

توالی‌یابی RNA

(RNA Sequencing)

تالیف:

دکتر آرشد قلیان چی لنگرودی

(هیئت علمی گروه میکروبیولوژی و ایمنی شناسی، دانشکده دامپزشکی دانشگاه تهران)

دکتر بهلا آقائیان

(گروه میکروبیولوژی و ایمنی شناسی، دانشکده دامپزشکی دانشگاه تهران)

دکتر ریحانه - مردوست

(گروه میکروبیولوژی و ایمنی شناسی، دانشکده دامپزشکی دانشگاه تهران)

دکتر حسین حسینی

(هیئت علمی دانشکده دامپزشکی دانشگاه آزاد اسلامی واحد کرج)

Introduction on RNA Sequencing

عنوان و نام پدیدآور	توالی‌یابی RNA Sequencing = تالیف آرش قلیانچی لنگرودی... [و دیگران].
مشخصات نشر	تهران: سازمان جهاد دانشگاهی تهران، انتشارات، ۱۳۹۸.
مشخصات ظاهری	۱۴۵ ص: مصور، جدول.
شابک	۹۷۸-۶۰۰-۱۳۳-۴۲۸-۳
وضعیت فهرست نویسی	فیا
یادداشت	تالیف آرش قلیانچی لنگرودی، لیلا آفانیان، ریحانه خضردوست، حسین حسینی.
یادداشت	: کتابنامه.
موضوع	: نوکلئوتیدها -- توالی
موضوع	Nucleotide sequence
موضوع	: آر. ان. ا.
موضوع	RNA
شناسه افزوده	: قلیانچی لنگرودی، آرش، ۱۳۵۸ -
شناسه افزوده	: سازمان جهاد دانشگاهی تهران، انتشارات
رده بندی کلاسیک	QP ۶۲۵
رده بندی دیویی	۵۷۴/۸۷۳۲۸
شماره کتابخانه ملی	۵۷۹۷۸۲۵



- ♦ نام کتاب: توالی‌یابی RNA
- ♦ تالیف: آرش قلیانچی لنگرودی - لیلا آفانیان - ریحانه خضردوست - حسین حسینی
- ♦ ناشر: انتشارات سازمان جهاد دانشگاهی تهران
- ♦ ناظر چاپ: احمد آرش
- ♦ طراح جلد: سعید صحابی
- ♦ صفحه‌آرا: ساسی (رها)
- ♦ چاپ و صحافی: ...
- ♦ لیثوگرافی: زاویه نور
- ♦ قیمت: ۲۵۰۰۰۰ ریال
- ♦ شمارگان: ۱۰۰ نسخه

♦ نوبت چاپ: اول - تابستان ۱۳۹۸

ISBN: 978-600-133-428-3

♦ شابک: ۹۷۸-۶۰۰-۱۳۳-۴۲۸-۳

www.jahat.ir

♦ نشانی: تهران - صندوق پستی ۱۸۶-۱۳۱۴۵

en.jahat@gmail.com

♦ تلفن: ۶۶۹۵۴۳۸ تلفکس: ۶۶۶۶۴۹۴۱

♦ تلفن مراکز پخش: ۶۶۹۵۵۷۵۳-۶۶۹۵۴۳۸ خانه کتاب دانشگاه: ۶۶۹۷۳۴۲۳

این اثر، مشمول قانون حمایت مولفان، مصنفان و هنرمندان مصوب ۱۳۴۸ است. هرکس تمام یا قسمتی از این اثر را بدون اجازه مولف یا ناشر تکثیر و کپی برداری نماید، مورد پیگرد قانونی قرار خواهد گرفت.

فهرست مطالب

۵.....	مقدمه ناشر
۱۵.....	مقدمه
۱۷.....	صل اول / معرفی RNA
۱۷.....	۱-۱- RNA
۱۸.....	۲-۱- مقایسه RNA با DNA
۱۸.....	۳-۱- ساختمان
۲۰.....	۴-۱- طول RNA
۲۱.....	۵-۱- سنتز RNA
۲۱.....	۶-۱- ترجمه
۲۲.....	۷-۱- RNAهای تنظیمی
۲۳.....	۸-۱- پردازش RNA
۲۴.....	۹-۱- ژنوم های RNA
۲۴.....	۱۰-۱- رونویسی وارونه
۲۵.....	۱۱-۱- RNA دورشته ای
۲۵.....	۱۲-۱- کشف های کلیدی در زیست شناسی RNA
۲۹.....	صل دوم / معرف داده های امیکس و روش های آنالیز
۳۰.....	۱-۲- امیکس
۳۱.....	۲-۲- ژنوم و ژنومیکس
۳۱.....	۱-۲-۲- خصوصیات ژنوم و گستره های ژنتیکی
۳۲.....	۲-۲-۲- فناوری ها و روش های آنالیز ژنتیکی
۳۳.....	۳-۲- ترنسکرپتوم و ترنسکرپتومیکس

