

Screening of Missense SNPs in Coding Regions of COX-2 as a Key Enzyme Involved in Cancer

Jahanbakhsh-Godehkahriz S^{1*}, Naderi E¹, Mohamadkhani A²

¹ Department of Agronomy and Plant Breeding, School of Agriculture, University of Mohaghegh Ardabili, Ardabil, Iran

² Digestive and Liver Research Center, Shariati Hospital, Tehran University of Medical Science, Tehran, Iran

*Corresponding Author: Tel: 09143544213 Fax: 0451-5512204 E-mail: jahanbakhsh@uma.ac.ir

Received: 14 Jan 2013

Accepted: 6 May 2013

ABSTRACT

Background & Objectives: Non-synonymous single nucleotide polymorphism (nsSNPs) which results in disruption of protein function are used as markers in linkage and association of human proteins that might be involved in diseases and cancers.

Methods: To study the functional effect of nsSNP in cyclooxygenase-2 (COX2) amino acids, the nucleotide sequences encoding COX-2 gene in cancers were extracted from the NCBI (gi|223941909) data bank (283 cases) and analyzed by SIFT, I-Mutant 2.0, SNP and GO, PANTHER and FASTSNP servers. These servers involve programs that predict the effects of amino acid substitution on protein function, stability and missense.

Results: COX-2 is an essential enzyme for the production of pro-inflammatory prostaglandins which are relevant to cancer development and progression. The substitutions in some positions such as R228H and S428A of COX-2 in most of cancers linked to reformed protein function through disruption in enzyme active site.

Conclusion: Amino acid substitutions as a consequence of COX-2 nsSNPs have important role in human disease. Substitutions which are located in catalytic domain are important for the enzymatic function of COX-2 and associated with higher expression of COX-2.

Key words: Cancer; Cyclooxygenase-2; Protein Coding Regions; SNP

غربالگری SNP های زیان آور در ناحیهی کدکننده آنزیم کلیدی COX-2 دخیل در سرطان

سدابه جهان بخش گده کهریز^{۱*}، الناز نادری^۱، اشرف محمدخانی^۲

^۱ گروه زارعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه محقق اردبیلی، اردبیل، ایران

^۲ مرکز تحقیقات گوارش و کبد، بیمارستان شریعتی، دانشگاه علوم پزشکی تهران، تهران

*نویسنده مسئول. تلفن: ۰۹۱۴۳۵۴۴۲۱۳ فاکس: ۰۴۵۱-۵۵۱۲۲۰۴ پست الکترونیک: jahanbakhsh@uma.ac.ir

چکیده:

زمینه و هدف: امروزه، پلی مورفیسم‌های تک نوکلئوتیدی غیر مترادف (nsSNPs)، که سبب تغییر در عملکرد پروتئین‌ها می‌گردند را بعنوان مارکرهایی در جهت شناسایی پروتئین‌های درگیر در بیماری‌های انسانی و سرطان‌ها بکار می‌برند.

روش کار: در این مطالعه، به جهت بررسی تأثیر عملکردی nsSNPها در ناحیهی کدکنندهی آنزیم COX-2، توالی نوکلئوتیدی این آنزیم از طریق پایگاه اطلاعاتی NCBI با رکورد (gi|223941909) به همراه تمام SNPهای گزارش شده موجود (۲۸۳ مورد) در این پایگاه برای سرطان‌های مختلف جمع‌آوری گردید و توسط سرورهای SIFT، I-Mutant 2.0، SNP & GO، PANTHER و FASTSNP مورد آنالیز قرار گرفت. این سرورها، ابزارهایی در جهت پیش‌بینی اثر جایگزینی‌های اسید آمینه‌ی بر عملکرد، پایداری و قدرت بیماری‌زایی پروتئین‌ها به حساب می‌آیند.

یافته‌ها: COX-2 آنزیمی اساسی در سنتز پروستاگلاندین‌ها و بروز التهاب می‌باشد که نقش آن در ایجاد و گسترش انواع سرطان به اثبات رسیده است. طبق نتایج بدست آمده در این مطالعه، همه‌ی جایگزینی‌های موجود در پایگاه اطلاعاتی NCBI سبب تغییر در عملکرد و پایداری پروتئین COX-2 نمی‌گردد و برخی از جایگزینی‌های اسید آمینه‌ی شامل S428A و R228H است که با ایجاد تغییر در عملکرد و پایداری پروتئین COX-2 بطور قطع می‌توانند در ایجاد و گسترش بسیاری از انواع سرطان‌ها نقشی اساسی را ایفا کنند.

نتیجه‌گیری: جایگزینی‌های اسید آمینه‌ی COX-2 بعنوان عامل بسیار مهمی در زمینه‌ی ابتلا به بیماری‌های مختلف به ویژه سرطان‌ها به شمار می‌آیند. جایگزینی‌های غیر مترادفی که در دومن کاتالیتیک رخ می‌دهند به جهت تغییر در فعالیت آنزیم از اهمیت ویژه‌ای نسبت به سایر نواحی برخوردارند و از آن‌ها می‌توان بعنوان مارکری مستقیم برای شناسایی هرچه سریعتر زمینه‌ی ابتلا به بیماری‌های مختلف استفاده نمود.

کلمات کلیدی: سرطان، نواحی کدکننده پروتئین، COX-2، nsSNP

دریافت: ۹۱/۱۰/۲۵ پذیرش: ۹۲/۲/۱۶

مقدمه

دو ایزوفرم از آنزیم سیکلواکسیژناز (COX-2) و COX-1 وجود دارد که دارای اهمیت بیشتری هستند، هریک از این ۲ ایزوفرم، دارای ساختار، توزیع بافتی و میزان بیان خاص خود هستند. COX-1 دارای عملکرد مراقبتی در بدن می‌باشد و همیشه بیان می‌شود و در فرآیندهای فیزیولوژیکی مانند: تنظیم آزادسازی اسید معده، کارکرد کلیوی، تجمع پلاکتی و تمامیت مخاط گوارشی نقش دارد [۲]. در مقابل،

پروستاگلاندین‌ها زیر مجموعه خیلی مهم از هورمون‌های پاراکرین مشتق شده از اسیدهای چرب غیراشباع با زنجیره طویل هستند که در یکی از مراحل اولیه سنتز آن‌ها آنزیمی بنام سیکلواکسیژناز (COX) دخالت دارد. در بین بسیاری از اثرات فیزیولوژیکی پروستاگلاندین‌ها می‌توان به درد و التهاب همراه با آرتزیت اشاره کرد [۱].

