

بررسی فیلوزنیک و تنوع ژنتیکی ویروس مولد کووید-۱۹ (SARS-CoV-2)

نویسنده‌گان:

معصومه جانثار^۱* و سیدمهدي سيدی^۱*

چکیده

زمینه و هدف: ظهور بیماری کووید-۱۹ توسط کرونا ویروس-2 SARS-CoV-2 با علائم سندرم حاد تنفسی در ماه دسامبر سال ۲۰۱۹ در شهر ووهان کشور چین، با شیوع بسیار سریع همراه بوده و به سرعت به مرحله همه‌گیری جهانی رسید. این مطالعه با هدف بررسی و معرفی تنوعات ژنتیکی ویروس مولد کووید-۱۹ صورت گرفته است.

روش: بررسی فیلوزنیکی و تنوع ژنتیکی ویروس‌های SARS-CoV-2 شیوع یافته در ایران انجام و با شش کشور منتخب مقایسه شد.

یافته‌ها: بررسی فیلوزنیکی ویروس SARS-CoV-2 نشان داد که تنوع‌های ژنتیکی ایجاد شده در ژنوم ویروس در حال حاضر به حدی نیست که بتواند اختلافات فیلوزنیکی قابل توجهی را بین ژنوم ویروس‌های به دست آمده از مناطق ژئوگرافیایی گوناگون (در مجموع ۱۹ کشور از قاره‌های مختلف) ایجاد کند ولی می‌تواند خوشبندی‌های متنوعی را شکل دهد. بررسی جهش‌های ویروس SARS-CoV-2 نشان داد که ویروس‌های شیوع یافته در ایران دارای شش تنوع ژنتیکی تک نوکلئوتیدی (SNPs) در ژن‌های غیرساختاری و ساختاری و یک اضافه شدن (Insertion) شش نوکلئوتیدی در ناحیه ژن RNA-dependent RNA polymerase می‌باشند. مقایسه تنوع‌های ژنتیکی ویروس‌های ایران با شش کشور منتخب مورد بررسی (هند، آمریکا، نپال، ایتالیا، اسپانیا و فرانسه) نشان از منحصر به فرد بودن تنوعات ژنتیکی ویروس‌های شیوع یافته در ایران دارد.

نتیجه‌گیری: نتایج مطالعات نشان می‌دهد که ارتباط مستقیمی بین جهش‌های رخ داده در ژنوم ویروس SARS-CoV-2 و تغییر در بیماری‌زایی ویروس وجود دارد. بنابراین می‌توان نتیجه گرفت که بررسی منظم تنوع ژنتیکی ویروس در طول زمان شیوع این بیماری همه‌گیر می‌تواند داده‌های ارزشمندی را برای ساخت واکسن و دارو در اختیار قرار دهد.

کلیدواژه‌ها: تنوع ژنتیکی، سندرم حاد تنفسی، کووید-۱۹.

مقدمه

انسان رو به رو کرد. این کروناویروس، عامل شیوع ویرانگر سندرم حاد تنفسی (SARS) در سال ۲۰۰۳-۲۰۰۲ بود. ظهور ناگهانی SARS باعث شد تا تحقیقات جدیدی برای درک مکانیسم‌های اصلی تکثیر و بیماری‌زایی اعضا این خانواده با هدف کنترل این ویروس‌ها در جهان آغاز شود(۴). شیوع SARS-CoV باعث ابتلای ۸۰۹۶ نفر و مرگ ۷۹۴ نفر شد که درصد مرگ‌ومیر این ویروس ۹/۸ درصد اعلام شد. پس از آن در سال ۲۰۱۲، MERS-CoV شیوع پیدا کرد و در کشورهای خاورمیانه، ۲۲۶۰ نفر را آلوده کرد و میزان مرگ‌ومیر ۳۵/۵ درصد را در انسان موجب شد(۵). هفت سال پس از ظهور بیماری MERS، ظهور بیماری کووید-۱۹ توسط -RNA ویروس SARS-CoV-2، در ماه دسامبر سال ۲۰۱۹ در شهر ووهان استان هویی، کشور چین با علایم سندرم حاد تنفسی و شیوع گستردگی و

کرونا ویروس‌ها (CoVs) در پستانداران و پرندگان به طور گستردگی دیده می‌شوند و به طور کلی باعث بروز بیماری‌های تنفسی، گوارشی و گاه بیماری‌های عصبی و یا هپاتیت می‌شوند. کرونا ویروس‌ها معمولاً میزبان خود را به صورت درون گونه‌ای آلوده می‌کنند و عفونت‌های حادث شده می‌توانند حاد یا مزمن باشند. عفونت‌ها معمولاً از طریق مسیر تنفسی و یا از طریق مدفع و مسیر دهان منتقل می‌شوند(۱).