۳۳	۲-۳-۱- فناوری‌ها و روش‌های آنالیز ترنسکرپتوم
۳۴	۲-۳-۲- معرفی روش‌های بررسی ترنسکرپتوم
۳۵	۲-۳-۱- نوردن بلات
۳۷	۲-۳-۲- Real-time PCR
۳۹	۲-۳-۳- روش ریزآرایه
۴۱	۲-۳-۴- تجزیه و تحلیل سریالی بیان ژن (SAGE)
۴۳	۲-۵-۱- روش MPSS
۴۳	۲-۳-۶- RNA-seq
۴۴	۲-۴-۱- پروم و پرومیکس
۴۵	۲-۴-۱- داده‌های کلی
۴۶	۲-۵-۱- کاربرد آمیکس
۴۶	۲-۵-۱- آمیکس‌ها در تحقیقات دارویی
۴۶	۲-۵-۲- آمیکس در پزشکی
۴۷	۲-۵-۳- کاربردهای فناوری ریزآرایه
۴۷	۲-۵-۴- آشنایی با مزیت‌ها و کاربردهای RNA-Seq
۴۷	۲-۵-۵- مقایسه روش RNA-Seq با روش میکروآرایه
۴۷	۲-۵-۵-۱- مزیت‌های روش RNA-Seq در مقایسه با میکروآرایه
۴۸	۲-۵-۵-۲- معایب روش RNA-Seq
۴۸	۲-۵-۶- آشنایی با کاربردهای داده‌های حاصل از روش RNA-Seq
۴۸	کاربرد تکنیک RNA-Seq در نوسازی ترانسکرپتوم ترانسکرپتوم
۴۹	کاربرد تکنیک RNA-Seq در سنجش بیان ژن و واریانت آنها
۵۰	کاربرد تکنیک RNA-Seq در آنالیز بیان ژن‌ها در نمونه‌های مختلف
۵۱	فصل سوم / توالی‌یابی RNA
۵۱	۳-۱- ترنسکرپتومیکس و RNA-seq
۵۱	۳-۲- RNA-Seq

۵۲	۳-۳- روند کار RNA-seq
۵۳	۳-۳-۱- طراحی تجربی
۵۴	۳-۳-۲- شناسایی هدف اولیه آزمایش
۵۵	۳-۴- حاشیه نویسی
۵۷	۳-۵- بررسی بیان ژن هایی که در بین نمونه ها دارای تفاوت بیان هستند (DGE)
۵۸	۳-۵-۱- تکثیر و بررسی DGE
۵۸	۳-۵-۲- مقایسه تکرار فنی و زیستی
۵۹	۳-۵-۳- چند تکرار باید توالی یابی شود؟
۶۰	۳-۵-۴- عمود توالی یابی
۶۱	۳-۵-۵- پیچیدگی های آزمایشی
۶۱	۳-۵-۶- ویژگی های رونویسی هدف
۶۲	۳-۵-۷- اندازه اثر
۶۲	۳-۶- نبود جایگزین برای داده های پایدار (ایتنایی)
۶۵	فصل چهارم / استخراج RNA
۶۵	۴-۱- آماده سازی RNA
۶۶	۴-۲- کار با RNA
۶۶	۴-۳- اقدامات احتیاطی هنگام کار با RNA
۶۷	۴-۴- پایدار سازی RNA
۶۹	۴-۵- جداسازی و خالص سازی RNA
۶۹	۴-۵-۱- محلول سازی
۷۰	۴-۵-۲- همگن سازی مکانیکی
۷۰	۴-۵-۳- نکات ملاحظات نوع بافت
۷۰	۴-۵-۳-۱- بافت عضلانی و پوست
۷۱	۴-۵-۳-۲- بافت های غنی از چربی
۷۱	۴-۵-۴- بازیابی RNA از لیزات