۳- دومن متصل شونده به غشاء (اسید آمینه های ۵۶-۱۰۴) یک سطح هیدروفوبیک از هلیکس ها را فراهم می کند که اجازه شناور بودن در سطح لیومن و دسترسی بیشتر به سوبسترا را برای آنزیم فراهم می کند و علاوه بر این دهانه های هیدروفوبیک باریکی را برای جایگاه فعال سیکلواکسیژنازی فراهم می کند.

۴- دومن کاتالیتیکی در برگرندهی ۸۰٪ پایانی ساختار آنزیم COX-2 است و ۲ جایگاه فعال مجزا را در برمی گیرد:

جایگاه فعال پراکسیدازی: بخشی کروی است که از ۲ لوب مجزا تشکیل شده و خط اتصال این لوبها یک شکاف کم عمق در سطح بالایی آنزیم که مکان فعال پراکسیدازی است ایجاد می کند و در مجاورت قسمت غیرپروتئینی و حاوی آهن (Heme) است و یک رابطه آهن-هیستیدین را فراهم می کند که برای فعالیت پراکسیدازی و واکنش با PGG₂ و دیگر لیپید پراکسیدازها ضروری است.

جایگاه فعال سیکلواکسیژنازی: دهانه ی طولانی، باریک و بن بست که بوسیله ی هلیکس های آمفی پاتیک دومن متصل شوند به غشاء ایجاد شده، جایگاه فعال سیکلواکسیژنازی است [۳-۵].

در میان بیماری هایی که به دلایل ژنتیکی بروز می کنند، بیماری سرطان، دارای نشانه های بسیار پیچیده است و علائم بیماری تا زمان تحریک شدن با برخی محرک های متابولیکی یا محیطی، پنهان باقی می ماند. پی بردن به اهمیت جهش های سوماتیکی که در انکوژن ها رخ می دهد، سبب برنامه ریزی برای چندین طرح وسیع توالی یابی از انواع سرطان چندگانه شده است که نتایج آن اطلس ژنوم سرطان^۱ (TCGA) و پروژه ی ژنوم سرطان (CGP)^۲ را کرده است [۱]. کنسر سیوم ژنوم سرطان بین المللی (ICGC)^۳ در ارتباط با پروژه های توالی یابی سرطان

بیان ژن COX-2 بوسیله ی محرک های التهابی القاء می شود و در سلول های التهابی و گلبول سفید یافت می شود و در شرایط پاتوبیولوژیکی مانند التهاب، استرس، دژنراسیون اعصاب و مرگ برنامه ریزی شده سلول های عصبی، سرطان و متاستازها و سیتوکین ها افزایش ناگهانی پیدا می کند و به همین دلیل ژن COX-2 را ژن پاسخ فوری می نامند [۳]. به دلیل نقشی که COX-2 در ایجاد التهاب و تحریک میتوز دارد، افزایش بیان آن در تمامی بیماری های التهابی گوارشی، بیماری های عصبی نظیر اسکروز متعدد (MS)، آلزایمر و پارکینسون، بیماری های التهابی مفاصل و اکثر سرطان ها بویژه سرطان های دستگاه گوارش به اثبات رسیده است [۴]. توالی ژن COX-2 شامل ده اگزون و نه اینترون است که mRNA این ژن ۶۰۴ آمینواسید را کد می کند. آنزیم COX-2 دارای ۴ دومن با عملکرد متفاوت است.

۱- ۱۷ اسید آمینه ابتدایی ساختار آنزیم COX-2، مسئول هدایت آنزیم های تازه تولید شده به سمت محیط لیومن اندوپلاسمیک هستند. عملکرد صحیح آنزیم COX-2 نیازمند محیطی است که امکان اکسیداسیون را برای آنزیم فراهم می کند. از اینرو قرار گرفتن COX-2 در محیط لیومن به سبب فعالیت آن الزامی است چراکه میزان اکسیداسیون در محیط لیومن نسبت به سیتوزول بسیار بیشتر است.

۲- در ادامه توالی (اسید آمینه های ۵۶-۱۸ و ۱۲۸-۱۰۴)، دومن دایمریزیشن قرار دارد که توسط واکنش های هیدروفوب، پیوندهای هیدروژنی و پل های نمکی در این ناحیه سبب نگهداری ساختاری دایمری COX-2 می شود. یکی از چهار پیوند دی سولفیدی در این ناحیه این دومن را به دومن کاتالیتیکی کروی متصل می کند. حضور باندهای دی سولفیدی نیازمند یک محیط ترکیب شدن با اکسیژن است و این مطابق با مکان قرار گرفتن COX-2 (لیومن) می باشد.

¹ The Cancer Genome Atlas

² Cancer Genome Project

³ International Cancer Genome Consortium

بکار روند. SNPهایی که در نواحی کدکننده رخ می‌دهند از اهمیت بیشتری برخوردارند چراکه بر روی بیان ژن اثر گذاشته و سبب تغییر دادن عناصر تنظیمی الگوهای اتصال‌دهی، افزایش و کاهش تنظیمی Exonic Splice Enhancers (ESE) و Intronic Splice Enhancers (ISE) می‌شوند.

توانایی ایجاد شده به واسطه SNPهای غیر مترادف که عملکرد پروتئین حامل را مختل می‌کنند به دو صورت رخ می‌دهد:

۱- به واسطه تغییر مستقیم در ریشه‌های فعال (اسیدهای آمینه در جایگاه‌های فعال): برای مثال جایگزینی اسید آمینه در یک ریشه درگیر در کاتالیز، Alloestric regulation, Ligand binding و تغییرات پس از ترجمه.

۲- به واسطه تأثیر بر روی اسکلت پروتئین و ساختمان آن: برای مثال تغییر فرم و یا عدم پایداری در جایگاه اتصال و ساختار پروتئین [۱۰].

