



دانشگاه شیراز



انتشارات مرکز منطقه‌ای
اطلاع‌رسانی علوم و فناوری

تجزیه و تحلیل داده‌های

توالی‌یابی نسل جدید

تألیف:

زینکن وانگ

ترجمه:

سید ضیاءالدین میرحسینی، شاهرخ قوتی، زهرا پزشکیان

سرشناسه	وانگ، شینکون، ۱۹۷۲-م.
	-Wang, Xinkun, 1972
عنوان و نام پدیدآور	تجزیه و تحلیل داده‌های توالی‌یابی نسل جدید/ تالیف زینکن وانگ؛ ترجمه سیدضیاءالدین میرحسینی، شاهرخ قوتی، زهرا پزشکیان.
مشخصات نشر	شیراز: مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری، انتشارات، ۱۴۰۰.
مشخصات ظاهری	۲۷۷ ص.
شابک	978-622-97759-8-1
وضعیت فهرست نویسی	فیبا
یادداشت	عنوان اصلی: Next-generation sequencing data analysis, 2016.
یادداشت	کتابنامه: ص. ۲۵۷-۲۷۷.
موضوع	ژنگان‌شناسی
موضوع	Genomics
موضوع	ژنوم
موضوع	Genomes
شناسه افزوده	میرحسینی، ضیاءالدین، ۱۳۳۷-، مترجم
شناسه افزوده	قوتی، شاهرخ، ۱۳۵۸-، مترجم
شناسه افزوده	پزشکیان، زهرا، ۱۳۶۲-، مترجم
شناسه افزوده	ایران: وزارت علوم، تحقیقات و فناوری. مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری
شناسه افزوده	Iran. Ministry of Science, Research and Technology Regional Information Center for Science and Technology
رده بندی کنگره	QH۴۴۷
رده بندی دیویی	۸۶۲۹/۵۷۲
شماره کتابشناسی ملی	۸۴۳۴۵۱۹
اطلاعات رکورد کتابشناسی	فیبا



دانشگاه گیلان



انتشارات مرکز منطقه‌ای
اطلاع‌رسانی علوم و فناوری

عنوان: تجزیه و تحلیل داده‌های توالی‌یابی نسل جدید
مترجمان: سید ضیاءالدین میرحسینی، شاهرخ قوتی، زهرا پزشکیان
ناشر: انتشارات مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری
شابک: ۹۷۸-۶۲۲-۹۷۷۵۹-۸-۱
تیراژ: ۲۰۰ نسخه

حق چاپ برای مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری و دانشگاه گیلان محفوظ است.

اداره انتشارات مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری

شیراز بلوار جمهوری اسلامی خیابان جام جم Publication@rice.ac.ir

تلفن: ۰۷۱۳۶۴۶۸۴۵۲

بخش ۱- مقدمه‌ای بر بیولوژی سلولی و مولکولی

۱. سیستم سلولی و کد حیات ۱
- ۱.۱ چالش سلولی ۱
- ۲.۱ سلول‌ها چگونه با چالش برخورد می‌کنند ۲
- ۳.۱ مولکول‌ها در سلول‌ها ۳
- ۴.۱ ساختارها یا فضاهای داخل سلولی ۴
- ۴.۱.۱ هسته ۴
- ۴.۱.۲ غشا سلولی ۶
- ۴.۱.۳ سیتوپلاسم ۶
- ۴.۱.۴ اندوزوم، لیزوزوم و پراکسیزوم ۷
- ۴.۱.۵ ریبوزوم ۸
- ۴.۱.۶ شبکه آندوپلاسمی ۹
- ۴.۱.۷ دستگاه گلژی ۱۰
- ۴.۱.۸ اسکلت سلولی ۱۰
- ۴.۱.۹ میتوکندری ۱۱
- ۴.۱.۱۰ کلروپلاست ۱۳
- ۵.۱ سلول به عنوان یک سیستم ۱۴
- ۵.۱.۱ سیستم سلولی ۱۴
- ۵.۱.۲ زیست‌شناسی سیستم سلول ۱۵
- ۵.۱.۳ چگونگی مطالعه سیستم سلولی ۱۶
۲. توالی DNA: اساس ژنوم ۱۸
- ۲.۱ مارپیچ دو رشته‌ای DNA و توالی باز ۱۸
- ۲.۲ چگونه مولکول‌های DNA همانند سازی می‌کنند و فرم اصلی را حفظ می‌نمایند ۱۹
- ۲.۳ چگونه اطلاعات ژنتیکی ذخیره شده در DNA به پروتئین انتقال می‌یابند ۲۱
- ۲.۴ چشم انداز ژنومی ۲۳
- ۲.۴.۱ ژنوم حداقلی ۲۳
- ۲.۴.۲ اندازه‌های ژنومی ۲۳
- ۲.۴.۳ مناطق کد کننده پروتئین در ژنوم ۲۴
- ۲.۴.۴ عناصر ژنومی غیر کد کننده ۲۵
- ۵.۲ بسته بندی DNA، دسترسی به توالی و تعاملات DNA-پروتئین ۲۸