- ۷۱..... ۴-۵-۴-۱- استخراج آلی
- ۷۲..... ۴-۵-۴-۲- استخراج فاز جامد
- ۷۳..... ۴-۵-۴-۳- نکات مقایسه روش استخراج آلی با روش فاز جامد
- ۷۳..... ۴-۵-۴-۴- در خصوص RNA های کوچک
- ۷۳..... ۴-۵-۴-۵- حذف آلودگی های ایجاد شده توسط DNA ژنومی
- ۴-۵-۴-۶- استفاده از ابزارهای شیمیایی برای حل مشکل آنالیز و شناخت
- ۷۴..... فرایندهایی زیستی
- ۷۴..... ۱-۵-۵- قواعد کروماتوگرافی و استخراج فاز جامد
- ۷۴..... ۴-۵-۱- قواعد کروماتوگرافی
- ۷۸..... ۴-۵-۲- قواعد استخراج فاز جامد
- ۷۹..... ۴-۵-۵-۲- روش استخراج ارگانیک
- ۸۰..... ۴-۵-۵-۴- روش spin blot براساس فیلتر
- ۸۰..... ۴-۵-۵-۵- روش magnetic particle
- ۸۱..... ۴-۵-۵-۶- روش direct lysis
- ۸۲..... ۴-۵-۵-۷- ارزیابی کمی و کیفی
- ۸۷..... ۴-۵-۹- spectroscopy
- ۹۰..... ۴-۶-۶- غنی سازی توالی هدف
- ۹۰..... ۴-۶-۱- انتخاب mRNAs بالغ با کمک انتهای آن
- ۹۱..... ۴-۶-۲- غنی سازی SuperSAGE برحسب برجسب سه پرمی mRNA
- ۹۱..... ۴-۶-۳- حذف توالی ریبوزومی از طریق هیبریدیزاسیون
- ۴-۶-۴- نرمال سازی تعداد توالی های تکراری از طریق استفاده از آنزیم نوکلئاز
- ۹۲..... مخصوص دوسویه (DSN)
- ۹۲..... ۴-۶-۵- غنی سازی توالی هدف از طریق انتخاب اندازه
- ۹۲..... ۴-۶-۷- تجزیه و تکه تکه کردن RNA
- ۹۳..... ۴-۷-۱- روش آنزیمی

۹۳	۴-۷-۲- استفاده از یون‌های فلزی
۹۴	۴-۷-۳- استفاده از گرما
۹۵	فصل پنجم / آماده‌سازی کتابخانه
۹۵	۵-۱- سنتز اولین رشته
۹۶	۵-۱-۱- بررسی بیشتر ساخت اولین رشته
۹۷	۵-۱-۲- آماده‌سازی رشته اول با استفاده از oligo-dt
۹۷	۵-۱-۳- آماده‌سازی رشته اول با استفاده از oligos تصادفی
۹۸	۵-۱-۴- سنتز رشته اول با استفاده از oligo پیش از مرحله اتصال
۹۸	۵-۲- سنتز رشته دوم
۹۸	۵-۲-۱- بررسی بیشتر سنتز دومین رشته
۹۸	سنتز دومین رشته از طریق جابه‌جایی RNA
۹۹	سنتز دومین رشته با استفاده از oligoهای متصل به انتهای پنج پریم RNA الگو
۹۹	سنتز دومین رشته با استفاده از oligo-Dg و template switching (SMART)
۹۹	۵-۳- شکستن cDNA
۱۰۰	۵-۳-۱- بررسی بیشتر خرد کردن cDNA
۱۰۰	۵-۴- آداپتورهای توالی‌یابی
۱۰۲	۵-۴-۱- تسهیم نمونه‌ها با استفاده از بارکد / اندیس
۱۰۲	۵-۴-۲- توالی‌یابی دوطرفه
۱۰۲	۵-۵- افزودن آداپتورها
۱۰۳	۵-۵-۱- بررسی بیشتر آداپتورها
۱۰۳	۵-۵-۱-۱- اضافه شدن آداپتور از طریق PCR آداپتور Y
۱۰۳	۵-۵-۱-۲- افزودن آداپتور از طریق RT / PCR
۱۰۴	۵-۵-۱-۳- افزودن آداپتورها از طریق اتصال
۱۰۴	۵-۶- آماده‌سازی کتابخانه‌های stranded
۱۰۵	۵-۶-۱- بررسی بیشتر پروتکل‌های مربوط به stranded