امروزه ابزارهای بیوانفورماتیکی مانند سرور SIFT^۱ قادرند ۹۰٪ از SNPهای زیان‌آور را پیش‌بینی کنند. این ابزارهای پیش‌بینی به ما در کاهش دادن میزان بررسی SNPهای کاندید از میان نواحی ژنومی بسیار بزرگ درگیر در ایجاد بیماری که سبب خسارت می‌شود کمک می‌کند [۱۱].

اهمیت ژن‌ها در نوع و عملکرد پروتئین‌هایی که بیان می‌کنند نهفته است همچنین عملکرد مولکول‌های بزرگ پروتئینی بستگی زیادی به شکل فضایی و ساختار فضایی آن‌ها دارد. دانش ساختار ۳-بعدی محصول یک ژن، نقش اصلی را در پیش‌بینی و درک عملکرد پروتئین، نقش آن در داخل سلول و نقش آن در بیماری را ایفا می‌کند و سرور هایی مانند MMDB^۲ در این زمینه بسیار کمک کننده می‌باشند. پروتئین‌هایی که در توالی آن‌ها جهش رخ داده،

در سراسر جهان همانند CGP و TCGA برای بیش از ۵۰ نوع مختلف سرطان و زیر شاخه‌های آن‌ها راه اندازی شده است [۶].

اگر این بیماری‌ها از طریق مارکرهایی سریع‌ا شناسایی و یا پیش‌بینی شوند خطر ابتلا را می‌توان با کنترل شیوه‌ی زندگی بیمار یا حامل تا حد زیادی کاهش داد و به این وسیله می‌توان شانس ابتلا به بیماری را به حداقل رساند [۷]. امروزه یکی از موضوعات جالب توجه در پزشکی بالینی، مطالعه جهش‌هایی در توالی اسید آمینه ای است که بتواند عملکرد پروتئین را تغییر دهد. در سطح مولکولی حتی تغییر یک اسید آمینه می‌تواند اهمیت داشته باشد و سبب بروز بیماری و یا نقص در فعالیت پروتئین گردد. یک اسید آمینه‌ی جهش یافته می‌تواند آرایش طبیعی پروتئین، انعطاف پذیری ساختاری، انرژی و یا حرکت مولکولی و خاصیت انتخاب کنندگی (تمایل) آنزیم‌ها نسبت به سوبسترا و مهارکننده را کاهش دهد و یا از بین ببرد. جایگاه فعال هر آنزیم در مقایسه با اندازه کلی مولکول پروتئین از نظر اندازه کوچک است. فاصله و زاویه بین ریشه‌های کاتالیتیک یک آنزیم و سوبسترا باید صحیح باشد تا کاتالیز بتواند انجام شود، به نحوی که تغییر در یک اسید آمینه با فاصله از جایگاه فعال، می‌تواند سبب تغییر فعالیت آنزیم شود. جایگاه‌های خاصی در توالی اسید آمینه وجود دارند که در بین جمعیت‌های مختلف متغیر هستند [۸].

این جایگاه‌های ترادفی خاص، زمانی که فقط در برگیرنده تغییرات مربوط به کدون اصلی یک اسید آمینه باشند پلی مورفیسم تک نوکلئوتیدی (SNP) نامیده می‌شوند که می‌توانند به فهم علت بروز پاسخ‌های متفاوت به بیماری یا درمان در جمعیت‌های انسانی کمک کند. SNPها می‌توانند بوسیله‌ی تغییر عملکرد یک ژن بطور مستقیم سبب بیماری شوند و یا بعنوان مارکرهایی در شناسایی بیماری که توسط خانواده‌ای از ژن‌ها کنترل می‌شود

¹ Sequence Independent Filtering Tools

² Molecular Modeling Data Base

توالی UniProt KB/Swiss-Pro^۶ بدست آمد. توالی نوکلئوتیدی آنزیم COX-2 از بانک اطلاعاتی NCBI با شمارهی (gi|223941909) و نیز از بانک اطلاعات توالی پروتئین Uni-Prot KB/Swiss Prot با رکورد (P05979) به دست آمد. ساختار سه بعدی پروتئین در بانک اطلاعاتی طبقه بندی ساختار پروتئین ها^۷ PDB به شماره (ID: 1PTH) جستجو شد. در این بانک اطلاعات مربوط به ساختار سه بعدی ماکرومولکول های زیستی ذخیره شده است.

ابزارهای پیش بینی اثر موتاسیون بر عملکرد پروتئین: SIFT (Sorting SIFT) (نرم افزار آنلاین intolerant from tolerant) است که پیش بینی های جایگزینی اسید آمینه ای اثر گذار بر عملکرد پروتئین را بر اساس همولوژی توالی و ویژگی های فیزیولوژیکی اسیدهای آمینه بازگو می کند. این سرور از طریق همولوژی توالی مورد نظر با ژن های مرتبط و دومن ها از میان گونه ها، امکان تماس همی ۲۰ اسید آمینه ممکن در یک جایگاه را بررسی می کند و این امکان را به کاربران می دهد که تشخیص دهند که کدام nsSNP می تواند دارای اهمیت بیشتری باشد. نتایج برای هر جایگاه بصورت های جابجایی های قابل تحمل (tolerant) و زیان آور یا غیر قابل تحمل (intolerant) در هر کدون بیان می کنند (جدول ۱). بر اساس نتایج این سرور برای تک تک کدون های آنزیم COX-2 از ۵۰ مورد SNP لیست شده در جدول ۱ تنها ۲۱ مورد با قاطعیت بالا بر روی عملکرد آنزیم COX-2 تاثیر گذارند که در مراحل بعد تنها این ۲۱ جیش مورد بررسی های دیگر قرار گرفتند.

۲- I-Mutant 2.0: بعد از گذشت چندین سال، روش های محاسباتی کامپیوتری بسیاری با استفاده از تکنیک های (Support Vector Machines (SVM).

اجازه ی تشکیل ساختار ۳- بعدی پروتئین را که در سایت PDB^۱ در دسترس است را نمی دهد [۱۲]. بنابراین روش مناسب برای تأیید موتاسیون های زیان آور قراردادن آن ها در الگوی ساختار ۳- بعدی پروتئین است. در نتیجه به جهت اهمیت بسیار آنزیم COX-2 در مطالعات بالینی، بیماری های مختلف و گزارشات متعدد در مورد آن، به غربالگری اولیه برای یافتن SNP های زیان آور در ناحیه کدکننده پروتئین، آنزیم COX-2 می پردازیم.