از جنبه تاریخی، کروناویروس‌ها به عنوان یک خانواده ویروسی جدید در دهه ۱۹۶۰، در پی کشف چندین پاتوژن جدید تنفسی انسان، به طور رسمی شناسایی و معرفی شدند(۲). این ویروس‌ها توسط حاشیه (Fringe) مربوط به ساختارهایی در سطح خود به نام اسپایک (Spikes) شناخته می‌شوند(۳). به طور تقریبی چهار دهه بعد از شناخت اولیه این ویروس‌ها، ویروسی با مورفولوژی مشابه، جهان را با ظهور پاتوژن جدید تنفسی در

۱. اعضو کارگروه مرجعیت علمی فرهنگستان علوم پزشکی

۲. دانشیار، پژوهشگاه ملی مهندسی ژنتیک و زیست‌فناوری

بالاست ولی با این حال هنوز ارتباط مستقیم بین هر جهش و تغییر قدرت بیماریزایی ویروس مشخص نشده است. علاوه بر این، آن‌ها شواهدی را ارائه کردند که نشان می‌دهد جهش‌های ویروس SARS-CoV-2 قادر به ایجاد تغییرات قابل توجهی در قدرت بیماری‌زایی آن‌هاست^(۷). برآورده دقیق میزان، نوع و محل تنوع ژنتیکی در ژنوم ویروس برای درک تکامل ویروس‌ها و مبارزه با آن‌ها بهوینه در زمان مواجهه با یک بیماری همه‌گیر و همچنین ویروسی با خصوصیات ناشناخته بسیار مهم است. در پژوهش حاضر بررسی‌های فیلوجنتیکی و تنوع ژنتیکی بر روی ویروس‌های شیوع یافته در ایران و مقایسه آن‌ها با سایر کشورهای جهان با هدف روش نمودن اهمیت و ضرورت چنین مطالعاتی هم در حوزه درمانی و هم در حوزه تولید واکسن صورت گرفته است.

روش کار جمع‌آوری داده‌ها

توالی ژنوم دو ویروس مولد کووید-۱۹ ثبت شده در بانک ژن NCBI توسط جمهوری اسلامی ایران در تاریخ‌های ۱۳۹۹/۱/۲۲ (ویروس اول) و ۱۳۹۹/۲/۱۹ (ویروس دوم) به ترتیب با شماره دسترسی‌های MT320891 و MT447177 SARS-CoV-2 ثبت شده در بانک ژن شامل کشورهای چین، هند، ایتالیا، آمریکا، نپال، فرانسه، اسپانیا، ژاپن، هنگ‌کنگ، ویتنام، کلمبیا، بربزرگ، کره جنوبی، فلسطین اشغالی، پاکستان، یونان، پرو و آفریقای جنوبی به ترتیب با شماره MT344948، MT066156.1، MT050493.1، NC_045512.2، MT263074، MT328032، MT240479 و MT114412 از پایگاه داده NCBI به دست آمد. Virus به دست آمد. ژنوم ویروس SARS-CoV-2 به دست آمده از شهر ووهان کشور چین به عنوان رفرنس برای بررسی واگرایی ژنتیکی و تنوع ژنتیکی مورد استفاده قرار گرفت. توالی ژن‌ها و پروتئین‌های ویروس SARS-CoV-2 از پایگاه داده ViPR^(۸) به دست آمد.

آنالیز فیلوجنتیک ویروس مولد کووید-۱۹

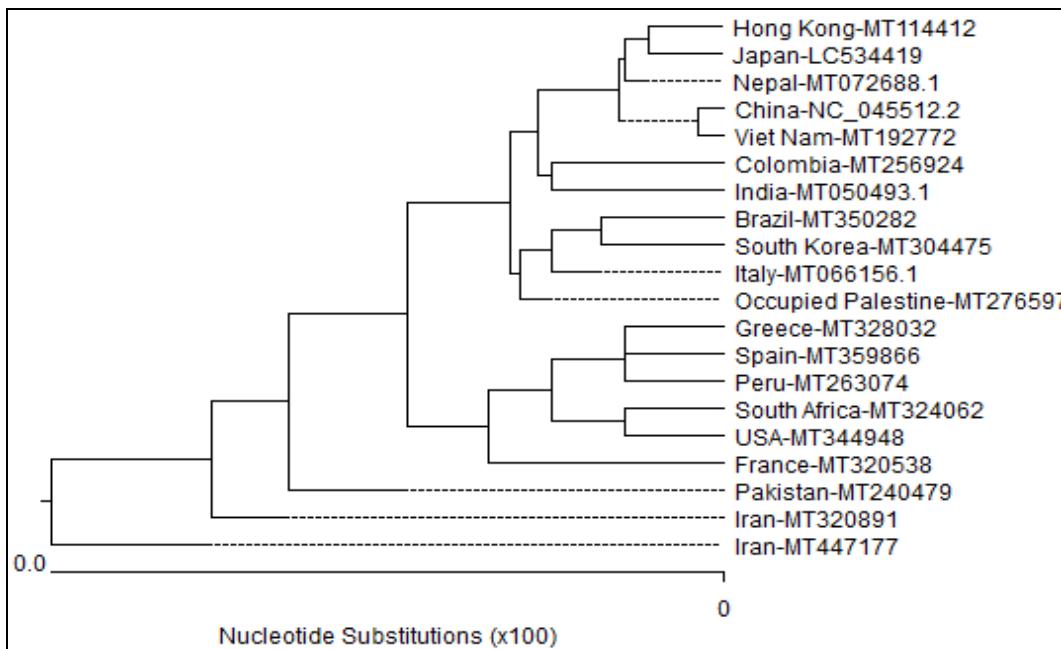
برای بررسی تنوع واگرایی ژنتیکی ژنوم‌های ویروسی به دست آمده از مناطق جغرافیایی مختلف در مقایسه با ژنوم رفرنس (ژنوم ویروس به دست آمده از شهر ووهان کشور چین)، آنالیز فیلوجنتیک انجام گرفت. درخت فیلوجنتیک با استفاده از توالی تمام ژنوم‌های ویروسی ثبت شده در بانک ژن از طریق نرم افزار DNASTAR MegAlign Version 8.1.2 ترسیم شد. برای ترسیم درخت فیلوجنتیکی مربوط به از روش ClustalW مناطق جغرافیایی مختلف فاکتور همزمانی نمونه‌های مورد استفاده برای آنالیز تا حد ممکن رعایت شد.