- ۱۰۵..... ۷-۵- اعتبارسنجی و کمی‌سازی
- ۱۰۷..... فصل ششم / توالی‌یابی
- ۱۰۷..... ۱-۶- انتخاب پلت‌فرم توالی‌یابی
- ۱۱۰..... ۱-۱-۶- تکنولوژی Illumina / Solexa
- ۱۱۲..... ۲-۱-۶- تکنولوژی ABI / SOLiD
- ۱۱۳..... ۳-۱-۶- روش تعیین توالی نیمه‌رسانا Ion Torrent
- ۱۱۵..... ۴-۱-۶- SMRT
- ۱۱۶..... ۵-۱-۶- توالی‌یابی نانوپور
- ۱۱۶..... ۵-۱-۶- مزیت‌های فناوری تعیین توالی نانوپور
- ۱۲۰..... ۲-۶- آماد سازش اراده نمونه
- ۱۲۱..... فصل هفتم / آنالیز
- ۱۲۲..... ۱-۷- پردازش اولیه توالی
- ۱۲۲..... ۱-۱-۷- تفکیک نمونه‌ها براساس اکتاها
- ۱۲۲..... ۲-۱-۷- حذف آداپتورها
- ۱۲۳..... ۳-۱-۷- فیلترکردن و ترمیم کردن خوانده‌ها براساس کیفیت
- ۱۲۴..... ۴-۱-۷- فیلترکردن و نرمال‌سازی خوانده‌ها براساس میزان K-mer
- ۱۲۵..... ۲-۷- روش مونتاژ de novo ترنسکریپتوم
- ۱۲۶..... ۱-۲-۷- مونتاژ مبتنی بر گراف‌های de Bruijn
- ۱۲۷..... ۲-۲-۷- ابزارهای رایج مبتنی بر گراف برای مونتاژ
- ۱۲۷..... Velvet / Oases-۱-۲-۲-۷
- ۱۲۷..... Trans-AbySS-۲-۲-۲-۷
- ۱۲۷..... Trinity-۳-۲-۲-۷
- ۱۲۸..... ۴-۲-۲-۷- روش مونتاژ مبتنی بر OLC (Overlap - layout - consensus)
- ۱۲۸..... ۳-۲-۷- توضیحات بیشتر در خصوص مونتاژ OLC
- ۱۲۹..... ۴-۲-۷- نرم‌افزارهای رایج برای مونتاژ OLC

۱۲۹	 CAP3 روش ۱-۴-۲-۷
۱۲۹	 MIRA 3 روش ۲-۴-۲-۷
۱۲۹	 Roche GS De Novo Assembler ("Newbler") روش ۳-۴-۲-۷
۱۲۹	 هم‌ردیف‌سازی خوانده‌های RNA-seq با ژنوم مرجع ۳-۳-۷
۱۳۱	 هم‌ردیف‌سازی با ژنوم مرجع ۱-۳-۷
۱۳۱	 هم‌ردیف‌سازی با ترنسکریپتوم ۲-۳-۷
۱۳۲	 microRNA ابزارهای هم‌ردیف‌سازی برای ۳-۳-۷
۱۳۲	 ۴-۱-۷ فایل خروجی حاصل از هم‌ردیف‌سازی خوانده‌های کوتاه
۱۳۳	 ۴-۷ حاشیه نویسی و رونوشت‌ها
۱۳۳	 ۵-۷ بررسی تفاوت بیان ژن‌ها و ایزوفرم‌های ژنی
۱۳۴	 ۱-۵-۷ نرم‌افزارهای تعدیل خوانده‌ها
۱۳۵	 ۲-۵-۷ آنالیز تفاوت بیان ژن با استفاده از مدل‌های توزیع گسسته
۱۳۵	 ۳-۵-۷ آنالیز تفاوت بیان ژن با استفاده از مدل‌های توزیع پیوسته
۱۳۶	 ۴-۵-۷ آنالیز تفاوت بیان ژن با استفاده از مدل‌های غیرپارامتری
۱۳۶	 ۵-۵-۷ انتخاب نرم‌افزارهای آنالیز داده‌های بیان ژن
۱۳۹	 فصل هشتم / واژه‌نامه
۱۴۱	 منابع

