

پیشگفتار.....	۷
فصل اول: جستجو در منابع آنلاین ژنتیکی.....	۹
فصل دوم: چشم اندازی از توارث در انسان.....	۲۱
فصل سوم: خواندن، فهمیدن و رسم شجره نامه ی انسان.....	۳۳
فصل چهارم: سیتوژنتیک.....	۵۱
فصل پنجم: جستجو در داده پایگاه‌های توالی DNA, RNA، پروتئین و جستجوگرهای ژنوم....	۶۱
فصل ششم: جستجوی ابزارهای آنلاین بیوانفورماتیک.....	۷۷
فصل هفتم: بیماری‌های پیچیده ی شایع و توارثی چندعاملی.....	۸۵
فصل هشتم: نوروزنتیک و ژنتیک رفتاری.....	۹۳
فصل نهم: ژنتیک سرطان.....	۱۰۳
فصل دهم: مشاوره ی ژنتیک.....	۱۱۵
فصل یازدهم: ابزار پیشرفته ی ویرایش ژنوم: CRISPR-Cas.....	۱۲۱
واژه نامه	۱۲۷
واژه یاب.....	۱۳۹

تقدیم بہ عشق زندگیم
کوروش

ہمسرمہربان و عزیزم

يُؤْتِي الْحِكْمَةَ مَنْ يَشَاءُ وَمَنْ يُؤْتَ الْحِكْمَةَ فَقَدْ أُوتِيَ خَيْرًا كَثِيرًا وَمَا يَذَّكَّرُ إِلَّا أُولُو الْأَلْبَابِ (بقره/ ۲۹۶)

امروزه با پیشرفت علم ژنتیک و ابزارهای آنالیز، تشخیص بیماری‌ها با استفاده از ترکیبی از علوم مختلف مانند ژنتیک، پزشکی، کامپیوتر، و همچنین ادغام این علوم در داده‌پایگاه‌های آنالیز بسیار راحت شده است. علاوه بر این با آنالیز ژنوم و بررسی جهش‌ها و پلی‌مورفیسم‌های موجود در ژنوم هر فرد می‌توان راهکارهای درمانی دقیق‌تری را برای هر شخص به‌صورت پزشکی شخصی ارائه داد که دارای بازدهی بالاتری باشد. به همین منظور به کارگیری علم بیوانفورماتیک می‌تواند راه‌گشای درمان بسیاری از بیماری‌ها و همچنین بررسی و تعامل دستاوردهای علمی با سایر محققان و دانشمندان دنیا باشد. از این رو می‌توان این کتاب را به عنوان آموزش استفاده از بیوانفورماتیک در علم ژنتیک به محققان و دانشجویان این رشته پیشنهاد داد.

خدای بزرگ را سپاس می‌گویم بخاطر اینکه به من لیاقت داد تا بتوانم سومین کتاب خود را ترجمه کنم. مطمئنم نگاه و دعای پیامبر گرامی اسلام همیشه همراهم بوده و هست. بهترین پدر دنیا و بهترین مادر دنیا هرچه کنم نمی‌توانم زحمات شما عزیزان دلم را جبران کنم اما همیشه تمام تلاشم را خواهم کرد که لایق مهربانی‌های بی‌دریغ شما باشم. همسر مهربانم تو نمونه‌ی واقعی مردی بی‌نظیر و فداکار هستی. از تو عشق و مهر آموختم. ممنونم برای تمام حمایت‌هایت برای تمام لحظاتی که اگر خسته شده بودم نگذاشتی دست از تلاش بردارم، آنقدر تشویقم کردی و به من انگیزه دادی تا توانستم با موفقیت کارم را به پایان برسانم.

استاد ارجمندم جناب آقای دکتر محمدرضا مشایخی بسیار خرسندم که افتخار شاگردی شما را داشتم. برای جبران زحمات شما این کتاب شاید کمترین کاری بود که توانستم انجام دهم. همواره حق استادی بر گردن من دارید. امیدوارم شاگرد لایقی برای شما باشم.