سریع همراه بوده و طی چند ماه با آلوده کردن ۲۱۰ کشور همراه با غافلگیر کردن کشورهای درگیر به مرحله همه‌گیری جهانی رسید. کرونا ویروس‌ها متعلق به راسته Nidovirales و خانواده Coronaviridae و زیرخانواده Alphacoronaviruses Orthocoronavirinae و Betacoronaviruses Gammacoronaviruses Deltacoronaviruses RNA می‌شوند. همه ویروس‌های راسته Nidovirales پوشش‌دار و دارای ژنوم RNA رشتہ مثبت بدون تقسیم و همه آن‌ها دارای ژنوم‌های بسیار بزرگی هستند. کرونا ویروس‌ها دارای ژنوم تک رشتہ مثبت RNA با طولی در حدود ۲۶ تا ۳۲ هزار جفت باز هستند. یکی از بارزترین ویژگی این خانواده ویروسی اندازه ژنوم آن‌ها است. کروناویروس‌ها دارای بزرگ‌ترین اندازه ژنوم در بین تمامی ویروس‌های RNA دارار هستند. ژنوم حاوی یک ساختار کلاهک در سر ۵ و یک دمپلی (A) در سر^۳ است که به آن اجازه می‌دهد تا به عنوان mRNA برای ترجمه پلی‌پروتئین‌های همانندسازی عمل کند. ژن‌های replicatease که پروتئین‌های غیرساختاری (nsps) را رمزگذاری می‌کنند، برخلاف پروتئین‌های ساختاری و جانی، که فقط حدود ۱۰ kb از ژنوم را تشکیل می‌دهند، دو سوم ژنوم، یعنی اندازه‌ای در حدود ۲۰ kb دارند. ژنوم ویروسی در انتهای^۴ leader ۳' حاوی توالی و ناحیه‌ای غیرکدکننده (UTR) است که شامل چندین ساختار ساقه - حلقه مورد نیاز برای تکثیر و رونویسی RNA است. علاوه بر این، در ابتدای هر ژن ساختاری یا جانی، توالی‌های تنظیم‌کننده رونویسی (TRSS) قرار دارند که برای بیان و نیز تنظیم بیان هر یک از این ژن‌ها ضروری هستند. UTR ۳' دارای ساختارهای RNA مورد نیاز برای تکثیر و سنترا ویروسی است. به طور کلی سازمان‌بندی ژنوم کرونا ویروس‌ها به صورت زیر است^(۱):

5'-leader-UTR-replicase-S (Spike)-E (Envelope)-M (Membrane)-N (Nucleocapsid)-3'UTR-poly (A)

کروناویروس‌ها همچنین دارای تعدادی ژن جانی (Accessory) پخش شده بین ژن‌های ساختاری هستند که پروتئین‌های حاصل از این ژن‌ها به طور تقریبی برای تکثیر در کشت بافت غیرضروری هستند. با این حال نشان داده شده است که بعضی از آن‌ها نقش مهمی در بیماری‌زایی ویروس دارند. تعداد این ژن‌ها در ویروس‌های مختلف متفاوت است^[۲]. ویروس نوظهور SARS-CoV-2 گونه جدید از بتا کرونا ویروس‌ها است که ساختار ژنومی معمول کرونا ویروس‌ها را شامل پروتئین‌های ساختاری اسپایک (S)، پوششی (E)، غشایی (M) و نوکلئوکپسید (N) و همچنین چندین پروتئین غیرساختاری و پروتئین‌های جانی منحصر به فرد را دارد می‌باشد^(۶).

دانشمندان چینی در حال حاضر برای بررسی طیف جهشی ژنوم ویروس مولد کووید-۱۹، توالی یابی بسیار عمیق ژنوم ۱۱ ویروس جدا شده از ۱۱ فرد بیمار را انجام دادند. آن‌ها ۳۳ جهش را شناسایی کردند که از میان آن‌ها ۱۹ جهش به طور کامل جدید بوده است. نتایج این مطالعه نشان داد که فراوانی SNP‌ها در ژنوم ویروس مولد کووید-۱۹



شکل ۱- درخت فیلوزنیک حاصل از آنالیز و مقایسه ژنوم ویروس SARS-CoV-2 جدا شده از ۱۹ کشور جهان که در بانک ژن NCBI ثبت شده‌اند

مهم توالی‌یابی ویروس مولد کووید-۱۹ (SARS-CoV2) می‌توان به شناسایی منشأ ویروس، ساختار ژنومی آن، و نیز اختلافاتی که با ویروس‌های منتشر شده در سایر کشورهای جهان دارد، اشاره کرد. با توجه به اینکه ساختار ژنومی ویروس‌ها در معرض دو پدیده جهش (mutation) و نوترکیبی (recombination) قرار دارد؛ لذا می‌توان انتظار داشت که شرایط محیطی کشورها و نیز ژنوم افراد در گیر با ویروس بتواند بر این دو پدیده تأثیرگذار باشد. نکته بسیار مهم از جنبه کاربردی مکان‌یابی بروز این دو پدیده از طریق مقایسه توالی با ژنوم ویروس مرجع، (که در اینجا توالی ویروس جدا شده از بیمار ساکن ووهان چین است) می‌باشد. با توجه به اینکه اهمیت تمامی ژن‌های ویروس مولد بیماری کووید-۱۹ در بیماری‌زایی یکسان نیست، لذا مکان‌یابی ژن‌های جهش یافته از جنبه بیماری‌زایی ویروس و نیز مسیرهای درمانی از اهمیت خاصی برخوردار است.