۲۸.....	۲. ۵. ۱ بسته بندی DNA
۲۸.....	۲. ۵. ۲ دسترسی به توالی
۲۹.....	۲. ۵. ۳ تعاملات پروتئین- DNA
۳۰.....	۲. ۶ جهش و پلی مورفیسم توالی DNA
۳۳.....	۲. ۷ تکامل ژنوم
۳۴.....	۲. ۸ اپی ژنوم و متیلاسیون DNA
۳۶.....	۲. ۹ توالی یابی ژنوم و خطر بیماری
۳۶.....	۲. ۹. ۱ بیماری های مندلی (تک ژنی)
۳۷.....	۲. ۹. ۲ بیماری های پیچیده که ژن های متعدد را شامل می شوند
۳۷.....	۲. ۹. ۳ بیماری هایی که ناشی از ناپایداری ژنوم است
۳۸.....	۲. ۹. ۴ بیماری های اپی ژنومیک/اپی ژنتیکی
۴۰.....	۳. RNA: توالی رونویسی شده
۴۰.....	۳. ۱ RNA به عنوان پیام رسان
۴۰.....	۳. ۲ ساختار مولکولی RNA
۴۲.....	۳. ۳. ۱ الگوی DNA
۴۳.....	۳. ۳. ۲ رونویسی ژن های پروکاریوتی
۴۵.....	۳. ۳. ۳ رونویسی اولیه پیش ساز mRNA ژن های یوکاریوتی
۴۷.....	۳. ۳. ۴ بلوغ mRNA از پیش ساز mRNA
۵۰.....	۳. ۳. ۵ انتقال و محل یابی
۵۰.....	۳. ۳. ۶ پایداری و زوال
۵۲.....	۳. ۳. ۷ مراحل اصلی تنظیم سطح رونویسی mRNA
۵۳.....	۳. ۴ RNA بیشتر از یک پیام رسان است
۵۴.....	۳. ۴. ۱ رایبوزایم
۵۵.....	۳. ۴. ۲ snRNA و snoRNA
۵۵.....	۳. ۴. ۳ RNA برای همانندسازی تلومر
۵۶.....	۳. ۴. ۴ RNAi و RNA های غیر کد کننده کوچک
۶۰.....	۳. ۴. ۵ RNA های غیر کد کننده طویل
۶۱.....	۳. ۴. ۶ دیگر RNA های غیر کد کننده
۶۱.....	۳. ۵ چشم انداز ترانسکریپتی سلولی
	بخش ۲- مقدمه ای بر توالی یابی نسل جدید (NGS) و آنالیز داده های NGS
۶۴.....	۴. تکنولوژی های توالی یابی نسل جدید (NGS): ورودی ها و خروجی ها
۶۴.....	۴. ۱ نحوه توالی یابی DNA: از نسل اول تا نسل جدید

