

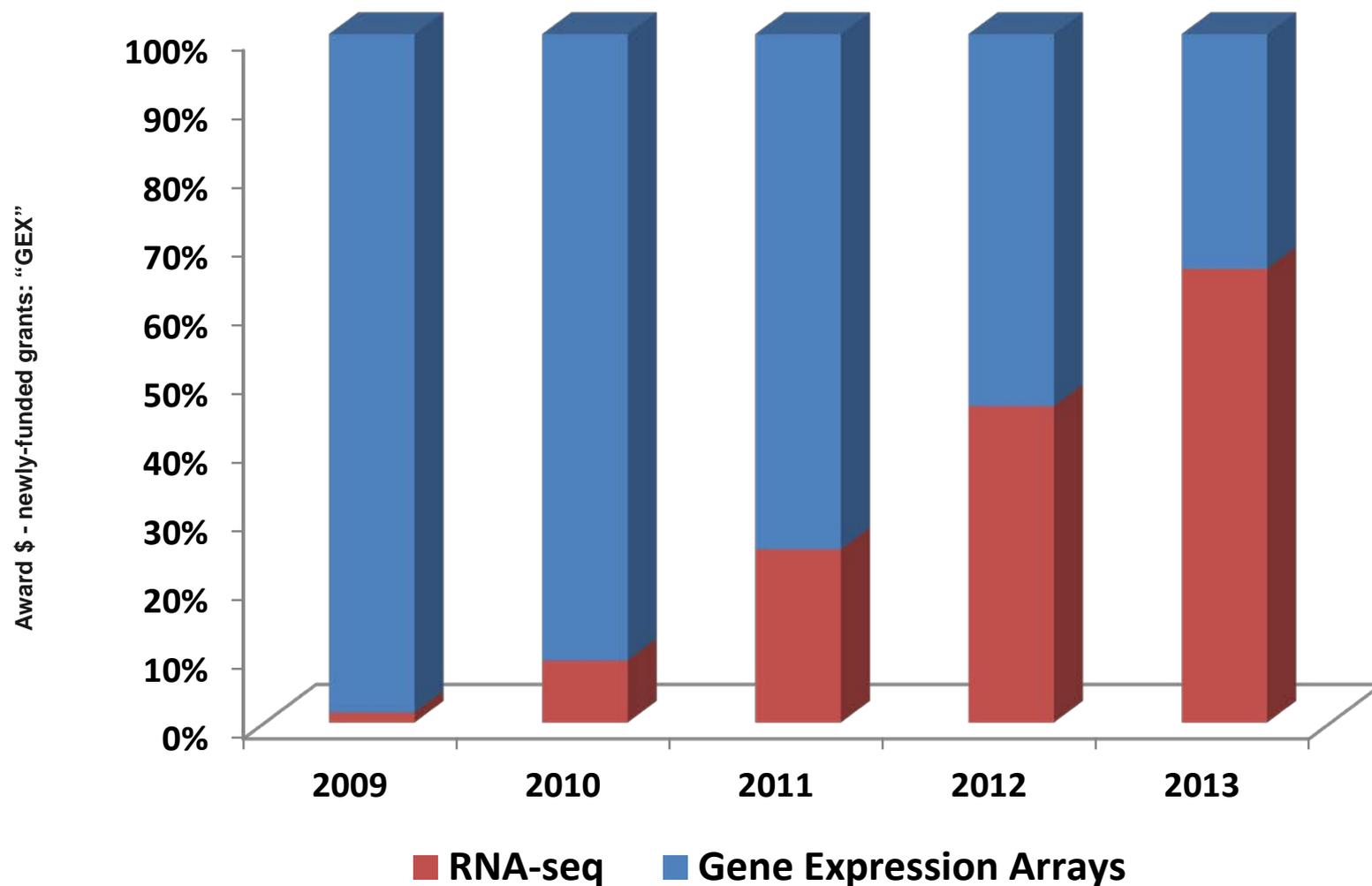
イルミナウェビナー
NextSeq 500シリーズ

RNA Seq時代到来： NextSeqが実現する 簡単、高速、安価な トランスクリプトーム解析

2014年9月16日
イルミナ株式会社
シーケンシング
スペシャリスト
鈴木 健介

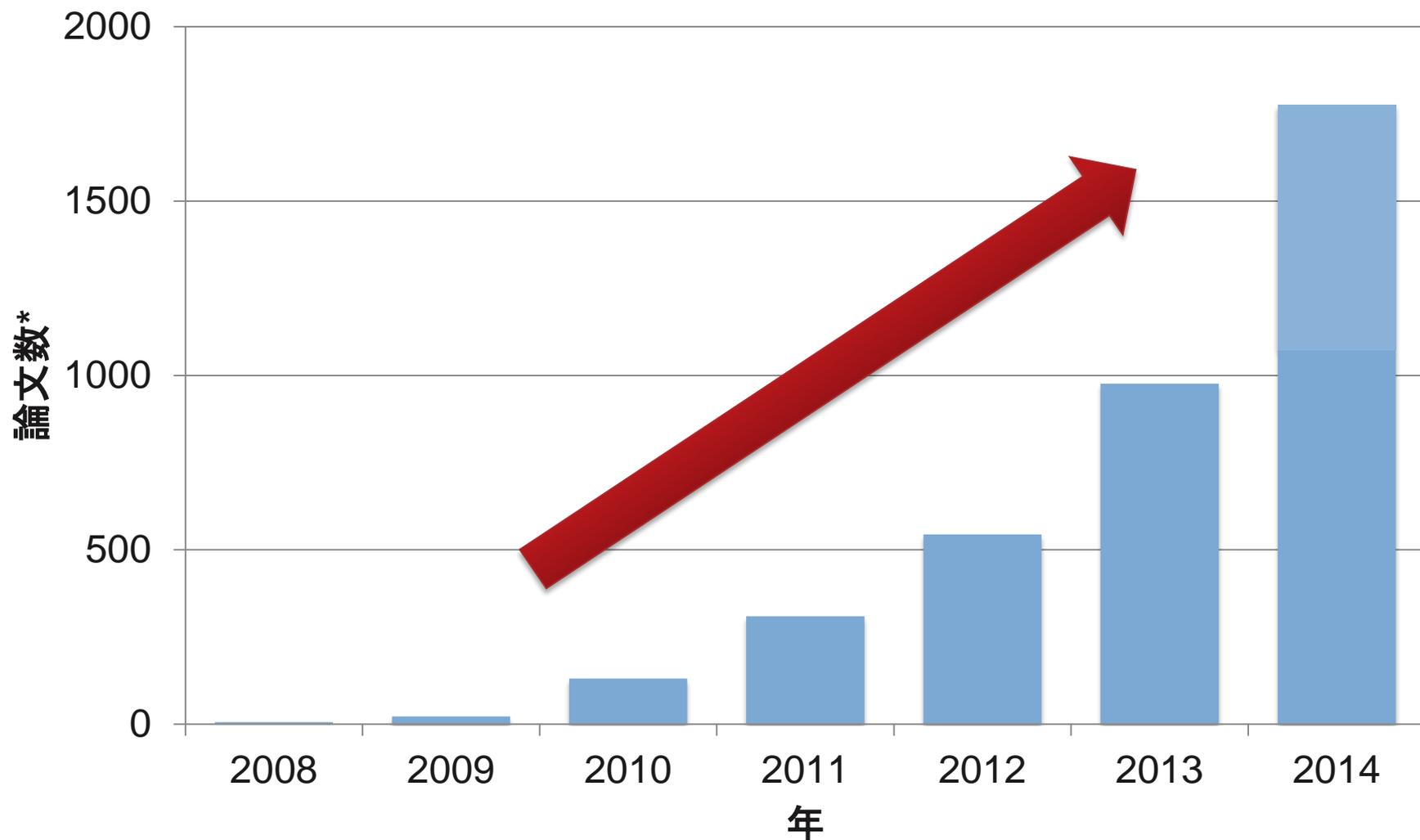


研究トレンドはRNA Seqに移行



※ NIHデータベースから新規遺伝子発現解析のグラントがRNA-Seqか発現アレイかを集計

急速に広がる RNA Seq 手法



* NCBI Pubmed を RNA-Seq で検索した結果を年ごとに集計。

本日の内容

- ▶ 今なぜRNA Seq?
 - マイクロアレイの原理と限界
 - RNA Seqの原理と長所
- ▶ NextSeq 500が実現する簡単、高速、安価なRNA Seq
 - NextSeq 500紹介
 - RNA Seqワークフロー
 - RNA Seqコスト
- ▶ 更に詳しく知りたい方へ