روش کار

جمع آوری داده ها:

موتاسیون های تک نوکلئوتیدی (SNPs) گزارش شده برای آنزیم COX-2 از طریق سایت NCBI^۲ و پایگاه اطلاعاتی SWISS-Prot^۳ بدست آمد. از ۲۸۳ مورد SNP گزارش شده در نواحی اینترون و اگزون آنزیم COX-2 تنها به بررسی دو دسته موتاسیون غیر مترادف (Non-synonymous) و مترادف (Synonymous) در نواحی کدکننده پروتئین پرداخته شد که در مجموع ۵۰ گزارش در پایگاه ها موجود بود. بر طبق تعریف، موتاسیون های غیر مترادف سبب تغییر اسید آمینه شده و مترادف ها اسید آمینه ای را تغییر نمی دهند. تمام ۵۰ مورد موتاسیون های غیر مترادف که در نواحی کدکننده COX-2 سبب تغییر اسید آمینه می شوند در جدول ۱ آمده است. اطلاعات مربوط به رابطه ی بین nsSNP ها و بیماری هایی که آنزیم COX-2 در آن ها دخیل است نیز از مطالعات *invivo* و *invitro* منتشر شده از طریق پایگاه PubMed^۴، OMIM^۵ و

¹ Protein Data Bank

² National Center for Biotechnology Information

³ <http://www.ca.expasy.org>

⁴ <http://www.ncbi.nlm.nih.gov/PubMed>

⁵ <http://www.ncbi.nlm.nih.gov/omim>

⁶ <http://www.expasy.org>

⁷ <http://www.rcsb.org/pdb>

(specific evolutionary conservation) صورت می‌گیرد. زمانیکه مقدار SubPSEC برابر صفر باشد جایگزینی با عملکرد طبیعی پروتئین همراه است و هرچه میزان آن منفی‌تر باشد امکان زیان‌آور بودن جایگزینی بیشتر می‌شود. حداکثر میزان SubPSEC برای زیان‌آور بودن ۳- است که قطعاً زیان‌آور است. ۳- FASTSNP: این سرور توانایی پیش‌بینی nsSNPهای زیان‌آور و SNPها در نواحی غیر ترجمه‌شوندگی را داراست. جستجو در این سرور با نام پروتئین آغاز می‌شود. ابزار FASTSNP در طبقه‌بندی و اولویت‌بخشیدن به خطرات فنوتیپی و اثرات زیان‌آور SNPها براساس تأثیر آنها بر روی تعیین ساختار پروتئینی، mRNA اولیه/ اتصال یافتن، انحراف در سطح رونویسی توالی، خاتمه‌ی زودرس در سطح ترجمه، انحراف در مکان‌ها در نواحی پروموتور برای اتصال فاکتور رونویسی و ... کاربرد دارد. در این مطالعه جستجو براساس Cyclooxygenase-2 صورت گرفت. ۱۷ مورد SNP برای این پروتئین که همگی ناحیه ی Coding را در می‌گیرد در جدول ۳ آمده است و اطلاعاتی در مورد SNPها در نواحی غیر ترجمه‌شوندگی آنزیم Cox-2 در اختیار ما قرار نداد.

یافته‌ها

از آنجا که، ژن COX-2 جزء رایج‌ترین ژن‌های جهش‌یافته در میان همه‌ی ژن‌های شناخته‌شده سرطان است، این جهش‌ها در کدون‌های متفاوت رخ می‌دهند و در نواحی به شدت ابقاء شده مستقر هستند. سرطان‌ها از طریق آپوپتوز، به وفور میزان کاهش‌یافته‌ی از مرگ سلولی را نشان می‌دهند و یک عامل عمده در برهم زدن فعالیت آپوپتوز آنزیم COX-2 است [۱۴]. چون تحقیقات مولکولی، پرزحمت و وقت‌گیر هستند بکار بردن ابزارهای کامپیوتری در فیلتر نمودن اولیه و اولویت‌بخشیدن

Neural Network و Decision Trees بصورت موفقیت‌آمیزی در پیش‌بینی ارتباط ساختمان-توالی بکار می‌رود. سروری که در این مطالعه بکار گرفته شد. I-Mutant 2.0 که یک تکنیک با روش SVM است و اساس پیش‌بینی آن پایداری پروتئین است بکار برده شد. فایل خروجی میزان تغییر انرژی آزاد گیبس (DDG) که بصورت زیر محاسبه می‌شود را نشان می‌دهد.

میزان انرژی آزاد گیبس غیرتاخوردگی (نوع وحشی) - میزان انرژی گیبس غیرتاخوردگی (پروتئین دارای موتاسیون) = DDG بر حسب kcal/mol

اگر میزان DDG مثبت باشد پروتئین موتاسیون یافته پایدارتر و اگر منفی باشد پایداری آن کمتر خواهد شد (انرژی آزاد گیبس مثبت باشد= پایداری) ۱- SNP & GO: برای پیش‌بینی اینکه اگر یک موتاسیون نقطه‌ای به ساختمان پروتئین راه یابد، آیا می‌تواند سبب مشارکت پروتئین حامل در ایجاد و گسترش بیماری گردد یا اینکه عملکرد پروتئین طبیعی باقی می‌ماند از سرور SNP & GO استفاده شد. این سرور براساس روش SVM که قبلاً توضیح داده شد به دسته‌بندی اثر موتاسیون می‌پردازد. اگر احتمال بیماری‌زایی بیش از ۰/۵ باشد موتاسیون بعنوان عامل بیماری شناخته می‌شود. نتایج مربوط به این روش در جدول ۲ آمده است. این سرور به جهت قطعیت بخشیدن به نتایج خود از سرورهای PhD-SNP و PANTHER نیز کمک گرفته و نتایج خود را با نتایج آنها مطابقت داده و در نهایت گزارش می‌کند.