۲.۴	گردش کار تجربی یک توالی‌یابی نسل جدید معمولی	۶۸
۳.۴	ورودی‌ها و خروجی‌های انواع پلت فرم‌های NGS	۷۱
۳.۴	۱.۳ توالی‌یابی خاتمه دهنده با رنگ برگشت‌پذیر ایلومینا	۷۲
۳.۴	۲.۳ توالی‌یابی نیمه‌هادی آیون تورنت	۷۸
۳.۴	۳.۳ توالی‌یابی زمان واقعی تک مولکولی پیسیفیک بایوساینسز (SMRT)	۸۰
۴.۴	۴.۴ اریبی‌ها و سایر عوامل اشتباه که ممکن است دقت داده‌های NGS را تحت تاثیر قرار دهند	۸۱
۴.۴	۱.۴ اریبی‌ها در ساخت کتابخانه	۸۳
۴.۴	۲.۴ اریبی‌ها و عوامل دیگر در توالی‌یابی	۸۴
۴.۴	۵. برنامه‌های کاربردی عمده NGS	۸۵
۴.۴	۵.۱ تعیین مشخصات ترانسکریپتوم و تشخیص تنوع پیرایشی (RNA-Seq)	۸۵
۴.۴	۵.۲ جهش ژنتیکی و کشف تغییرات	۸۵
۴.۴	۵.۳ اسمبل کردن ژنوم به روش از نو (د نوو)	۸۶
۴.۴	۵.۴ تجزیه و تحلیل متقابل پروتئین-DNA (ChIP-Seq)	۸۶
۴.۴	۵.۵ مطالعات اپی ژنومیک و متیلاسیون DNA (Methyl-Seq)	۸۶
۴.۴	۵.۶ متازنومیکس	۸۷
۵.۸	۵. آنالیز داده‌های توالی‌یابی نسل جدید (NGS) در مراحل اولیه: مراحل مشترک	۸۸
۵.۸	۱.۵ فراخوانی باز، فرمت فایل FASTQ، و نمره کیفیت باز	۸۹
۵.۸	۲.۵ کنترل کیفیت و پیش پردازش داده‌های NGS	۹۳
۵.۸	۳.۵ مکان‌یابی خوانش‌ها	۹۵
۵.۸	۳.۱ روش‌ها و الگوریتم‌های مکان‌یابی	۹۵
۵.۸	۳.۲ انتخاب الگوریتم‌های مکان‌یابی و توالی‌های ژنوم مرجع	۹۸
۵.۸	۳.۳ SAM / BAM به عنوان فرمت استاندارد مکان‌یابی	۹۹
۵.۸	۴.۳ بررسی و عملیات فایل مکان‌یابی	۱۰۲
۶.۸	۶. نیازمندی‌های محاسباتی برای مدیریت و تجزیه و تحلیل داده‌ها در توالی‌یابی نسل جدید (NGS)	۱۰۶
۶.۸	۱.۶ ذخیره سازی، انتقال و به اشتراک گذاشتن داده‌های NGS	۱۰۶
۶.۸	۲.۶ قدرت محاسباتی مورد نیاز برای تجزیه و تحلیل داده NGS	۱۰۹
۶.۸	۳.۶ نیازهای نرم افزاری برای تجزیه و تحلیل داده NGS	۱۱۰
۶.۸	۴.۶ مهارت‌های بیوانفورماتیکی مورد نیاز برای تجزیه و تحلیل داده NGS	۱۱۳

بخش ۳- تجزیه و تحلیل اختصاصی کاربردی داده‌های NGS

۱۱۷.....	۷. ترانسکریپتومیکس توسط توالی‌یابی RNA
۱۱۷.....	۷.۱ اصول RNA-Seq
۱۱۸.....	۷.۲ طراحی آزمایش
۱۱۸.....	۷.۲.۱ طرح فاکتوریل
۱۱۹.....	۷.۲.۲ تکرار و تصادفی بودن
۱۱۹.....	۷.۲.۳ آماده سازی نمونه
۱۲۱.....	۷.۲.۴ استراتژی توالی‌یابی
۱۲۲.....	۷.۳ تجزیه و تحلیل داده‌های RNA-Seq
۱۲۲.....	۷.۳.۱ کنترل کیفیت داده‌ها و مکان‌یابی خوانش‌ها
۱۲۵.....	۷.۳.۲ نرمال سازی داده‌های RNA-Seq
۱۲۸.....	۷.۳.۳ شناسایی ژن‌ها با بیان متفاوت
۱۳۱.....	۷.۳.۴ آنالیز پیرایشی متفاوت
۱۳۲.....	۷.۳.۵ تجسم داده‌های RNA-Seq
۱۳۳.....	۷.۳.۶ آنالیز عملکردی ژن‌های شناسایی شده
۱۳۳.....	۷.۴ RNA-Seq به عنوان ابزار کشف
۱۳۵.....	۸. توالی‌یابی RNA کوچک
۱۳۶.....	۸.۱ تولید داده‌ها و پردازش‌های بالادستی RNA کوچک توسط توالی‌یابی نسل جدید (NGS)
۱۳۶.....	۸.۱.۱ تولید داده‌ها
۱۳۸.....	۸.۱.۲ پیش پردازش
۱۳۹.....	۸.۱.۳ مکان‌یابی
۱۴۱.....	۸.۱.۴ شناسایی گونه‌های RNA کوچک شناخته شده و احتمالی
۱۴۱.....	۸.۱.۵ نرمال سازی
۱۴۲.....	۸.۲ شناسایی RNAهای کوچک با بیان متفاوت
۱۴۲.....	۸.۳ آنالیز عملکردی RNAهای کوچک شناسایی شده
۱۴۴.....	۹. تعیین ژنوتیپ و کشف تغییرات ژنوم توسط توالی‌یابی مجدد کل ژنوم
۱۴۶.....	۹.۱ پیش پردازش، مکان‌یابی، هم‌ردیفی مجدد، و کالیبراسیون مجدد داده‌ها
۱۴۷.....	۹.۲ تنوع تک نوکلئوتیدی (SNV) و فراخوانی ایندل
۱۴۷.....	۹.۲.۱ فراخوانی SNV
۱۴۹.....	۹.۲.۲ شناسایی جهش‌ها به روش از نو یا د نوو
۱۵۱.....	۹.۲.۳ فراخوانی کردن ایندل
۱۵۱.....	۹.۲.۴ فراخوانی تنوع داده‌های RNA-Seq