به منظور بررسی فیلوزنیک ویروس‌ها SARS-CoV-2 تعیین توالی شده از جمهوری اسلامی ایران، توالی کامل ژنوم ویروس مولد کووید-۱۹ ثبت شده از کشورمان و همه کشورهایی که تاکنون توالی این ویروس را در بانک ژن ثبت کرده‌اند که شامل چین، هند، ایتالیا، آمریکا، نپال، فرانسه، اسپانیا، ژاپن، هنگ‌کنگ، ویتنام، کلمبیا، بربزیل، کره جنوبی، فلسطین اشغالی، پاکستان، یونان، پرو و آفریقای جنوبی است، استخراج و برای رسم درخت فیلوزنیک مورد استفاده قرار گرفت که نتیجه در شکل ۱ نشان داده شده است. درخت فیلوزنیک رسم شده در شکل ۱ نشان می‌دهد که تمامی ژنوم‌های مورد استفاده در آنالیز تحت کlad SARS-CoV-2 (clade) قاره‌ای گیرند. این امر بدین معنی است که شباهت توالی ژنتیکی بین این ژنوم‌ها به حدی زیاد است که نمی‌تواند موجب

بررسی تنوع ژنتیکی ویروس کووید-۱۹

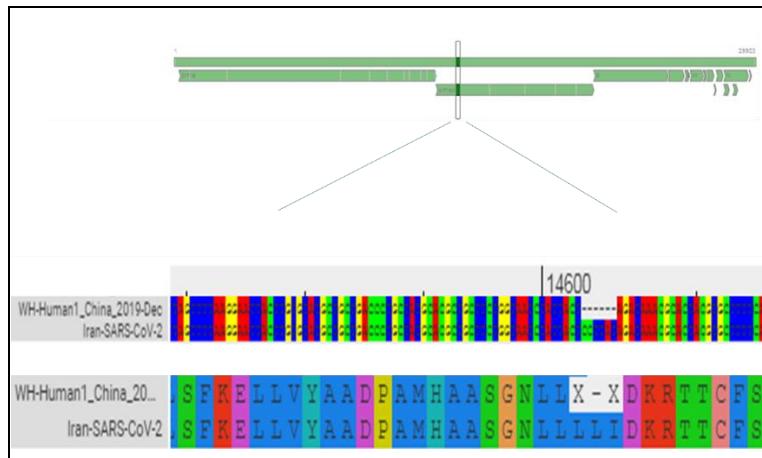
برای بررسی تنوع ژنتیکی ویروس‌های کووید-۱۹ شیوع یافته در ایران و مقایسه آن با سایر کشورهای جهان از توالی‌های ژنومی ویروس SARS-CoV-2 مربوط به کشورهای ایران، هند، ایتالیا، آمریکا، نپال، فرانسه و اسپانیا استفاده شد. ژنوم‌ها در فرمت FASTA با استفاده از نرم‌افزار Detective Coronavirus Typing Tool (version 1.1.3) (Genome قرار گرفته و جهش‌های آن‌ها در مقایسه با ژنوم ویروس به دست آمده از شهر ووهان کشور چین مورد شناسایی و ارزیابی قرار گرفت.

به منظور ارزیابی تأثیر SNP‌ها (تنوعات ژنتیکی تک نوکلئوتیدی) بر پایداری پروتئین‌های ویروسی، از دو روش پیش‌بینی پایداری پروتئین استفاده شد. در روش اول از سرور I-MUTANT server (۱۰) برای پیش‌بینی پایداری توالی پروتئین در pH=۷ و دمای ۲۵ درجه سانتی‌گراد و در روش پیش‌بینی دوم از سرور MuPro (۱۱) استفاده شد.

یافته‌ها

بررسی فیلوزنیک

با شیوع بیماری کرونا در جهان، کشورهای مختلف در گیر با این بیماری برای شناسایی بهتر ویروس و نیز در دست داشتن توالی ژنتیکی آن، اقدام به توالی‌یابی این ویروس از بیماران کشور خود کردند. بسیاری از کشورها توالی ویروس‌های خود را در بانک ژن NCBI ثبت نموده تا مورد استفاده سایر کشورها نیز قرار گیرد. برخی کشورها به توالی‌یابی ویروس جدا شده از یک فرد بستنده نکرده و اقدام به توالی‌یابی تعداد زیادی ویروس کرونا نموده‌اند که از جمله می‌توان به آمریکا و چین اشاره داشت. از دلایل



شکل ۲- هم‌ردیف‌سازی توالی نوکلئوتیدی ژنوم ویروس (ویروس اول) مولد بیماری کووید-۱۹-شیوع یافته در ایران با ویروس ووهان، جایگاه Insertion شش نوکلئوتیدی و تغییرات آمینو اسیدی احتمالی را در توالی ژنوم ویروس ایران نشان می‌دهد.

ویروس دارد (۱۷). بنابراین جهش در این ژن‌ها می‌تواند بر میزان ویرولانس ویروس مولد بیماری کووید-۱۹ تاثیر گذار باشند. علاوه بر این، nsp ۱۵ orf1ab بزرگ‌ترین ژن SARS-CoV-2 بوده که پلی‌پروتئین pp1ab RNA-dependent RNA polymerase یک پروتئین را رمزگذاری می‌کند. ویروسی با نقش آنژیمی است که در تکثیر ویروس نقش اساسی دارد. علاوه بر جهش‌های تک نوکلئوتیدی، مقایسه ژنوم ویروس اول به دست آمده از جمهوری اسلامی ایران با ژنوم ویروس ووهان، اضافه شدن شش نوکلئوتید را به ترتیب در جایگاه‌های ۴۷۸۱ و ۳۸۹ پلی‌پروتئین pp1ab و RNA-dependent RNA polymerase نشان داد (شکل ۲).