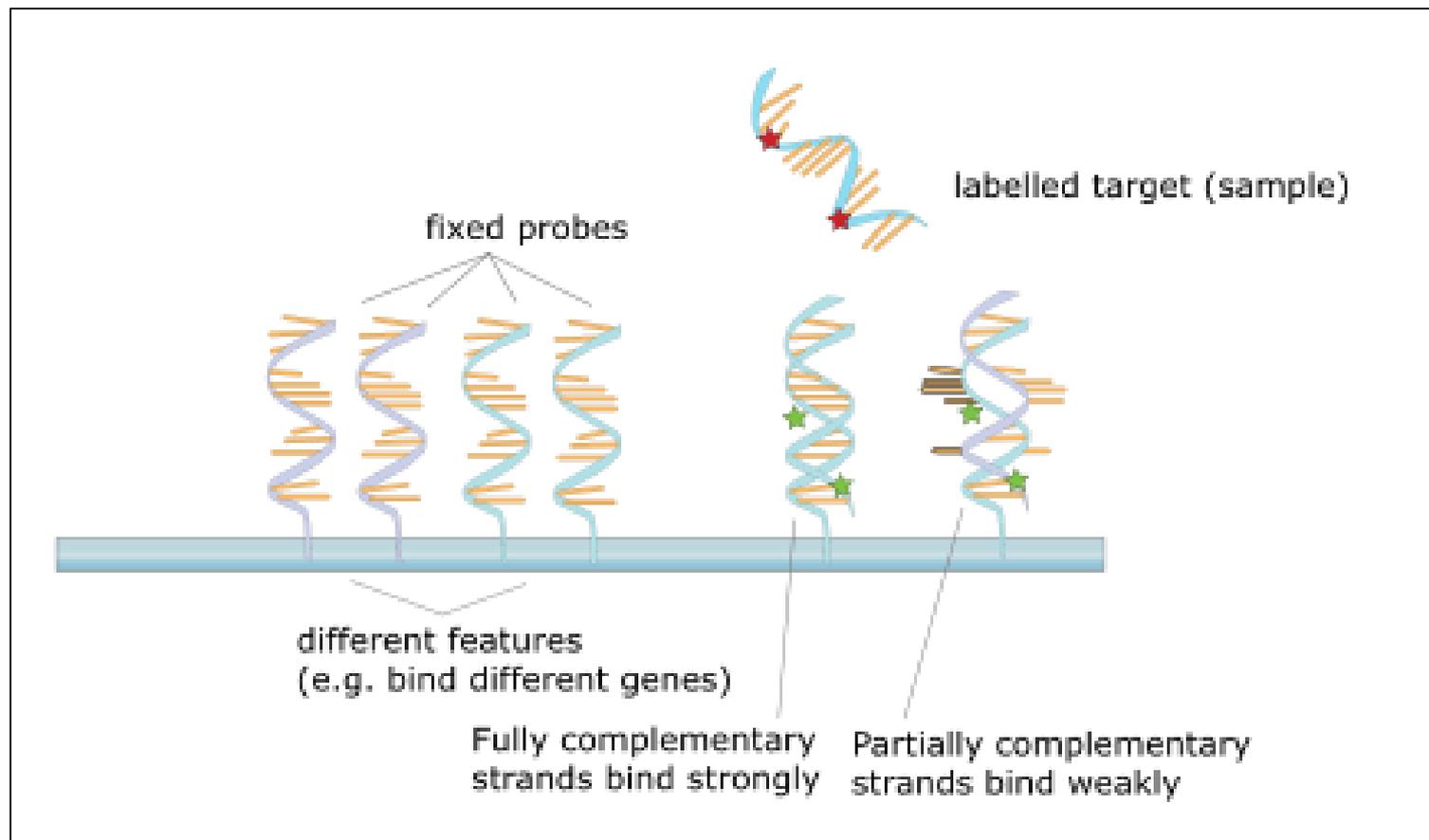
マイクロアレイの原理と限界



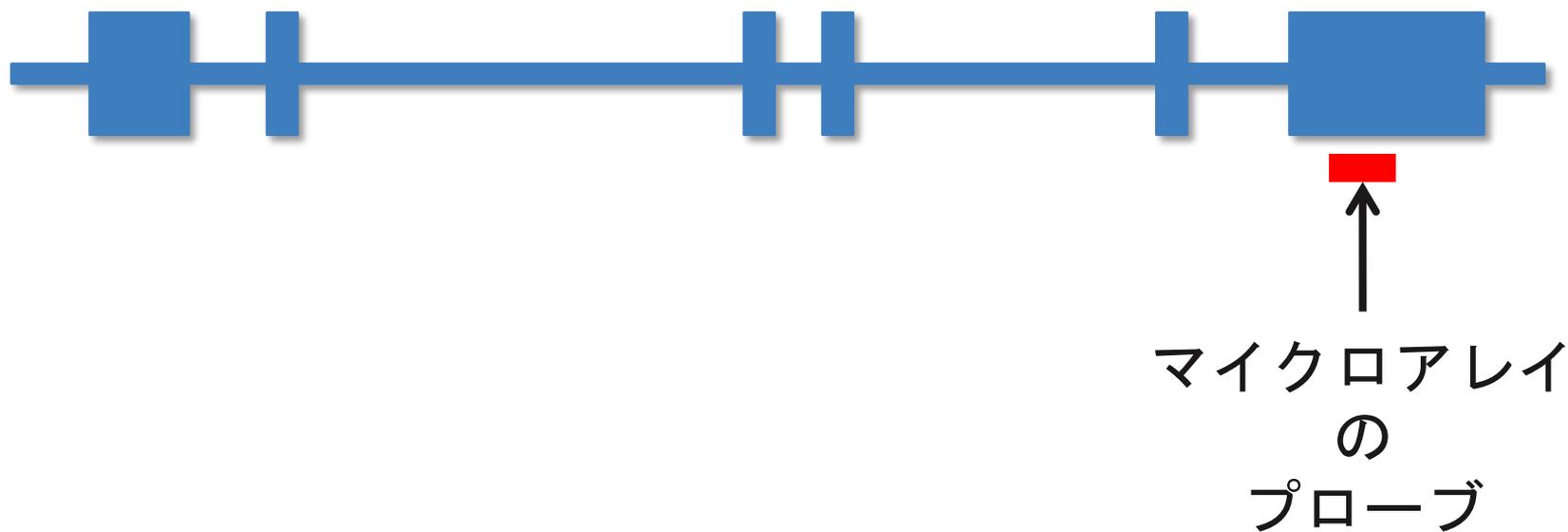
DNA microarray

From Wikipedia, the free encyclopedia

http://en.wikipedia.org/wiki/DNA_microarray



