

# بیوانفورماتیک

جلد اول - پایگاه‌های داده و ترازایی

«ویراست چهارم»

آندره باکسوانیس - گری بیدر - دیوید ویشارت

دکتر ایوب ترکیان

نیاز دانش

## فهرست

صفحه	عنوان
۱	فصل اول: پایگاه‌های داده بیولوژیکی
۱	مقدمه
۲	پایگاه‌های داده نوکلئوتید
۲	فایل‌های توالی نوکلئوتید
۱۱	پایگاه‌های داده پروتئین
۱۵	خلاصه
۱۷	فصل دوم: بازیابی اطلاعات از پایگاه‌های داده بیولوژیکی
۱۷	مقدمه
۱۸	بازیابی اطلاعات یکپارچه: سیستم Entrez
۱۹	روابط بین درایه‌های پایگاه داده: همسایگی
۳۱	پایگاه‌های داده پزشکی
۳۵	پایگاه‌های توالی ارگانسمی و رای NCBI
۴۱	خلاصه
۴۳	فصل سوم: ارزیابی مشابهت توالی جفتی: FASTA و BLAST
۴۳	مقدمه
۴۴	ترازیابی توالی فراگیر و محلی
۴۵	ماتریس‌های نمره‌دهی
۵۲	BLAST
۶۳	توالی‌های BLAST <sub>۲</sub>
۶۳	MegaBLAST
۶۶	PSI-BLAST
۶۹	BLAT
۷۴	FASTA
۷۹	خلاصه
۸۰	فصل چهارم: مرورگرهای ژنوم
۸۰	مقدمه

۸۰	.....	مرورگر ژنوم UCSC
۹۵	.....	مرورگر ژنوم ENSEMBL
۱۰۷	.....	Biomart آنسمل
۱۱	.....	خلاصه

۱۱۳	.....	فصل پنجم: تبیین ژنوم
۱۱۳	.....	مقدمه
۱۱۴	.....	روش‌های پیش‌بینی ژن
۱۱۵	.....	پیش‌بینی از ابتدای ژنوم‌های پروکاریوتی
۱۲۱	.....	پیش‌بینی ژن از ابتدا در ژنوم‌های یوکاریوتی
۱۲۷	.....	میزان کارایی پیش‌بینی‌کنندگان ژن
۱۳۰	.....	تولید شواهد برای تبیین ژنوم
۱۳۲	.....	تبیین ژن و تولید شواهد با استفاده از پایگاه‌های داده توالی پروتئین
۱۳۳	.....	تبیین ژن و تولید شواهد با استفاده از پیش‌بینی ژن تطبیقی
۱۳۹	.....	زنجیره تبیین ژنوم
۱۴۲	.....	خلاصه

۱۴۵	.....	فصل ششم: روش‌های پیش‌بینی با استفاده از توالی‌های RNA
۱۴۵	.....	مقدمه
۱۴۸	.....	مرور پیش‌بینی ساختار ثانویه RNA با استفاده از ترمودینامیک
۱۵۱	.....	برنامه‌سازی دینامیک
۱۵۲	.....	صحت پیش‌بینی ساختار ثانویه RNA
۱۵۳	.....	پیش‌بینی ساختار ثانویه مشترک برای توالی‌های RNA چندگانه
۱۵۵	.....	مقدمه عملی روش‌های توالی تکی
۱۶۵	.....	مقدمه عملی روش‌های توالی چندگانه
۱۶۸	.....	دیگر روش‌های محاسباتی برای مطالعه ساختار RNA
۱۶۹	.....	مقایسه روش‌ها
۱۶۹	.....	پیش‌بینی ساختار ثالثیه RNA
۱۷۰	.....	خلاصه

۱۷۲	.....	فصل هفتم: روش‌های پیش‌بینی با استفاده از توالی‌های پروتئین
۱۷۲	.....	مقدمه
۱۷۳	.....	پیش‌بینی یک بُعدی ساختار پروتئین
۱۹۲	.....	پیش‌بینی کارکرد پروتئین
۲۰۵	.....	خلاصه