۲- (protein analysis through PANTHER evolutionary relationships likelihood) این نرم افزار می‌تواند nsSNPهایی را که می‌تواند سبب اثرگذاری بر عملکرد پروتئین شود را پیش‌بینی کند. این سرور براساس هم‌ریدیگی تکاملی در پروتئین‌های مرتبط از طریق جایگزینی جایگاه‌های خاص تکاملی حفاظت شده (SubPSEC (substitution position-

دقیق برای امکان پیش بینی اثرات فنوتیپی قبل از ورود به آزمایشگاه را فراهم کند. تنوع ابزارهای موجود در تلفیق با مشاهدات آزمایشگاهی این امکان را به محقق می‌دهد که ابزار مورد نظر خود را انتخاب و پیش از ورود به آزمایشگاه تا حد بسیاری

به بررسی جایگزینی‌های زیان آور که می‌توانند اثرات نامشابهی بر عملکرد پروتئین داشته باشد بسیار سودمند است [۷]. علاوه براین، مطالعات کامپیوتری روش‌هایی سریعتر و به نسبت ارزانتری در مقایسه با تحقیقات

جدول ۱. SNP های گزارش شده برای آنزیم COX-2 در سایت NCBI همراه با پیش بینی سرور SIFT

rs IDs	Allele change	AA position	Tolerance index (TI)	SIFT Prediction
rs20426	G/A	M11	0.00	Intolerant
rs200140544	C/G	R4G	0.00	Intolerant
rs200906212	G/T	A13S	0.79	Tolerant
rs199946406	A/G	T17A	0.03	Intolerant
rs199737603	C/A	S23Y	0.00	Intolerant
rs201588411	G/A	C32Y	0.07	Potentially intolerant
rs199661060	C/T	R46W	0.03	Intolerant
rs189809394	G/C	V63L	0.82	Tolerant
rs199619767	A/C	I91L	0.05	Potentially intolerant
rs201046652	C/A	P92T	0.04	Intolerant
rs200132172	G/A	R95Q	0.31	Tolerant
rs145215600	A/G	N96S	0.03	Intolerant
rs200842448	T/G	S160A	0.72	Tolerant
rs199513587	A/G	K172E	0.17	Border line
rs201493026	C/A	F186L	0.05	Potentially intolerant
rs141500962	G/A	R202Q	0.27	Tolerant
rs3218622	G/A	R228H	0.02	Intolerant
rs186230736	G/A	M247I	0.58	Tolerant
rs201723686	C/T	R270W	0.01	Intolerant
rs148160346	G/A	R293Q	0.03	Intolerant
rs200204426	G/A	D300N	0.02	Intolerant
rs201228846	G/A	G310D	0.33	Tolerant
rs199520386	A/G	K328R	0.02	Intolerant
rs200843740	A/G	H337R	0.21	Tolerant
rs200479241	A/G	N361S	0.28	Tolerant
rs201708424	T/C	Y390H	0.19	Border line
rs202093833	T/G	F408L	0.09	Potentially intolerant
rs201373237	C/T	A417V	0.03	Intolerant
rs4648279	T/G	S428A	0.23	Tolerant
rs202020013	G/A	A429T	0.63	Tolerant
rs139215497	G/T	A436S	0.01	Intolerant
rs139787656	A/G	K471E	0.02	Intolerant
rs141702884	G/A	E478K	0.02	Intolerant
rs189679860	G/A	D483N	0.00	Intolerant
rs5272	A/G	E488G	0.18	Border line
rs201405348	C/T	R499W	0.01	Intolerant
rs200408009	G/T	R499L	0.02	Intolerant
rs200926368	G/A	G505S	0.90	Tolerant
rs200030471	A/T	E510D	0.14	Border line
rs5273	T/C	V511A	1.00	Tolerant
rs201342852	G/A	V524I	0.02	Intolerant
rs199820825	A/T	I525L	0.31	Tolerant
rs201922079	C/T	A529V	0.08	Potentially intolerant
rs200583134	T/C	I545T	1.00	Tolerant
rs201730487	T/A	C561S	0.13	Border line
rs201606460	A/G	S567G	0.07	Potentially intolerant
rs201772944	A/C	G578R	0.04	Intolerant
rs201772944	C/A	S586Y	0.22	Tolerant
rs3218625	G/A	G587R	0.22	Tolerant
rs201299558	G/A	R600H	0.03	Intolerant

نتایج آزمایش خود را پیش بینی نماید. نتیجه‌ی آنچه که در مورد آنزیم COX-2 در این مطالعه صورت گرفت سبب شد که بجای سردرگمی در انتخاب و

آزمایشگاهی هستند که می‌توانند یک روش بسیار

در ایجاد و گسترش بیماری‌های که آنزیم COX-2 در آن‌ها دخیل است اهمیت دارند. به کمک ابزارهای بیوانفورماتیکی و با هزینه‌ی بسیار کم از ۲۸۳ مورد SNP، ۵-۶ جایگزینی با اهمیت بیشتر بدست آمد که این جایگزینی‌ها در آن‌ها با قطعیت بسیار عملکرد آنزیم را تحت تاثیر خواهد گذاشت. یکی از موارد پیشنهادی برای این نتایج این است که یک array براساس اسیدهای آمینه در این چند جایگاه با اهمیت طراحی کرد که بتوان جایگزینی در این جایگاه‌ها را در بیماری‌هایی که COX-2 در آن‌ها دخیل است بویژه سرطان‌ها را مورد سنجش قرار داد.