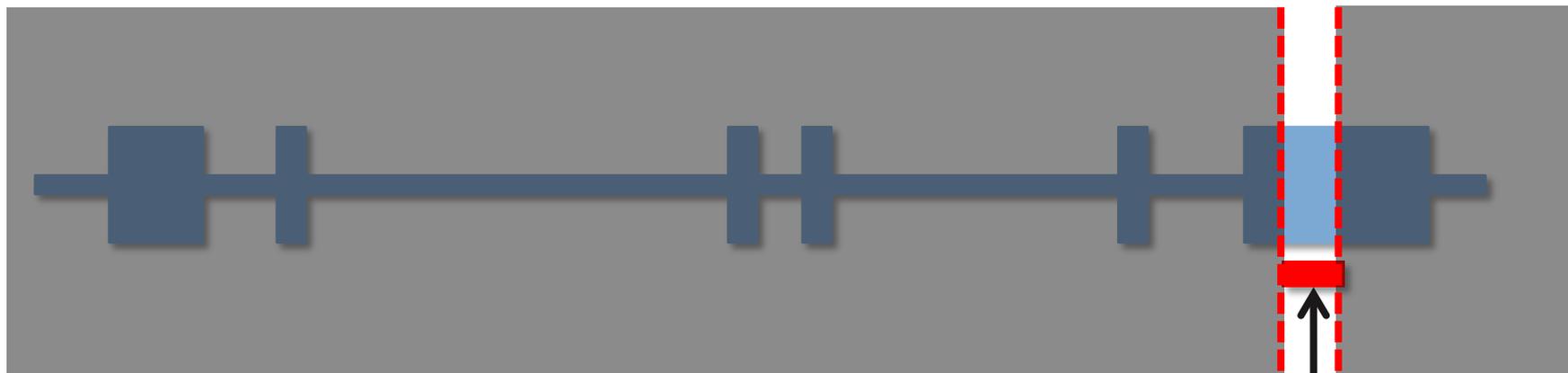
マイクロアレイの原理と限界



マイクロアレイ

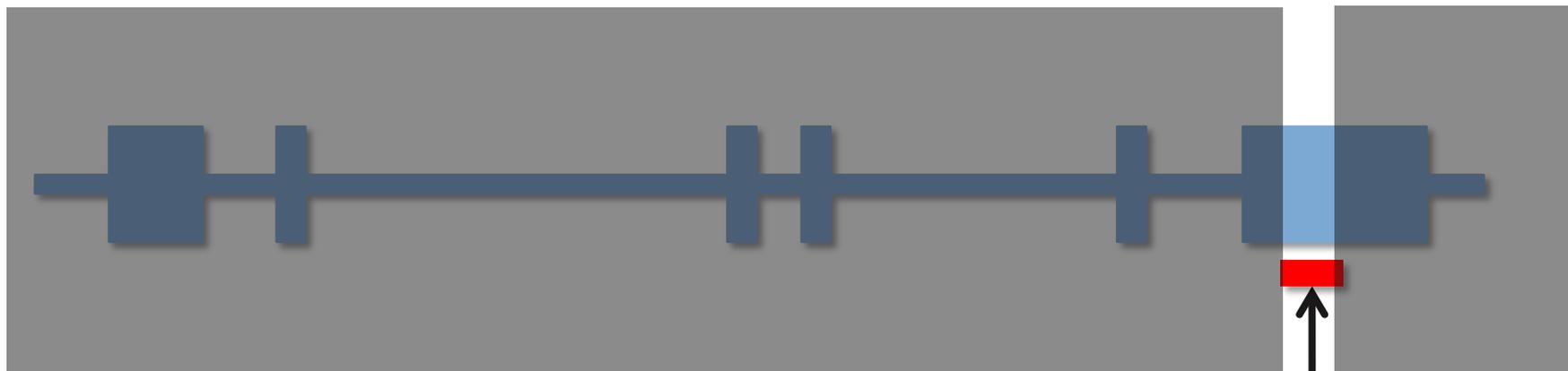
- 遺伝子の一部の配列からプローブを設計
- サンプルとハイブリダイゼーションさせ発現量を調べる