۱۵۲	۵.۲.۹ فایل فرمت فراخوانی تنوع (VCF).....
۱۵۴	۶.۲.۹ ارزیابی نتایج VCF.....
۱۵۵	۳.۹ فراخوانی تنوع ساختاری (SV).....
۱۵۵	۱.۳.۹ فراخوانی SV مبتنی بر خوانش جفتی.....
۱۵۶	۲.۳.۹ تعیین نقطه شکست.....
۱۵۷	۳.۳.۹ تشخیص SV مبتنی بر اسمبل به روش از نو.....
۱۵۷	۴.۳.۹ تشخیص CNV.....
۱۵۸	۵.۳.۹ آنالیز یکپارچه SV.....
۱۵۸	۴.۹ حاشیه نویسی تنوع‌های فراخوان شده.....
۱۵۹	۵.۹ آزمودن ارتباط تنوع با بیماری‌ها یا صفات.....
۱۶۱	۱۰.۱ اسمبل کردن ژنوم به روش از نو از خوانش‌های توالی‌یابی نسل جدید (NGS).....
۱۶۲	۱.۱۰ عوامل ژنتیکی و استراتژی‌های توالی‌یابی برای اسمبل کردن به روش از نو.....
۱۶۲	۱.۱.۱ عوامل ژنتیکی مؤثر بر اسمبل کردن به روش از نو.....
۱۶۳	۱.۱.۲ استراتژی‌های توالی‌یابی برای اسمبل کردن به روش از نو.....
۱۶۵	۲.۱۰ اسمبل کردن کانتیگ‌ها.....
۱۶۵	۱.۲.۱۰ پردازش داده‌های توالی‌یابی، تصحیح خطا و ارزیابی خصوصیات ژنوم.....
۱۶۷	۲.۲.۱۰ الگوریتم‌های اسمبلی کانتیگ.....
۱۷۰	۳.۱۰ اسکافولدینگ.....
۱۷۱	۴.۱۰ ارزیابی کیفیت اسمبلی.....
۱۷۲	۵.۱۰ بسته شدن شکاف.....
۱۷۳	۶.۱۰ محدودیت‌ها و توسعه آینده.....
۱۷۵	۱۱.۱ مکان‌یابی برهمکنش پروتئین و DNA با ChIP-Seq.....
۱۷۵	۱.۱۱ اصول ChIP-Seq.....
۱۷۷	۲.۱۱ طراحی آزمایش.....
۱۷۷	۱.۲.۱۱ کنترل آزمایش.....
۱۷۸	۲.۲.۱۱ عمق توالی‌یابی.....
۱۷۹	۳.۲.۱۱ تکرار.....
۱۷۹	۳.۱۱ مکان‌یابی خوانش، فراخوانی پیک (اوج) و مجسم‌سازی اوج.....
۱۷۹	۱.۳.۱۱ کنترل کیفیت داده‌ها و مکان‌یابی خوانش.....
۱۸۳	۲.۳.۱۱ فراخوانی اوج.....
۱۹۱	۳.۳.۱۱ مجسم‌سازی اوج.....
۱۹۱	۴.۱۱ آنالیز اتصال متفاوت.....