فصل اول

جستجو در منابع آنلاین ژنتیکی

نتایج یادگیری

پس از اتمام این فصل شما قادر خواهید بود:

- به راحتی منابع ژنتیکی آنلاین را جستجو، پیدا و استفاده کنید.
- به راحتی از طریق صفحه وب مرکز ملی اطلاعات بیوتکنولوژی (NCBI) پیمایش کنید.
- شروع به پاسخ دادن به سؤالات علوم زیست شناختی از طریق منابع آنلاین کنید.

پیش زمینه

امروزه اینترنت به یک ابزار اساسی برای جستجو و آموزش ژنومیک و ژنتیک انسانی تبدیل شده است و دسترسی به منابع آنلاین انبوهی را فراهم می کند که ذخیره سازی داده ها، دسترسی و آنالیز را برای هر کاربر از جمله دانشمندان، دانش آموزان، پزشکان، بیماران و خوانندگان عمومی را تسهیل کرده است [۱].

توسعه منابع آنلاین برای ژنومیک و ژنتیک انسانی از دهه ۱۹۸۰ شروع شد. از آن زمان منابع آنلاین که اطلاعات و داده ها را برای تحقیقات ژنومیک و ژنتیک فراهم می کنند به سرعت رشد کرده اند و اینترنت به یک منبع بسیار پر طرفدار برای گردآوری جدیدترین تحقیقات و نوآوری ها تبدیل شده است [۲].

طیف وسیعی از داده های ژنومیک و ژنتیک برای دسترسی و تجزیه و تحلیل به صورت آنلاین قرار دارد. برای مثال داده پایگاه هایی برای نقشه برداری ژنی، نقشه برداری مناطق کروموزومی، نقشه برداری اپی ژنوم، داده های توالی DNA و RNA و پروتئین، داده توالی اختصاصی گونه ها مقایسه بین گونه ها، داده اختصاصی بیماری، داده بیان ژن از بافت های سالم و بیمار، جستجوگرهای ژنوم و سایر اطلاعات مشابه را می توان نام برد. استفاده از این منابع می تواند برای پاسخگویی به سؤالات تحقیقاتی خاص توسط کاربران نهایی مناسب باشد.

یکی از پیشرفت ها بسیار مهم در حوزه ژنتیک، تولید داده های توالی کل ژنوم است. در دسترس بودن توالی کل ژنوم آرایه های وسیعی از داده های ژنتیکی را برای مناطق کد کننده، مناطق غیر کد کننده، ساختار کروموزوم و سازماندهی ژن ها، مناطق ژنی روی کروموزوم ها، تعداد ژن ها، تعداد جفت بازها، توالی نوکلئوتیدی و انواع DNA در اختیار ما قرار می دهد [۳].

تا کنون انواع مختلفی از موجودات از جمله گونه هایی از باکتری ها، یوکاریوت ها و دومین های آرکی باکتری ها توالی یابی شده است. برخی نمونه هایی که قبل تر بر روی آنها کار شده عبارتند از *هموفیلوس آنفلانزا* (۱۹۹۵)، *ساکارومایسزسوزیه* (۱۹۹۶) و *سنورابتیدیس الگانس* (۱۹۹۸) [۴-۱].

در راستای اهداف این کتاب، فقط پروژه های ژنوم انسان بررسی شده است.

پروژهی ژنوم انسان

پروژهی ژنوم انسان که در سال ۱۹۹۰ آغاز شد، یک پروژهی تحقیقاتی مشهور بین‌المللی است که برای تشخیص نقشه‌های ژنتیکی، فیزیکی و توالی ژنوم انسان توسط رمزگشایی کل توالی نوکلئوتیدها ایجاد شده است. رمزگشایی توالی کل ژنوم، تشخیص ژن‌های پراکنده در سرتاسر ژنوم را امکان پذیر کرده است [۷ و ۸]. دستیابی به توالی کل ژنوم اطلاعاتی را دربارهی ژن‌ها، عناصر تنظیمی و ساختار و سازمان‌دهی کروموزومی به ما نشان می‌دهد و این موضوع آغاز عصر جدیدی در حوزهی ژنتیک انسانی را تسهیل می‌کند [۸].