マイクロアレイの原理と限界

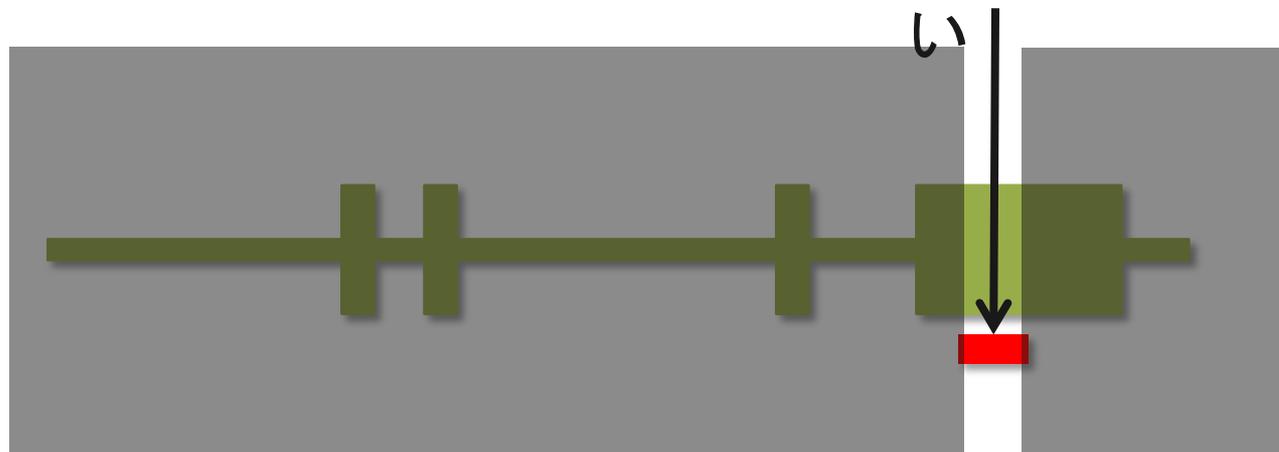


遺伝子の一部
(プローブ設計部
分)
しか見ていない

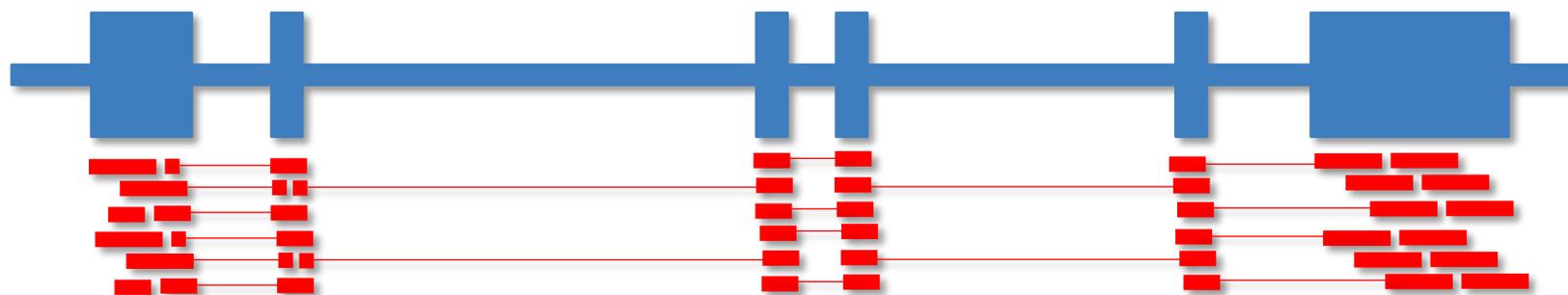
マイクロアレイの原理と限界