یکی از نکات بسیار مهم در خصوص بیماری کووید-۱۹، سرعت شیوع و بیماری‌زایی این ویروس است که می‌تواند به توالی ژن‌های ویروس ارتباط داشته باشد. با توجه به اینکه سرعت شیوع و بیماری‌زایی این ویروس در مناطق جغرافیایی مختلف، متفاوت گزارش شده، شاید بتوان این امر را تا حدودی با تغییرات ژنتیکی ایجاد شده در ژنوم ویروس طی شیوع بیماری مرتبط دانست. به منظور بررسی تنوع ژنتیکی وابسته به منشأ جغرافیایی، نتایج حاصل از بررسی تنوع ژنتیکی ویروس‌های مولد کووید-۱۹ با منشأ ایرانی با نتایج به دست آمده از ژنوم‌های ویروسی هند، آیتالیا، آمریکا، نپال، فرانسه و اسپانیا مورد مقایسه قرار گرفت و نتایج در جدول ۱ نشان داده شده است. در واقع این جدول فهرستی از SNP‌های توالی‌های ویروس و تغییراتی که در توالی ژن‌های هر یک و در نتیجه آمینو اسیدهای پروتئین‌ها رخ داده را براساس توالی ثبت شده ویروس در باank ژن جهانی نشان می‌دهد. نتیجه جالب توجه این است که ویروس‌ها با منشأ جغرافیایی متفاوت، دارای جهش‌های منحصر به فردی می‌باشند. مقایسه تعداد جهش‌ها رخ داده در بین ژنوم ویروس‌های مورد مطالعه (جدول ۱) نشان داد که بیشترین جهش‌ها بدون احتساب جهش‌های

جدایی فیلورزنیکی براساس مبدأ جغرافیایی آن‌ها شود، ولی گروه‌بندی‌های متنوعی بین کشورهای همسایه و غیرهمسایه قابل مشاهده است. به عنوان مثال، دو توالی متعلق به ایران و توالی پاکستان به عنوان دو کشور همسایه شاخه‌های نزدیک به یکدیگر را در درخت تشکیل داده‌اند. از طرف دیگر، آمریکا و آفریقای جنوبی به عنوان دو کشور غیرهمسایه موجود در دو قاره مجزا در یک خوشة قرار گرفته‌اند. علاوه بر این، درخت فیلورزنیک و اگرایی ژنتیکی ویروس SARS-CoV-2 منتشر شده در کشورهای مختلف نسبت به ویروس ووهان را نشان می‌دهد.

بررسی تنوع ژنتیکی

به منظور بررسی دقیق‌تر تنوع ژنتیکی ویروس، ژنوم ویروس‌های SARS-CoV-2 از جمهوری اسلامی ایران در مقایسه با ویروس ووهان (کشور چین) آنالیز شد. نتایج این بررسی نشان داد که هر دو ویروس توالی‌یابی شده از کشورمان ۹۹/۹ درصد با ویروس ووهان شباهت دارد. با این حال، آنالیز بیشتر ویروس توالی‌یابی شده از جمهوری اسلامی ایران در مقایسه با ویروس ووهان نشان داد که ژن‌های orf1ab polyprotein nsp2، orf1a polyprotein nsp6 ۳'-to-5' exonuclease و Nucleocapsid ۲'-O-ribose nsp2، orf1ab polyprotein phosphoprotein در ویروس دوم دارای SNP هستند (جدول ۱). ژن‌های ساختاری و غیر ساختاری فوق‌الذکر در فرآیند های حیاتی ویروس، مانند تکثیر، همانندسازی و ویرولانس آن، نقش اساسی دارند. به عنوان نمونه، گلیکوپروتئین اسپایک (Surface glycoprotein) موجود بر روی سطح بیرونی کرونا ویروس‌ها مسئول اتصال و ورود ویروس به سلول‌های میزان است. برهم‌کنش پروتئین اسپایک و گیرنده سلول میزان نقش تعیین‌کننده در آلوده کردن یک گونه میزانی خاص و همچنین کنترل تروپیسم بافتی

جدول ۱- مقایسه های ویروس SAR-CoV2 از جمهوری اسلامی ایران با هند، آمریکا، نیال، ایتالیا، اسپانیا و فرانسه. NA به معنی مشخص نشده می باشد