مقدمه

با توجه به اهمیت بسزای بهداشت، سلامت و افزایش کیفیت زندگی و چالش‌های جهانی پیش‌رو در حوزه سلامت، گسترش روزافزون بسیاری از بیماری‌های نوپدید ویروسی و ژنتیکی و سرطان‌ها، به منظور دستیابی به اطلاعات جدید در خصوص بروز تغییر و تحولات در موجودات زنده اعم از انسان و حیوان و نیز میکروارگانیسم‌ها، نیازمند روش‌های نوین علمی هستیم. اهمیت به کارگیری روش‌های نوین علمی در وجود روابط بسیار پیچیده در میان ارگانیسم‌ها و تعاملات گوناگون و متغیر آنهاست. از این رو، فهم تعاملات هرچند جزئی ولی با اهمیت موجود در دنیای اطراف نیازمند فناوری‌های بسیار دقیق و سریع است. در واقع با استفاده از فناوری و نیز علوم نوین، می‌توان اقداماتی به منظور کنترل انواع بیماری‌ها به واسطه واکسن‌ها و نیز داروهای جدید انجام داد. از اهداف اصلی زیست‌شناسی در زمانه‌ای در مطالعه و دستیابی به رهیافت‌های اصلی مبارزه با عوامل بیماری‌زا در این زمینه، در این زمینه، تولید واکسن‌های جدید یا بهینه‌سازی واکسن‌های گذشته، شناسایی شاخص‌های تشخیصی (Diagnostic Markers)، شناسایی اهداف دارویی کارآمد و جایگزین (Drug Targets & Drug Discovery) با استفاده از ژنومیک و پروتئومیک است. با توجه به رابطه موجود میان ژن‌ها و واحدهای عملکردی تولیدشده توسط آنها (پروتئین‌ها)، از طریق رونویسی mRNA و بیان پروتئین‌ها و با در نظر گرفتن نیمه عمر کوتاه و متفاوت mRNA رسیدن به اهداف مذکور از طریق آنالیز پروفایل ترنسکرپتومی به وسیلهٔ دو تکنیک ریزآرایه^۱ و (RNA-seq) و به همراه پروفایل محتوای پروتئینی (پروتئوم) از طریق دو تکنیک الکتروفورز دوبعدی و طیف‌سنجی جرمی^۲

1. Microarray

2. Mass spectrometry

امکان‌پذیر است. در این‌گونه مطالعات همراهی بررسی‌های تجربی و میدانی با تکنیک‌های بیوانفورماتیک و بررسی‌های *In Silico*، دانش و تجربه بسیاری را در اختیار قرار می‌دهد. تنظیمات ترانسکرپتومیک برای فرایندهای فیزیولوژیک، پاتولوژیکی و رشد و نمو سلول حیاتی است. پروفایل بیان mRNA می‌تواند ویژگی‌های سلول را در حالت خاص نشان دهد و به فعالیت‌های حال حاضر و آینده سلول کمک کند. اهداف ترانسکرپتومیکس عبارت‌اند از: جمع‌آوری انواع نسخه‌برداری‌ها، شامل mRNA، non-coding RNA و Small RNA تا بتواند ساختار نسخه‌برداری ژن‌ها را تعیین کند؛ به عبارت دیگر، جایگاه‌های شروع ۵' و انتهای ۳'، الگوها، برش (splicing patterns) و تغییرات پس از نسخه‌برداری (post-transcriptional modifications) و همچنین مقدار سطوح بیان هر نسخه‌برداری را در طی رشد و توسعه آن و در شرایط مختلف تعیین رسی کند. از طرف دیگر، رمزگشایی توالی DNA برای همه شاخه‌های علوم زیستی امری حیاتی است. دانشمندان با ظهور توالی‌یابی مبتنی بر الکتروفورز مویرگی سنگر، توانستند اطلاعات ژنتیکی در سیستم بیولوژیکی را مشخص کنند. این فناوری به‌طور گسترده در آزمایشگاه‌های سراسر دنیا به کار گرفته شده است.

اما محدودیت‌های داخلی در توان عملیاتی، مقیاس‌پذیری، سرعت و تفکیک‌پذیری، مانعی است که از دسترسی دانشمندان به داده‌های اساسی مورد نیاز جلوگیری می‌کند. برای غلبه بر این موانع، فناوری جدید توالی‌یابی نسل جدید (NGS) مورد نیاز است. این فناوری اساساً رویکرد متفاوتی در توالی‌یابی است که کمالات متعدد را سبب شده و تحولی را در علوم ژنومیک ایجاد کرده است. در کتاب پیش‌رو، ما به معرفی جامعی از انواع RNA و خلاصه‌ای از روش‌های بررسی بیان ژن و به‌طور تفصیلی در خصوص یکی از جدیدترین روش‌های بررسی بیان ژن یعنی تکنیک RNA-SEQ پرداخته‌ایم.