۲۰۸	فصل هشتم: ترازایی توالی چندگانه
۲۰۸	مقدمه
۲۱۰	سنجش کیفیت ترازایی چندگانه
۲۱۲	انجام ترازایی: مسایل عملی
۲۱۲	بسته‌های ترازایی رایج
۲۱۹	مشاهده ترازایی چندگانه
۲۲۱	خلاصه
۲۲۴	فصل نهم: تکامل مولکولی و تحلیل فیلوژنتیک
۲۲۴	مقدمه
۲۲۵	توالی به عنوان ساعت مولکولی
۲۲۶	اصطلاحات پس‌زمینه و مبانی
۲۳۵	نحوه ساخت درخت
۲۴۱	مطالعات تکامل مبتنی بر مارکر
۲۴۴	تحلیل فیلوژنتیک و یکپارچه‌سازی داده‌ها
۲۴۷	چالش‌های آتی
۲۴۹	فصل دهم: تحلیل بیان
۲۴۹	مقدمه
۲۵۰	گام صفر: انتخاب فناوری تحلیل بیان
۲۵۲	گام ۱: طراحی آزمایش
۲۵۴	گام ۲: جمع‌آوری و مدیریت داده‌ها و متاداده‌ها
۲۵۴	گام ۳: پیش‌پردازش داده‌ها
۲۵۵	گام ۴: کنترل کیفیت
۲۵۹	گام ۵: اثرات نرمال‌سازی و بیج
۲۶۴	گام ۶: تحلیل داده کاوشی
۲۷۳	گام ۷: تحلیل بیان دیفرانسیل
۲۷۸	گام ۸: کاوش مکانیسم‌ها از طریق تحلیل غنی‌سازی کارکردی
۲۸۰	گام ۹: توسعه دسته‌بند
۲۸۱	توالی‌یابی تک‌سلولی
۲۹۰	خلاصه
۲۹۱	پیوست‌ها



# فصل ۱

## پایگاه‌های داده بیولوژیکی

### مقدمه

در چند دهه گذشته، تلاش زیادی برای شناخت «کتاب حیات» در ابتدایی‌ترین سطح انجام شده است. بیولوژیست‌ها (و دانشمندان به طور عام) در خصوص نحوه ذخیره همه اطلاعات مورد نیاز توسط سلول برای انجام فرایندهای متابولیکی ضروری در حیات ارگانیسم و منتقل شده از نسل به نسل، بررسی‌های زیادی انجام داده‌اند. به منظور شناخت پایه نحوه تداوم موتور حیات توسط مجموعه نوکلئوتیدها، مقادیر زیادی داده‌های توالی باید به شیوه‌ای جمع‌آوری و ذخیره شده، که آنها را بتوان به سادگی جستجو و تحلیل کرد. در این راستا، تلاش زیادی برای طراحی و نگهداری پایگاه‌های داده توالی بیولوژیکی صورت گرفته است. این پایگاه‌های داده اثر قابل ملاحظه‌ای در بهبود شناخت بیولوژی از نقطه نظر محاسباتی و نیز از طریق استفاده یکپارچه، همراه با مطالعات آزمایشگاهی، داشته است.

اتمام توالی‌سنجی ژنوم انسان و ارگانیسم‌های مدل متعدد، و نیز وجود تعداد زیادی توالی به صورت عام، فرصتی طلایی برای متخصصین بیولوژی در استفاده از این داده‌های دارای ارزش ذاتی فراهم آورد. هم‌زمان، حجم بالا و فزاینده داده‌ها نیز برای فرد مبتدی مشکلاتی ایجاد می‌کند. در واقع، منظره توالی‌سنجی در سال‌های اخیر با توسعه فناوری‌های ظرفیت بالای بهتر، ارزان‌تر، و سریع‌تر به مقدار قابل ملاحظه تغییر کرده است. با توجه به ارزش نهفته داده‌های موجود در این پایگاه‌های داده توالی، تمرکز این فصل بر بسترسازی شناخت قوی پایگاه‌های اصلی به عنوان اولین گام در انجام تحلیل‌های ستبر و صحیح، می‌باشد.