Proteins	SARS-CoV2/Iran (MT320891)	SARS-CoV2/Iran (MT447177)	SARS-CoV2/India (MT050493.1)	SARS-CoV2/US (MT344948)	SARS-CoV2/Nepal (MT072688.1)	SARS-CoV2/Italy (MT066156.1)	SARS-CoV2/Spain (MT359866)	SARS-CoV2/France (MT320538)
orf1ab polyprotein	V378I (1397G>A), L3606F(11083G>T), T6038I(18377C>T)	V378I (1397G>A), G6875R (20887G>A)	I476V(1691A>G), P2079L(6501C>T), T5538I(16877C>T)	P4715L (14408C>T)	NA	L3606X(11083G>N)	P4715L (14408C>T)	Y118C (618A>G) P4715L (14408C>T)
Leader protein	NA	NA	NA	NA	NA	NA	NA	Y118C (618A>G)
nsp2	V198I(1397G>A)	V198I (1397G>A)	I296V(1691A>G)	NA	NA	NA	NA	NA
nsp3	NA	NA	P1261L(6501C>T)	NA	NA	NA	NA	NA
nsp4	NA	NA	NA	NA	NA	NA	NA	NA
3C-like proteinase	NA	NA	NA	NA	NA	NA	NA	NA
nsp6	L37F(11083G>T)	NA	NA	NA	NA	L37X(11083G>N)	NA	NA
nsp7	NA	NA	NA	NA	NA	NA	NA	NA
nsp8	NA	NA	NA	NA	NA	NA	NA	NA
nsp9	NA	NA	NA	NA	NA	NA	NA	NA
nsp10	NA	NA	NA	NA	NA	NA	NA	NA
RNA-dependent RNA polymerase	NA	NA	NA	P323L (14408C>T)	NA	NA	P323L (14408C>T)	P323L (14408C>T)
helicase	NA	NA	T214I(16877C>T)	NA	NA	NA	NA	NA
-3'-to-5' exonuclease	T113I(18377C>T)	NA	NA	NA	NA	NA	NA	NA
endoRNase	NA	NA	NA	NA	NA	NA	NA	NA
2'-O-ribose	NA	G77R (20887G>A)	NA	NA	NA	NA	NA	NA
Methyltransferase	NA	NA	NA	NA	NA	NA	NA	NA
orf1a polyprotein	V378I(1397G>A), L3606F(11083G>T)	NA	I476V (1691A>G), P2079L(6501C>T)	NA	NA	L3606X(11083G>N)	NA	Y118C (618A>G)
nsp11	NA	NA	NA	NA	NA	NA	NA	NA
Surface glycoprotein	NA	T22I (21627C>T)	A930V(24351C>T)	D614G (23403A>G)	NA	NA	D614G (23403A>G)	D614G (23403A>G)
ORF3a protein	NA	NA	NA	NA	NA	G251V(26144G>T)	NA	NA
Envelope protein	NA	NA	NA	NA	NA	NA	NA	NA
Membrane glycoprotein	NA	NA	NA	NA	NA	NA	NA	NA
ORF6 protein	NA	NA	NA	NA	NA	NA	N34S (27302A>G)	NA
ORF7a protein	NA	NA	NA	NA	NA	NA	NA	NA
ORF7b	NA	NA	NA	NA	NA	NA	NA	NA
ORF8 protein	NA	NA	L84S(28144T>C)	L84S(28144T>C)	NA	NA	NA	NA
Nucleocapsid phosphoprotein	NA	S186F (28830C>T)	NA	S194L (28854C>T)	NA	NA	R203K (28881G>A) 28882G>A), G204R (28883G>C)	NA
ORF10 protein	NA	NA	NA	NA	NA	NA	NA	NA

ناحیه ژن polymeraseRNA-dependent RNA و حذف شدن ۳ نوکلئوتید در ناحیه ژن surface glycoprotein (G107del (21881_21883delGGT)) مشاهده شده در ژنوم ویروس های شیوع یافته در ایران تنوعات ژنتیکی منحصر به فرد بوده و در ژنوم های ویروسی مربوط به ۶ کشور دیگر مشابه آن یافت نمی شود. وجود SNP در ژن های 3'-to-5' exonuclease (ویروس اول) و 2'-O-ribose (ویروس دوم) به صورت اختصاصی فقط در ژنوم های ویروسی کشورمان قابل مشاهده می باشد.

بررسی تأثیر جهش بر پایداری پروتئین

تأثیر تنوعات ژنتیکی تک نوکلئوتیدی در پایداری پلی پروتئین ها و

مربوط به orfها به ترتیب متعلق به ویروس به دست آمده از هند، اسپانیا، ایران (ویروس دوم) و آمریکا به ترتیب با ۵، ۴ و ۴ جهش و سپس فرانسه با ۳ جهش و کمترین جهش ها متعلق به ایتالیا (۲ جهش) می باشد. ویروس نپال فاقد جهش است و لذا شباهت ۱۰۰ درصدی به ویروس ووهان را نشان می دهد.

علاوه بر SNP، وجود تنوعات ژنتیکی چند نوکلئوتیدی به صورت حذف یا اضافه نیز در ژنوم های مورد مطالعه بررسی و مقایسه شد. از میان این ۷ ژنوم ویروسی مربوط به کشورهای مختلف فقط ایران (ویروس اول) و فرانسه به ترتیب هر کدام دارای یک تنوع ژنتیکی شامل اضافه شدن ۶ نوکلئوتید (X389L (14606_14607insCCTTAT) در

جدول ۲- پیش‌بینی اثر جهش‌های تک نوکلئوتیدی بر پایداری پروتئین‌های ویروس اول SARS-CoV-2 شیوع یافته در جمهوری اسلامی ایران با شماره دسترسی MT320891

Proteins	Single nucleotide polymorphisms (SNPs)	Protein stability
orf1ab polyprotein	V378I (1397G>A)	Decrease (G= -0.479)
	L3606F(11083G>T)	Decrease (G= -1.292)
	T6038I(18377C>T)	Decrease (G= -0.114)
nsp2	V1981(1397G>A)	Decrease (G= -0.479)
nsp6	L37F(11083G>T)	Decrease (G= -1.292)
3'-to-5'exonuclease3	T113I(18377C>T)	Decrease (G= -0.114)
orf1a polyprotein	V378I(1397G>A)	Decrease (G= -0.479)
	L3606F(11083G>T)	Decrease (G= -1.292)

جدول ۳- پیش‌بینی اثر جهش‌های تک نوکلئوتیدی بر پایداری پروتئین‌های ویروس دوم SARS-CoV-2 شیوع یافته در جمهوری اسلامی ایران با شماره دسترسی MT447177

Proteins	Single nucleotide polymorphisms (SNPs)	Protein stability
orf1ab polyprotein	V378I (1397G>A)	Decrease (G= -0.479)
	G6875R (20887G>A)	Decrease (G= -0.587)
nsp2	V378I (1397G>A)	Decrease (G= -0.479)
2'-O-ribose	G6875R (20887G>A)	Decrease (G= -0.587)
Surface glycoprotein	T22I (21627C>T)	Decrease (G= -0.320)
Nucleocapsid phosphoprotein	S186F (28830C>T)	Decrease (G= -0.47)

پژوهش حاضر و پژوهش یاد شده احتمالاً می‌تواند به همزمانی تقریبی نمونه‌ها در این دو پژوهش مربوط باشد.