در سال‌های اخیر پیشرفت‌ها در زمینهی تکنولوژی توالی‌یابی نسل بعدی، توالی‌یابی کل ژنوم با صرفه جویی در زمان و هزینه را امکان پذیر ساخته است. این تکنولوژی‌های توالی‌یابی با توان بالا، نقطه‌ی عطفی در تحقیقات ژنومیک و ژنتیک انسانی به شمار می‌رود. این تکنولوژی‌ها امکان بررسی بی نقص بودن ژنوم و تغییرات ژنتیکی را در ژنوم انسان را فراهم می‌کند و قول بزرگی را برای تشخیص آسان شرایط مندلی همانند ناهنجاری‌های چندعاملی به ما می‌دهد. علاوه بر این، این فن آوری‌ها با معرفی واریانت‌های ژنتیک انسانی به مراکز درمانی به منظور پیش بینی، پیشگیری و کشف نوع خاصی از داروها کمک می‌کنند [۸].

گردآوری حجم بالایی از داده‌های توالی‌یابی، تجزیه و تحلیل عملکردی را به صورت دستی غیر ممکن ساخته است. از این رو داده‌پایگاه‌ها و ابزارهای کامپیوتری برای آنالیز داده‌های توالی‌یابی با توان بالا به کار گرفته می‌شوند. باید به این نکته توجه داشت که بسیاری از موارد آنالیز کامپیوتری توالی در مورد توارث صفات و بیماری‌های انسانی، باید با مطالعات عملکردی بیشتری تکمیل شوند [۸].

در اثر تجزیه و تحلیل توالی‌ها با توان بالا، مقدار زیادی از داده‌های ژنومیک و اپی‌ژنومیک در پایگاه داده‌های همانند مرکز ملی بیوتکنولوژی (NCBI)، جستجوگر ژنوم دانشگاه سانتاکروز کالیفرنیا (UCSC)، دایره‌المعارف عناصر (DNA ENCODE) و پروژهی اپی‌ژنومیک نقشه‌ی راه در دسترس قرار گرفته است. این داده‌پایگاه‌ها در زیر شرح داده شده‌اند.

مرکز ملی اطلاعات بیوتکنولوژی (NCBI)

NCBI بخشی از کتابخانه‌ی ملی پزشکی ایالات متحده است (NLM)، شاخه‌ای از مؤسسه‌ی ملی سلامت است که به‌طور رسمی در سال ۱۹۸۸ تأسیس شد. هدف اولیه‌ی آن توسعه‌ی سیستم‌های بر پایه‌ی کامپیوتر برای جمع‌آوری، تفسیر و ارائه‌ی داده‌های به‌دست آمده در زیست‌شناسی مولکولی انسانی، ژنتیک و بیوشیمی بود، و سپس به عنوان یک الگوی کار بین‌المللی برای به اشتراک گذاری داده‌ها و آنالیزهای داده‌ها بر پایه‌ی کامپیوتر تکامل پیدا کرد و به صورت جامع اطلاعاتی را از همه‌ی موجودات مورد پوشش قرار داد [۹].

صفحه‌ی اصلی NCBI طیف وسیعی از اطلاعات زیست‌شناسی را دربرمی‌گیرد که در دسترس و قابل بارگزاری هستند. بخش Entrez لینک‌هایی را برای آرایه‌های وسیعی از داده‌پایگاه‌ها فراهم می‌کند مانند بخش **Gene** که داده‌پایگاهی با تمرکز مخصوص بر روی داده‌های مخصوص ژن‌ها است که اطلاعاتی درباره‌ی مکان ژنی، محصول ژن، توالی اسیدنوکلئیک، انواع توالی‌های حاشیه‌نویسی شده، نقشه‌های ژنی و همولوگ‌های ژنی را همراه با لینک‌های دیگری برای داده‌های بیشتر فراهم می‌آورد. NCBI دسترسی به **OMIM** را فراهم می‌کند (توارث مندلی آنلاین در انسان) که داده‌پایگاهی است که اطلاعاتی را درباره‌ی بیماری‌های توارثی، ژن‌های مسئول بیماری، نقشه‌ها و مکان‌های کروموزومی، رابطه‌ی فنوتایپ-ژنوتایپ و ویژگی‌های بالینی در اختیار ما قرار می‌دهد (همچنین جنبه‌های ژنتیک مولکولی و سیتوژنتیک را دربرمی‌گیرد). یکی دیگر از بخش‌های اساسی NCBI داده‌پایگاه **PubMed** است که دسترسی آنلاین به آرشیو مطالب زیست‌پزشکی و خلاصه مقالات را از طریق MEDLINE و مجلات علمی را فراهم می‌سازد. NCBI این اجازه را به ما می‌دهد تا بین داده‌پایگاه‌های مختلف ارتباط برقرار کنیم.