بررسی یک جایگاه خاص از میان ۲۸۳ مورد SNP گزارش شده در سایت NCBI برای این آنزیم، تنها به بررسی ۵۰ مورد پردازیم چراکه هدف ما بررسی SNPهای اثرگذار گزارش شده در ناحیه‌ی کدکننده‌ی این آنزیم بود. بر اساس نتایج مطالعه حاضر، از این ۵۰ مورد نیز بر طبق نتایجی که در جداول آمده است برخی جایگزینی‌ها از جمله: M1I, R46W, R228H و R293Q از اهمیت مطالعاتی بیشتری برخوردار بوده‌اند. این جایگزینی‌ها براساس نتایج سرور SIFT برای ساختار آنزیم COX-2 قابل تحمل نبوده (جدول ۱) و بر اساس سرور I-Mutant 2 (جدول ۲) نیز پایداری آنزیم را کاهش داده و همچنین براساس نتایج سرورهای FASTSNP و GO & SNP (جدول ۲ و ۳) نیز این جایگزینی‌ها

بحث

جدول ۲. لیست اثر پیش بینی‌های سرورهای I-Mutant 2.0 و GO & SNP برای SNPهای ناحیه‌ی کدکننده‌ی COX-2

rs IDs	Allele change	AA position	I-Mutant 2.0		GO & SNPs	
			Predicted impact	DDG	Predict	Probability
rs20426	G/A	M1I	Decrease stability	0.80	Neutral	0.028
rs200140544	C/G	R4G	Decrease stability	-0.44	Neutral	0.062
rs199737603	C/A	S23Y	Decrease stability	0.03	Disease	0.711
rs199661060	C/T	R46W	Decrease stability	0.02	Disease	0.950
rs201046652	C/A	P92T	Decrease stability	-2.09	Disease	0.519
rs145215600	A/G	N96S	Decrease stability	-1.52	Neutral	0.114
rs3218622	G/A	R228H	Decrease stability	-0.43	Neutral	0.016
rs201723686	C/T	R270W	Decrease stability	-0.50	Neutral	0.376
rs148160346	G/A	R293Q	Decrease stability	-1.39	Disease	0.781
rs200204426	G/A	D300N	Decrease stability	-0.54	Neutral	0.333
rs199520386	A/G	K328R	Decrease stability	-0.88	Disease	0.628
rs202093833	T/G	F408L	Decrease stability	-1.00	Neutral	0.062
rs201373237	C/T	A417V	Decrease stability	-0.47	Disease	0.633
rs139215497	G/T	A436S	Decrease stability	-0.75	Neutral	0.223
rs139787656	A/G	K471E	Decrease stability	0.79	Disease	0.619
rs189679860	G/A	D483N	Decrease stability	-0.71	Neutral	0.127
rs201405348	C/T	R499W	Decrease stability	-1.34	Disease	0.882
rs200408009	G/T	R499L	Decrease stability	-0.39	Disease	0.588
rs201342852	G/A	V524I	Decrease stability	-1.04	Neutral	0.117
rs201772944	A/C	G578R	Decrease stability	-2.40	Disease	0.458
rs201299558	G/A	R600H	Decrease stability	-0.69	Neutral	0.327

DDG < 0: Decrease Stability DDG > 0: Increase Stability DDG: DG (New Protein) - DG (Wild Type) in Kcal/mol

جدول ۳. لیست SNPهای دارای اهمیت عملکردی برای آنزیم COX-2 که توسط سرور FASTSNP پیش بینی شده‌اند

rs	Allele change	AA position	Impact	Probability
rs 5272	E488G	Coding	Missense(non-conservative), splicing regulation	Medium-High (3-4)
rs3218625	G587R	Coding	Missense(conservative)	Low-Medium (2-3)
rs5273	V511A	Coding	Missense(conservative), splicing regulation	Low-Medium (2-3)
rs61757788	-	Coding	Sense/synonymous, splicing regulation	Low-Medium (2-3)
rs4648279	S428A	Coding	Missense(conservative)	Low-Medium (2-3)
rs5279	-	Coding	Sense/synonymous, splicing regulation	Low-Medium (2-3)
rs3218624	-	Coding	Sense/synonymous, splicing regulation	Low-Medium (2-3)
rs3218623	-	Coding	Sense/synonymous, splicing regulation	Low-Medium (2-3)
rs3218622	R228H	Coding	Missense(conservative)	Low-Medium (2-3)
rs4648265	-	Coding	Sense/synonymous, splicing regulation	Low-Medium (2-3)
rs20426	M1I	Coding	Missense(conservative)	Low-Medium (2-3)
rs5274	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)
rs4648288	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)
rs5278	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)
rs4648264	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)
rs20429	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)
rs5277	-	Coding	Sense/synonymous, splicing regulation	Very Low- Very Low (1-1)

Risk Definition: 0= No effect, 1= very low, 2= low, 3= medium, 4= high, 5= very high

$\Delta G = \text{protein unfolding energy}$

تنها حدود ۵۲K از nsSNP ها در جمعیت انسانی شناخته شده است این در حالی است که یک مجموعه ی K ۶۷-۲۰۰ از nsSNP ها در انسان مورد انتظار است [۱۴]. اگرچه بررسی تأثیرات عملکردی ایجاد شده به سبب جایگزینی SNP های غیر مترادف، با استفاده از بررسی تغییر در ساختمان ۳ بعدی پروتئین نسبت به بررسی تنها توالی قابل اطمینان تر است، اما متأسفانه ساختارهای فضایی قابل آزمایش فقط برای درصد اندکی از پروتئین ها موجود است. از اینرو تخمین زدن میزان کاهش پایداری پروتئین در طی تغییرات انرژی غیر تاخوردگی بصورت آزمایشی بسیار ساده تر و کم هزینه تر است نسبت به بررسی این تغییرات از طریق ساختار، ترجمه شدن بصورت مستقیم تحت شرایط فیزیولوژیکی طبیعی و از این طریق می توان به تغییرات بزرگ مورد انتظار که کاهش دهنده ی پایداری اند دست یافت [۱۵].

کاربردهای دیگر روش های پیش بینی جایگزینی-**های آمینو اسیدی**

- کاربرد روش های پیش بینی جایگزینی های آمینو اسیدی در بیماری های ساده مندلی
روش های پیش بینی جایگزینی های آمینو اسیدی می تواند در تشخیص دادن تغییرات غیر مترادف که سبب بیماری های ساده مندلی اند از تغییرات غیر مترادف خنثی که اثر خاصی ندارند، بکار روند. جهت تأیید صحت روش های پیش بینی جایگزینی های اسید آمینه ای، از نتایج قبلی ثبت شده در پایگاه های اطلاعاتی OMIM، HGMD و Swiss-Prot بهره می گیرند.