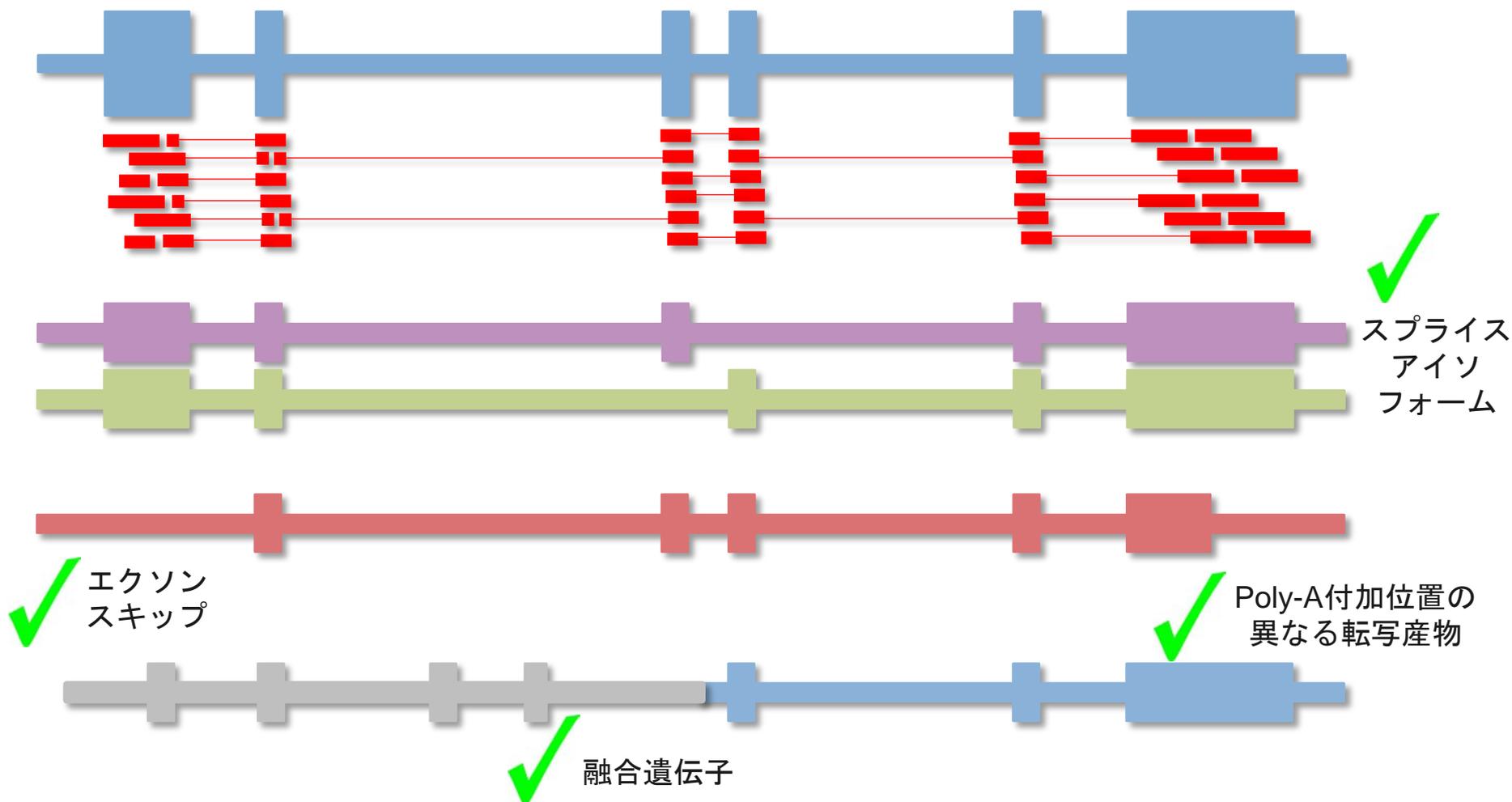
他の遺伝子や偽遺伝子に
同じ配列があると区別できな
い



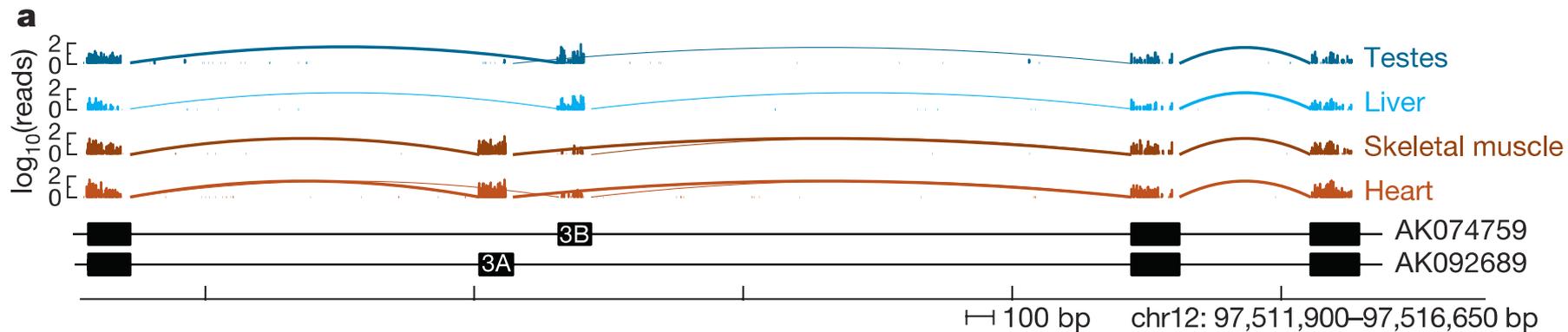
RNA Seqの原理と長所