داده‌پایگاه‌های دیگری نیز در NCBI جای گرفته‌اند که در صفحه‌ی اصلی لیست شده‌اند و استفاده از آنها در بخش آموزش و تمرین NCBI نشان داده شده است. در جدول ۱-۱ برخی از این داده‌پایگاه‌ها لیست شده‌اند.

جستجوگر ژنوم UCSC

پس از اتمام اولین پیش‌نویس‌های پروژه‌ی ژنوم انسان در سال ۲۰۰۰، ژنوم ایجاد شده در جستجوگر ژنوم UCSC قابل دسترس گردیده شد [۱۰]. امروزه جستجوگر ژنوم به صورت وسیعی برای رویت و تجزیه و تحلیل داده‌های ایجاد ژنوم مورد استفاده قرار می‌گیرد که از UCSC و یا سایر منابع گرفته می‌شود (داده‌هایی که توسط کاربران ایجاد می‌شود). UCSC ژنوم ایجاد شده‌ی بیش از ۹۳ موجود زنده همانند انسان، شامپانزه، موش و ماهی گورخری را در خود جای داده است [۱۰]. جستجوگر ژنوم UCSC دسترسی به داده‌های زیر را امکان‌پذیر می‌سازد:

- توالی‌یابی و نقشه‌برداری (مانند مکان بازها، میزان GC، محتویات و داربست)
- پیش‌گویی ژنها و محصولات آنها (مانند توالی کدکننده مورد توافق CDS و RNA غیر کد کننده)
- فنوتایپ و مفهوم آن
- mRNA و برچسب نمایانگر بخشهایی از ژنهای بیان شده (EST)
- بیان ژنها
- تنظیم ژنها (مانند تنظیم همبستگی از ENCODE، جزایر CpG)
- ژنومیک مقایسه‌ای
- بررسی تنوع (مانند پلی‌مورفیسم تک نوکلئوتیدی [SNP]) و تکرارها (مانند میکروساتلیت‌ها) [۱۰ و ۱۱].
- علاوه بر این ابزارهای متنوع آنالیزی در جستجوگر ژنوم UCSC وجود دارد که شامل:
- جستجوگر ژنوم که قادر است داده‌های ژنومی را شبیه‌سازی کند.
- BLAT که قابلیت هم‌ترازی توالی‌ها را دارد.
- *in-silico* PCR که توالی‌های پرایمر را در ژنوم هم‌تراز می‌کند.
- Gene Sorter که قادر است توسط معیارهای بیانی ژن‌های مشابه را جستجو کند.

دائرةالمعارف عناصر DNA (ENCODE)

پروژه‌ی مشترک ملت‌های مختلف است که توسط مؤسسه‌ی ملی ژنوم انسان برای پیشبرد پروژه‌ی ژنوم انسان تأسیس شده است و هدفش بررسی توالی عملکردی DNA ژنوم انسان است. ENCODE اطلاعاتی درباره‌ی منشأ ژنوم انسان را برای ما مهیا کرده است همانند:

- مناطق کد کننده و غیر کد کننده
- ژن‌های کد کننده‌ی پروتئین
- رونوشت‌های غیر کد کننده
- RNA های غیر کد کننده‌ی طویل
- ژن‌های کاذب
- محلی شدن تحت سلولی رونوشت‌ها در انسان به وسیله‌ی توالی‌یابی
- متیلاسیون DNA
- شاخص دسترسی کروماتین در نواحی تنظیمی DNA
- جایگاه‌های اتصال شونده فاکتورهای رونویسی
- واکنش‌های متقابل در فضای سه بعدی
- یکپارچه سازی SNP های به دست آمده از مطالعات مرتبط با کل ژنوم (GWAS) [۱۲-۱۴].

