	Second QTL	
<b>True location (cM)</b>	<b>43</b>	<b>47</b>	<b>43</b>	<b>47</b>	
Location (cM)	40.14 ± 0.38	49.57 ± 0.79	45.59 ± 4.28		
Additive effect (0.4) (-0.4)	0.30 ± 0.13	-0.28 ± 0.09	0.01 ± 0.51		
Dominance effect (0.2) (-0.2)	0.22 ± 0.12	-0.19 ± 0.18	0.01 ± 0.16		
Detection		7/100	27/100		66/100
Location (cM)	40.96 ± 1.06	49.44 ± 0.71	44.19 ± 4.84		
Additive effect (0.8) (-0.8)	0.34 ± 0.12	-0.37 ± 0.09	0.06 ± 0.50		
Dominance effect (0.4) (-0.4)	0.27 ± 0.13	-0.24 ± 0.11	0.01 ± 0.19		
Detection		25/100	32/100		43/100
Location (cM)	41.29 ± 1.18	48.76 ± 1.67	45.50 ± 4.53		
Additive effect (1.2) (-1.2)	0.51 ± 0.14	-0.51 ± 0.14	-0.03 ± 0.44		
Dominance effect (0.6) (-0.6)	0.31 ± 0.16	-0.33 ± 0.19	0.00 ± 0.38		
Detection		59/100	30/100		11/100
Location (cM)	41.46 ± 1.18	48.40 ± 1.24	46.70 ± 3.08		
Additive effect (1.6) (-1.6)	0.66 ± 0.18	-0.56 ± 0.18	-0.27 ± 0.47		
Dominance effect (0.8) (-0.8)	0.35 ± 0.20	-0.35 ± 0.21	-0.13 ± 0.51		
Detection		87/100	10/100		3/100
Location (cM)	42.19 ± 1.48	47.89 ± 1.38	45.50 ± 4.55		
Additive effect (2.0) (-2.0)	0.83 ± 0.23	-0.84 ± 0.23	-0.16 ± 0.33		
Dominance effect (1.0) (-1.0)	0.52 ± 0.26	-0.52 ± 0.25	-0.25 ± 0.17		
Detection		98/100	2/100		0/100

### نتیجه گیری کلی

GA را می توان یک روش جستجوی مؤثر در فضاهای بسیار وسیع و بزرگ نامید که در نهایت منجر به جهت گیری به سمت یک جواب بهینه می شود که شاید نتوان در مدت زمان زندگی یک فرد به آن جواب بهینه دست یافت. می توان از GA در بهینه سازی طرح ریزی برنامه های انتخاب و حذف دام استفاده کرد.

از GA برای بهینه سازی هزینه خوراک دام استفاده کرد تا حداقل هزینه را مطابق با احتیاجات حیوان به مواد مغذی و استفاده مؤثر از مواد خوراکی را در بر داشته باشد. از GA می توان برای نقشه یابی همزمان اثرات متقابل چندگانه QTL و شناسایی Multiple QTL خیلی پیوسته استفاده کرد.

### منابع

۱. علیرضا، م. ۱۳۸۶. الگوریتم های ژنتیک و کاربرد های آن. انتشارات زانوس

2. Abbass, H., M, Towsey., E, Kozan., J, Van der werf. 2003. Genetic Algorithms for a Large Scale Dynamic Allocation Problem. Journal of Applied Systems Studies, 4(2).

3. Akif shaman., M. M, Cunkas., S, Inal., F, Inal., B, Coskun., U, Taskiran. 2009. Cost optimization of feed mixes .by genetic algorithms. Journal Elsevier. 965-974.

4. Alberd De Vies. 2002. Optimization of dairy heifer purchasing decisions under herd constrains with a genetic algorithm . Anim .sci.76:3118-3125

5. An Introduction to Genetic Algorithms ,Mitchell Melanie, A Bradford Book The MIT Press Cambridge ,England Fifth printing, 1999

6. Carborg, O., L, Andersson., B, kinghorn., 2000. The Use of a Genetic Algorithm for Simultaneous Mapping of Quantitative Trait Loci . Genetics Society of America. 155: 2003-2010 Multiple Interacting



## Apply the genetic algorithm (G.A) to detection of quantitative trait loci (QTL,s) in quantitative genetic

Mehdi Momen\*<sup>1</sup>, Khadije Naseri<sup>2</sup>

Master of Science Students- Livestock and Poultry Department-Tehran University-AbuRaihan

Campus . \* Corresponding E-mail address: mmomen@ut.ac.ir

### Abstract:

Genetic algorithm (G.A) can called a researcher method that fallow to find a optimized response the base on of “remind best idea”. Genetic algorithm (G.A) using the Darwin’s natural selection theory to find optimized formula to predicting the best responses. Genetic algorithms are best choice for technique prediction base on regression and are non –parametric method. Especially say that the Genetic algorithm is a programming technique that use the genetic development as a plan to solving the question. The question that must be solving is input and the solution method is the metric and fitness function that estimate the many solution method that are random.

**Keywords:** Genetic algorithms, multiple interactions, QTL, non – parametric method