مطالعات نشان می‌دهد که دو ژن ساختاری اسپایک (S) و نوکلئوپسید (N) در انتقال ویروس بسیار مهم هستند به نحوی که جهش در این ژن‌ها می‌تواند در پایداری و بیماری‌زایی ویروس مؤثر باشد (۱۵). در میان هفت کشور مورد بررسی در این مطالعه، ایران (ویروس دوم) و آمریکا دارای یک جهش و اسپانیا دارای دو جهش در ژن نوکلئوپسید است. در مورد ژن اسپایک نیز، ایران (ویروس دوم)، هند، آمریکا، اسپانیا و فرانسه دارای جهش می‌باشند. با توجه به نقش اساسی پروتئین اسپایک در اتصال به گیرنده سلول میزان و تعیین تروپیسم یافته ویروس و در نتیجه بیماری‌زایی ویروس، برخی از محققان هندی بر این باورند که دلیل پایین بودن بیماری‌زایی این ویروس در هند می‌تواند به این جهش منحصر به فرد ژن اسپایک ویروس هند وابستگی داشته باشد. بررسی Anti-viral-SARS-CoV-2 در هند نشان داد که به طور ویژه hsa-mir-27b تنها به ژنوم ویروس هندی و به طور اختصاصی به ناحیه جهش یافته در پروتئین اسپایک اتصال می‌یابد. در مطالعات قبلی اثر بازدارنده این miRNA بر همانندسازی ویروس HIV به اثبات رسیده بود که شاید روش‌کننده دلیل مؤثر بودن تیمارهای ضد HIV برای درمان بیماری کووید-۱۹ در هند باشد (۱۶).

مطالعه الگو و نوع جهش‌ها به منشأی‌ای ویروس، کمک زیادی می‌کند. مطالعه SNP (V1981(1397G>A)) مربوط به پروتئین nsp2 مشترک بین نمونه‌های ویروسی اول و دوم از کشورمان، منشأ مشترک این دو ویروس را نشان

پروتئین‌های ویروسی با استفاده از روش محاسباتی و نرم افزاری، مورد بررسی و پیش‌بینی قرار گرفت. تجزیه و تحلیل جهش‌های تک نوکلئوتیدی همان‌طور که در جدول ۲ و ۳ قابل مشاهده است نشان داد که تمام جهش‌های تک نوکلئوتیدی ویروس‌های SARS-CoV-2 شیوع یافته در ایران می‌تواند پایداری پلی‌پروتئین‌ها و پروتئین‌های ویروسی را کاهش دهد. بنابراین، ادامه این جهش‌ها طی زمان شاید بتواند بر عملکرد پروتئین بیشتر تأثیرگذار باشد.

نتیجه‌گیری

توجه به این نکته ضروری است که سرعت جهش در ویروس‌های RNA (۱۲) از جمله SARS-CoV-2 بالاست. میزان جهش ویروس SARS-CoV-2 $10^{-3} \times 10^{-3} \text{ تا } 10^{-3} \text{ تا } 1/31$ تعویض در هر سایت در هر سال، محاسبه شده است که تقریباً با میزان جهش تخمین زده شده برای ویروس مولد بیماری مرس (MERS-CoV) مشابه است (۱۳ و ۱۴) که این امر می‌تواند از ویروس SARS-CoV-2 شیوع یافته اولیه یک جمعیت ویروسی بزرگ با ژنوم‌های تغییر یافته را ایجاد کند. میانگین جهش‌های SNP این هفت کشور مورد مطالعه در پژوهش حاضر ۳/۲۵ جهش است. این نتیجه با نتایج دانشمندان چینی که ۳۳ جهش را در ۱۱ نمونه ویروسی به دست آمده از بیماران مختلف یعنی به طور میانگین ۳ جهش برای هر ژنوم ویروسی گزارش کردند مشابه است (۷). با توجه به اینکه ارتباط زیادی بین زمان نمونه گیری و میزان تنوع ژنتیکی در ژنوم ویروس کووید-۱۹ وجود دارد، مطابقت میانگین تنوع جهش‌های محاسبه شده در

بین جهش‌های رخ داده در ژنوم ویروس SARS-CoV-2 و تغییر در بیماری‌زایی ویروس وجود دارد (۷). الگوی جهش‌ها و فراوانی آن‌ها در میزان‌ها، کشورها و زمان‌های مختلف می‌تواند بسیار متفاوت باشد. بنابراین می‌توان نتیجه گرفت که بررسی منظم و دقیق تنوع ژنتیکی ویروس در هر کشور یا منطقه جغرافیایی در طول زمان شیوع این بیماری پاندمیک می‌تواند اطلاعات ارزشمندی را از تغییرات ژنتیکی وابسته به زمان شیوع در اختیار قرار دهد که مطمئناً این تغییرات با قدرت بیماری‌زایی و سرعت شیوع ویروس مطابقت خواهد داشت. از طرف دیگر این داده‌ها برای ساخت واکسن و دارو و حتی برای تجویز داروی مناسب نیز بسیار ضروری و سودمند هستند.