پروژه‌ی اپی‌ژنومیک نقشه‌ی راه

اطلاعات به‌دست آمده از توالی ژنوم به تنهایی برای توصیف فرآیندهای فیزیولوژیکی و پاتولوژیکی بدن انسان کافی نیست. هنگامی که درباره‌ی بیان ژن و تنظیم آنها در درون ژنوم می‌اندیشیم، اپی‌ژنتیک مورد توجه قرار می‌گیرد. اپی‌ژنتیک تغییرات عملکرد ژن را در طول بیان ژن بدون تغییر در توالی نوکلئوتید اصلی، مطالعه می‌کند. پروژه‌ی اپی‌ژنوم برای نشان دادن ساختار اپی‌ژنومیک ژنوم انسان شروع به کار کرد. این پروژه، مکانیسم‌های اپی‌ژنتیک شامل متیلاسیون DNA و تغییرات هیستونی را مورد بررسی قرار می‌دهد. فهمیدن این مکانیسم‌ها می‌تواند دید بهتری را نسبت به بیان ژن و بیماری‌های انسان ایجاد کند [۱۵ و ۱۶].

این نکته را باید مورد توجه داشت که این فصل به‌طور کامل تمام قسمت‌های مرتبط با منابع آنلاین ژنتیک را پوشش نداده است.

جدول ۱-۱: لیستی از داده‌پایگاه‌های انتخابی قابل دسترس از طریق صفحه‌ی اصلی NCBI.

URL	توضیحات	داده‌پایگاه
http://www.ncbi.nlm.nih.gov/clinvar	تنوع ژنی تعیین شده و شرح مصور موارد مربوط به سلامت انسان.	ClinVar
http://www.ncbi.nlm.nih.gov/dbvar	واریانت‌های ساختاری ژنومیک حاشیه‌نویسی شده شامل حذف و اضافه، مضاعف‌سازی، معکوس شدگی، ترنس‌لوکیشن‌ها و سایر نوآرایی‌های پیچیده.	dbVar
http://www.ncbi.nlm.nih.gov/gene	اطلاعات فیزیکی و تشریحی را درباره‌ی ژن مورد نظر ارائه می‌دهد.	Gene
http://www.ncbi.nlm.nih.gov/genome	داده‌های ژنومیک را از هر گونه‌ای مهیا می‌کند و امکان دسترسی به منابع ژنومی انسان، میکروب‌ها، اندامک‌ها، ویروس‌ها و ژنوم رفرنس پروکاریوت‌ها را امکان‌پذیر می‌سازد.	Genome
http://www.ncbi.nlm.nih.gov/medgene	اطلاعاتی را از جنبه‌ی بالینی درباره‌ی ژنتیک انسانی مهیا می‌کند.	MedGen
http://www.ncbi.nlm.nih.gov/nucleotide	توالی نوکلئوتیدی برگرفته شده از چندین منبع را در اختیار ما قرار می‌دهد. ژنوم، ژن و توالی‌های رونوشت از بخش‌هایی هستند که در بخش Nucleotide سایت NCBI در دسترس هستند.	Nucleotide
http://www.ncbi.nlm.nih.gov/omim	اطلاعاتی تشریحی را درباره‌ی بیماری‌های وراثتی و ساختار ژنتیکی آنها فراهم می‌کند.	OMIM
http://www.ncbi.nlm.nih.gov/protein	توالی‌های پروتئینی را از منابع انبوهی در دسترس قرار می‌دهد.	Protein
http://www.ncbi.nlm.nih.gov/pmc	دسترسی رایگان به منابع زیست‌پزشکی و علوم زیستی را در NIH و NLM امکان‌پذیر می‌سازد.	PMC
http://www.ncbi.nlm.nih.gov/pubmed	توصیفات کامل و جامع تحقیقات زیست‌پزشکی از MEDLINE و مجلات علوم زیستی و کتاب‌های آنلاین را شامل می‌شود.	PubMed
http://www.ncbi.nlm.nih.gov/snp	شامل پلی‌مورفیسم‌های تک‌نوکلئوتیدی (SNP) و واریانت‌های کمتری از سایر موارد می‌باشد.	SNP

تمرین‌های پایان فصل

۱. در این تمرین شما باید از منابع آنلاین ژنتیک به منظور درک بهتر داده‌پایگاه‌های ژنتیکی که معمولاً برای آموزش و تحقیقات دانشمندان به عنوان ابزارهای بیوانفورماتیکی استفاده می‌شود، استفاده کنید. به این منظور اسکرین‌شات‌هایی از صفحه‌ی NCBI برای