۱۹۴	۵.۱۱ تجزیه و تحلیل عملکردی
۱۹۵	۶.۱۱ تجزیه و تحلیل موتیف
۱۹۶	۷.۱۱ تجزیه و تحلیل یکپارچه داده‌های ChIP-Seq
۱۹۸	۱۲. تجزیه و تحلیل اپی ژنومیکس و متیلاسیون DNA توسط توالی‌یابی نسل جدید (NGS)
۱۹۹	۱۱.۱ استراتژی‌های توالی‌یابی متیلاسیون DNA
۱۹۹	۱۲.۱.۱ توالی‌یابی بی‌سولفیت کل ژنوم (WGBS)
۲۰۱	۱۲.۱.۲ توالی‌یابی با کاهش بازنمایی بی‌سولفیت (RRBS)
۲۰۱	۱۲.۱.۳ توالی‌یابی متیلاسیون بر اساس غنی‌سازی DNA متیله
۲۰۲	۱۲.۱.۴ تمایز متیلاسیون سیتوزین از محصولات دمتیلاسیون در توالی‌یابی بی‌سولفیت
۲۰۳	۱۲.۲ تجزیه و تحلیل داده‌های توالی‌یابی متیلاسیون DNA
۲۰۳	۱۲.۲.۱ کنترل کیفیت و پیش پردازش
۲۰۳	۱۲.۲.۲ مکان‌یابی خوانش
۲۰۶	۱۲.۲.۳ تعیین مقدار متیلاسیون DNA
۲۰۸	۱۲.۲.۴ تجسم داده‌های متیلاسیون DNA
۲۱۰	۱۲.۳ تشخیص سیتوزین‌ها یا مناطقی که به صورت متفاوت متیله شده
۲۱۱	۱۲.۴ تأیید، اعتبارسنجی و تفسیر داده‌ها
۲۱۲	۱۳. تجزیه و تحلیل متاژنوم توسط توالی‌یابی نسل جدید (NGS)
۲۱۳	۱۳.۱ طراحی آزمایش و تهیه نمونه
۲۱۵	۱۳.۱.۱ مجموعه نمونه متاژنوم
۲۱۵	۱۳.۱.۲ پردازش نمونه متاژنوم
۲۱۶	۱۳.۲ روش‌های توالی‌یابی
۲۱۷	۱۳.۳ بررسی کلی تجزیه و تحلیل داده‌های توالی‌یابی متاژنوم شات گان کل ژنوم (WGS)
۲۱۹	۱۳.۴ کنترل کیفیت و پیش پردازش داده‌های توالی‌یابی
۲۲۰	۱۳.۵ خصوصیات تاکسونومیک یک جامعه میکروبی
۲۲۰	۱۳.۵.۱ اسمبل کردن متاژنوم
۲۲۱	۱۳.۵.۲ گذاشتن توالی در پیاله
۲۲۳	۱۳.۵.۳ فراخوانی قاب‌های خوانش باز (ORF) و سایر عناصر ژنومی از توالی‌های متاژنومی
۲۲۴	۱۳.۵.۴ تجزیه و تحلیل نشانگر ژن فیلوژنتیکی
۲۲۵	۱۳.۶ خصوصیات عملکردی یک جامعه میکروبی
۲۲۵	۱۳.۶.۱ حاشیه نویسی عملکرد ژن
۲۲۶	۱۳.۶.۲ بازسازی مسیر متابولیک
۲۲۶	۱۳.۷ آنالیز متاژنومی مقایسه‌ای

۲۲۷.....	۱.۷.۱۳ نرمال سازی داده‌های توالی‌یابی متازنوم
۲۲۷.....	۲.۷.۱۳ شناسایی گونه‌های فراوان یا واحدهای تاکسونومیکی عملیاتی (OTU ها) مختلف
۲۲۸.....	۸.۱۳ خط مشی تجزیه و تحلیل داده‌های متازنومیکس
۲۲۸.....	۹.۱۳ مخازن داده متازنومیکس
بخش ۴- تغییر چشم انداز فن آوری‌های توالی‌یابی نسل جدید و تجزیه و تحلیل داده‌ها	
۲۳۱.....	۱.۱۴ آینده توالی‌یابی نسل جدید (NGS) چیست؟
۲۳۱.....	۱.۱۴ منظره در حال تغییر توالی‌یابی نسل جدید (NGS)
۲۳۳.....	۲.۱۴ تکامل سریع و رشد ابزارهای بیوانفورماتیک برای تجزیه و تحلیل داده‌های توالی‌یابی با توان بالا
۲۳۵.....	۳.۱۴ استاندارد سازی و کارآمد سازی خط مشی تحلیلی NGS
۲۳۶.....	۴.۱۴ محاسبات موازی
۲۳۷.....	۵.۱۴ محاسبات ابری
پیوست الف: انواع فایل‌های رایج مورد استفاده در تجزیه و تحلیل داده‌های توالی‌یابی نسل جدید	
۲۴۰.....	(NGS)
۲۴۴.....	پیوست ب: واژه نامه
۲۵۷.....	منابع