می‌دهد. علاوه بر این، های (P323L (14408C>T)) مربوط به پروتئین RNA-dependent RNA polymerase D614G (23403A>G) مربوط به پروتئین اسپاکیک بین نمونه‌های ژنوم ویروسی فرانسه، اسپانیا و آمریکا کاملاً مشترک هستند. با توجه به اینکه طبق گزارش‌های WHO (۱۷) اولین گزارش بیماری کووید-۱۹ قاره اروپا به این سازمان از فرانسه در تاریخ ۲۴ ژانویه مربوط به سه بیمار سفر کرده به ووهان بوده است، می‌توان نتیجه گرفت که بخشی از ویروس شیوع یافته در اسپانیا از فرانسه منشأ گرفته است. از طرف دیگر می‌توان نتیجه گرفت که بخشی از ویروس شیوع یافته در آمریکا از اروپا منشأ گرفته و به طور مستقیم منشأ آن به چین باز نمی‌گردد.

نتایج مطالعات دانشمندان چینی نشان می‌دهد که شواهد مستقیمی

Phylogenetic and Genetic Diversity of SARS-CoV-2 Virus

Masoomeh Jannesar¹, Seyed Mahdi Seyed^{1,2*}

Abstract

Background: The incidence of COVID-19 by the coronavirus SARS-CoV-2 with symptoms of acute respiratory syndrome in December 2019 in Wuhan, China, has spread rapidly and change to a global epidemics.

Methods: SARS-CoV-2 virus phylogenetic study performed to identify the genetic diversity of the prevalent viruses in Iran and compared with 6 selected countries.

Results: Phylogenetic study of SARS-CoV-2 virus showed the genetic variations generated in the virus genome are not currently sufficient to distinguish significant phylogenetic differences between the genome of viruses obtained from different geographical areas (19 countries from different continents) but can form variety of clusters. A study of the SARS-CoV-2 virus mutations showed that the prevalent viruses in Iran had six single-nucleotide (SNP) genetic variations in non-structural and structural genes and one insertion of six nucleotides in the RNA-dependent RNA polymerase gene region. Comparing the genetic variations of viruses in Iran with the six selected countries (India, the United States, Nepal, Italy, Spain and France) indicate the uniqueness of genetic variations of the prevalent viruses in Iran.

Conclusion: Studies show that is a direct connection between mutations in the SARS-CoV-2 virus genome and changes in pathogenesis. Therefore, regular assessment of the virus genetic diversity in each country can provide valuable data for the development of vaccines and drugs.

Keywords: Acute Respiratory Syndrome, COVID-19, Genetic Variation

منابع

1. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Coronaviruses* 2015; 1282: 1-23.
2. Almeida JD, Tyrrell D. The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture. *Journal of General Virology* 1967; 1(2): 175-178.
3. Berry D, Cruickshank J, Chu H, Wells R. The structure of infectious bronchitis virus. *Virology* 1964; 23(3): 403-407.
4. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *New England journal of medicine* 2003; 348(20): 1953-1966.
5. Kim K, Tandi T, Choi JW, Moon J, Kim M. Middle East respiratory syndrome coronavirus (MERS-CoV) outbreak in South Korea, 2015: epidemiology ,characteristics and public health implications. *Journal of Hospital Infection* 2017; 95(2): 207-213.
6. Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research* 2020; 24: 91-98.
7. Yao H, Lu X, Chen Q, Xu K, Chen Y, Cheng L, et al. Patient-derived mutations impact pathogenicity of SARS-CoV-2. 2020. Available at: medrxiv.org
8. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, et al. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic acids research* 2012; 40(D1): D593-D598.
9. Cleemput S, Dumon W, Fonseca V, Karim WA, Giovanetti M, Alcantara LC, et al. Genome Detective Coronavirus Typing Tool for rapid identification and characterization of novel coronavirus genomes. *Bioinformatics* 2020.
10. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic acids research* 2005; 33(suppl_2): W306-W310.

1. Member of Scientific Leadership Working Group, Academy of Medical Sciences, IR Iran

2. *National Institute of Genetic Engineering and Biotechnology, Tehran, IR Iran

11. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins: Structure, Function, and Bioinformatics* 2006; 62(4): 1125-1132.
12. Leaf-nosed bat. Encyclopædia Britannica: Encyclopædia Britannica Online. 2009. Available at: britannica.com.
13. Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palser AL, Ong SH, et al. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *MBio* 2014; 5(1): e01062-13.
14. Li X, Zai J, Zhao Q, Nie Q, Li Y, Foley BT, et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *Journal of medical virology* 2020; 92: 602-611.
15. Baric RS, Yount B, Hensley L, Peel SA, Chen W. Episodic evolution mediates interspecies transfer of a murine coronavirus. *Journal of virology* 1997; 71(3): 1946-1955.
16. Sardar R, Satish D, Birla S, Gupta D. Comparative analyses of SAR-CoV2 genomes from different geographical locations and other coronavirus family genomes reveals unique features potentially consequential to host-virus interaction and pathogenesis. 2020. Available at: [biorxiv.org](https://www.biorxiv.org).
17. Fehr, A.R., Perlman S., Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* 2015; 1282: 1-23.
18. <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid19/news/news/> 2020/01/2019-ncov-outbreak-first-cases-confirmed-in-europe