پیش فرض لازم این بوده است که، این جایگزینی ها در نمونه ها بر روی عملکرد پروتئین اثر گذاشته اند و در نهایت سبب بیماری شده اند (بیماری نتیجه ی این جایگزینی هاست). اکثر بیماری های نشان داده شده بوسیله ی ژن ها در این پایگاه های اطلاعاتی که از نوع

امروزه مشخص شده است که تعدادی از بیماری ها و احتمالاً اکثر سرطان ها، ناشی از اثرات مرکب انتقال عمودی جهش ها (زمینه ی ارثی) و همچنین انتقال افقی جهش های القاء شده می باشند. پیشگیری یا تشخیص اولیه و زود هنگام برای سرطان، هدف نهایی غربالگری افراد در معرض خطر است. ابزارهای پیش بینی برای سرطان های معینی می تواند مشتمل بر سبک زندگی یا رژیم غذایی، دارو درمانی، جراحی پیشگیری کننده یا غربالگری باشد. علاوه بر این، غربالگری می تواند مشتمل بر آزمون های تشخیصی باشد که به شکل غیر مستقیم فنوتیپ را آشکار می کند و به جستجوی دیگر ویژگی های بالینی که گواه بر حضور یا فقدان ژن است، می پردازد [۱۳].
تأثیر همه ی جایگزینی های اسیدهای آمینه بصورت انفرادی در عملکرد پروتئین زیان آور نیست چرا که مطالعه ی ساختمان و عملکرد پروتئین ها نشان می دهد که تمام طول توالی یک پروتئین برای عملکرد آن الزامی نیست، بلکه بخش هایی از آن از اهمیت ویژه ای برخوردارند، برای مثال با بررسی ساختار هر آنزیم (پروتئین) در می یابیم که جایگاه اتصال و واکنش در عملکرد هر آنزیم نقش اساسی دارد، از اینرو می توان استنباط کرد که این توالی های خاص در طی تکامل حفاظت شده اند (Conserved Sequence) و جهش در این جایگاه ها مرگ آور است [۱۰].
تغییرات ساختمانی ایجاد شده بعلت موتاسیون در دو زمینه می تواند ارزیابی شود:

۱- بعنوان جایگزینی های قابل اندازه گیری در ساختار سوم پروتئین به واسطه ی پیش بینی از طریق مطالعات کریستالوگرافی

۲- بعنوان تغییراتی در پایداری پروتئین، برای مثال از طریق اندازه گیری انرژی غیر تاخوردگی^۱ [۷]

$$\Delta\Delta G = \Delta G_{\text{mutant}} - \Delta G_{\text{normal-type}}$$

¹ Unfolding

استفاده از روش‌های پیش بینی جایگزینی‌های آمینو اسیدی بصورت ۱۰۰٪ می‌تواند بیمار و کنترل را از یکدیگر تفکیک کنند.

علاوه بر این ۲ ژن از طریق این پیش بینی‌ها بعنوان دارندگان بیشترین missense در مقایسه با کنترل شناسایی شدند. در حالیکه همگی ژن‌ها بصورت تجربی یک نسبت مشابهی از تغییرات missense را داشتند. فقط با بکار بردن روش‌های پیش بینی جایگزینی‌های اسیدآمینه‌ی محققین قادر بودند ۲ ژن درگیر در بیماری پارکینسون را شناسایی کنند، چرا که میزان موتاسیون در DNA میتوکنندری با افزایش سن بیشتر می‌شود. تغییرات زیان آور نادر نقش مهمی را در بیماری‌های مخصوص سنین بالا مانند پارکینسون و آلزایمر بازی می‌کند. با گذشت زمان و دوباره توالی‌یابی کردن ژنوم میتوکندریایی این پیش بینی اثبات گردید [۱۵].

- کاربرد روش‌های پیش بینی جایگزینی‌های آمینو اسیدی در مقایسات درون گونه‌ای:

مدل‌های پیش‌بینی جایگزینی‌های اسیدآمینه‌ی در بررسی پلی‌مورفیسم‌های درون گونه‌ای اغلب بکار می‌روند. همچنین در جایگزینی‌های ثابت در میان گونه‌ها نیز بکار می‌رود. مقایسات درون گونه‌ای نشان داد که در گونه‌های اهلی تعداد بیشتری از SNP های تخریب کننده نسبت به گونه‌های وحشی وجود دارند. مطالعات بیشتر بیان می‌کند که پیش بینی کردن جایگزینی‌های اسیدآمینه‌ای می‌تواند به درک و آگاهی نسبت به تفاوت‌های فنوتیپی مشاهده شده در میان گونه‌ها منجر گردد [۱۰].

- کاربرد روش‌های پیش بینی جایگزینی‌های اسیدآمینه‌ی در مطالعه‌ی موتاسیون در مقیاس وسیع:

روش‌های پیش بینی کردن جایگزینی‌های اسیدآمینه‌ای می‌تواند در مقیاس بزرگ نیز بکار رود، مثلا وقتی ژن مورد نظر ما از موتاسیون‌های زیان‌آور بسیاری اشباع شده است اندازه گیری تمام این موتاسیون‌ها

مندلی بودند، جداسازی شدند که علت ایجاد آن‌ها آسیب‌های حاصل از جایگزینی‌های انفرادی است [۸]. از طریق روش‌های پیش بینی جایگزینی‌های اسیدآمینه‌ی از میان این لیست ۷۰-۹۰٪ از این جایگزینی‌ها مخرب و زیان آور پیش بینی شدند، علاوه بر این تنها ۲۰-۱۰٪ از تغییرات خنثی از طریق این پیش بینی‌ها مخرب شناخته شد. این نشان دهنده‌ی آن است که روش‌های پیش‌بینی جایگزینی‌های اسیدآمینه‌ی می‌تواند میان جایگزینی‌هایی که سبب بیماری می‌شوند با جایگزینی‌های خنثی تفاوت قائل شود [۱۱].

اکثر بیماری‌های مندلی ساده بصورت نادر وجود دارند زیرا یک گزینش خالص سازی بصورت گزینش فوق غالبیت میان آن‌ها رخ می‌دهد و گزینش فوق غالبیت زمانی رخ می‌دهد که حامل هتروزیگوت سازگاری برتری نسبت به موتانت‌ها و هموزیگوت‌های نرمال داشته باشد.

- کاربرد روش‌های پیش‌بینی جایگزینی‌های اسیدآمینه‌ای در بیماری‌های کمپلکس

علاوه بر شناسایی SNP های غیر مترادف درگیر در بیماری‌های ساده مندلی، سودمندی روش‌های پیش‌بینی جایگزینی‌های اسیدآمینه‌ی در بیماری‌های کمپلکس مانند دیابت، Hypertension و بیماری قلبی در حال کشف و شناسایی است. بیماری‌های ژنتیکی کمپلکس، آن‌هایی هستند که تنها به علت یک مکان منفرد رخ نمی‌دهند بلکه حاصل واکنش بین لوکوس‌های چندگانه یا واکنش با محیط و یا هر دو آن‌ها هستند [۷].

برای توضیح Common disease-rare variant می‌توان به تحقیقی در مورد روش‌های پیش بینی جایگزینی‌های اسیدآمینه‌ی در بررسی نقش موتاسیون‌های میتوکندریایی در ژن‌های موثر در بیماری پارکینسون اشاره کرد. هفت ژن در ژنوم میتوکندریایی از نمونه‌های نرمال و نمونه‌های دارای بیماری پارکینسون توالی‌یابی شدند [۸]. محققین با

بیماری‌های خاص گردد و در حالت‌هایی که جایگزینی‌های قابل تحمل رخ می‌دهد می‌تواند موجب تفاوت‌های درون گونه‌ای گردد و برخی از حاملین دارنده‌ی جایگزینی خاص دارای برتری یا ضعف نسبت به سایر افراد گونه شوند [۱۳].

نتیجه گیری

جایگزینی‌های اسیدآمینه‌ی COX-2 بعنوان عامل بسیار مهمی در زمینه‌ی ابتلا به بیماری‌های مختلف به ویژه سرطان‌ها به شمار می‌آیند. جایگزینی‌های غیرمترادفی که در دومن کاتالیتیک رخ می‌دهند به جهت تغییر در فعالیت آنزیم از اهمیت ویژه‌ای نسبت به سایر نواحی برخوردارند و از آن‌ها می‌توان بعنوان مارکری مستقیم برای شناسایی هر چه سریعتر زمینه‌ی ابتلا به بیماری‌های مختلف استفاده نمود.

هزینه بر و وقت گیر است. روش‌های پیش بینی کردن جایگزینی‌های اسیدآمینه‌ی می‌تواند در اولویت بخشیدن به بررسی موتاسیون‌های زیان آور مشابه از نظر اثر گذاشتن بر عملکرد پروتئین و در نهایت فنوتیپ موجود بکار رود.

در نهایت با مقایسه‌ی جمعیت‌های کنترل و بیمار برای برخی ژن‌های خاص می‌توان به جایگزینی‌های تک نوکلئوتیدی منحصر بفردی دسترسی پیدا کرد که بعنوان بیومارکرهایی برای ابتلا یا مصونیت از آن بیماری خاص بتوان از آن‌ها بهره گرفت و از این طریق می‌توان اقدامات پیشگیرانه‌ی لازم را برای حامل‌های اینگونه بیومارکرها قبل از ابتلا به بیماری در نظر گرفت [۱۱].

در جایگاه‌هایی که هیچگونه جایگزینی برای حامل بدون تغییر در عملکرد پروتئین قابل تحمل نیست، در صورت رخ دادن جایگزینی می‌تواند موجب توقف عملکرد پروتئین مربوطه گردد و سبب ایجاد نقص و

References

- 1- Dewitt DL. Prostaglandin endoperoxidase synthase-regulation of enzyme expression. *Biochimica Et Biophysica Acta- Biomembranes*. 2008 Oct; 1083: 121-134.
- 2- Copeland RA, Williams JM, Giannaras J, Nurnberg S, Covington M, Pinto Donald, et al. Mechanism of selective inhibition of the inducible isoform of prostaglandin G/H synthase. *Proceedings of the National Academy of Sciences*. 2011 Nov; 91: 11202-6.
- 3- Dewitt DL, Smith WL. Primary structure of prostaglandin G/H synthase from sheep vesicular gland determined from the complementary DNA sequence. *Proc Natl Acad Sci U S A*. 2008 Mar; 85: 1412-6.
- 4- Dong J, Dai J, Zhang M, Hu Z, Shen H. Potentially functional COX-2 -1195G>A polymorphism increases the risk of digestive system cancers: a meta-analysis. *Journal of Gastroenterology and Hepatology*. 2010 Jun; 25: 1042-50.
- 5- Cho H, Yun CW, Park WK, Kong JY, Kim KS, Park Y, et al. Modulation activity of pro-inflammatory enzyme, COX-2 and iNOS, by derivatives. *Pharmacological Research*. 2003 Jan; 49: 37-43.
- 6- Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics*. 2006 Nov; 22: 2729-34.
- 7- Schaefer Ch, Rost B. Predict impact of single amino acid change upon protein structure. Schaefer and Rost *BMC Genomics*. 2012 Jun; 13: 1-10.
- 8- Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. *Annu. Rev. Genom. Human Genet*. 2006 Jul; 7: 61-80
- 9- Bromberg Y, Rost B. Correlating protein function and stability through the analysis of single amino acid substitutions. *BMC Bioinformatics*. 2009 Aug; 10: 1-9.

- 10- Zhang Z, Miteva MA, Wang L, Alexov E. Analyzing effects of naturally occurring missense mutations. *Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine*. 2012 Feb; 805827: 1-15.
- 11- Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*. 2003 Jul; 13: 3812–3814.
- 12- Zhou H, Zhou Y. Quantifying the effect of burial of amino acid residues on protein stability. *PROTEINS: Structure, Function, and Bioinformatics*. 2004 Feb; 54:315–322.
- 13- Magesh R, George Priya Doss C. Computational methods to work as first-pass filter in deleterious SNP analysis of Alkaptonuria. *The Scientific World Journal*. 2012 Apr; 738423: 1-9.
- 14- Lin H, Wang H, Ding H, Chen YL, Li QZ. Prediction of subcellular localization of apoptosis protein using Chou's pseudo amino acid composition. *Acta Biotheoretica*. 2009 Sep; 57: 321-330.
- 15- Indran IR, Tufo G, Pervaiz S, Brenner C. Recent advances in apoptosis, mitochondria and drug resistance in cancer cells. *Biochim. Biophys. Acta*. 2011 Jun; 1807: 735–45.