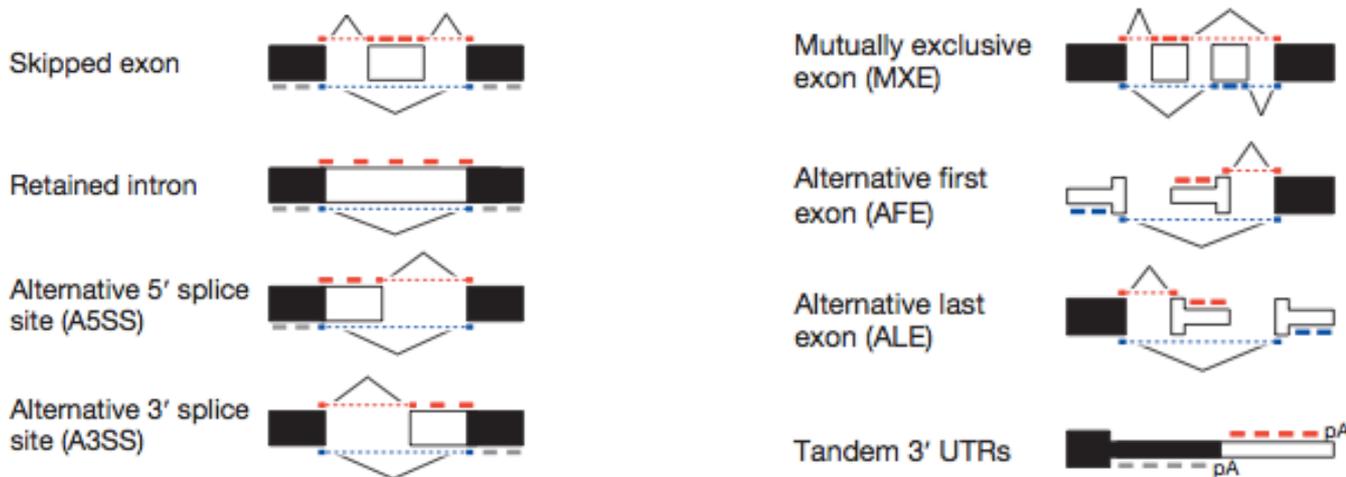
RNA Seqの原理と長所



臓器特異的なスプライスアイソフォームの発現

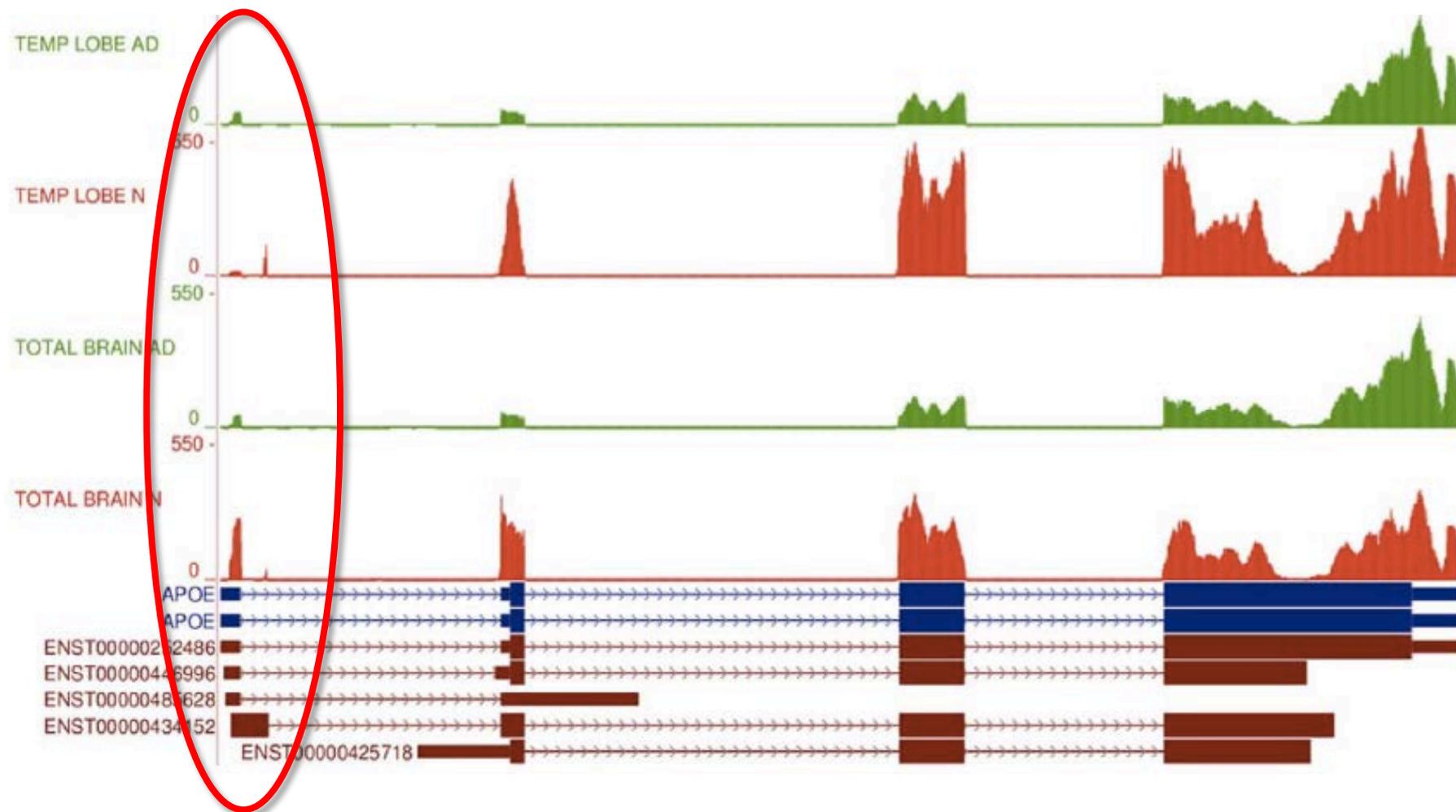


様々なバリエーション



アルツハイマーにおけるスプライスアイソフォーム

- 患者由来の脳RNA Seq : APOE の解析結果



癌において発現する偽遺伝子

Cell

Resource

Expressed Pseudogenes in the Transcriptional Landscape of Human Cancers

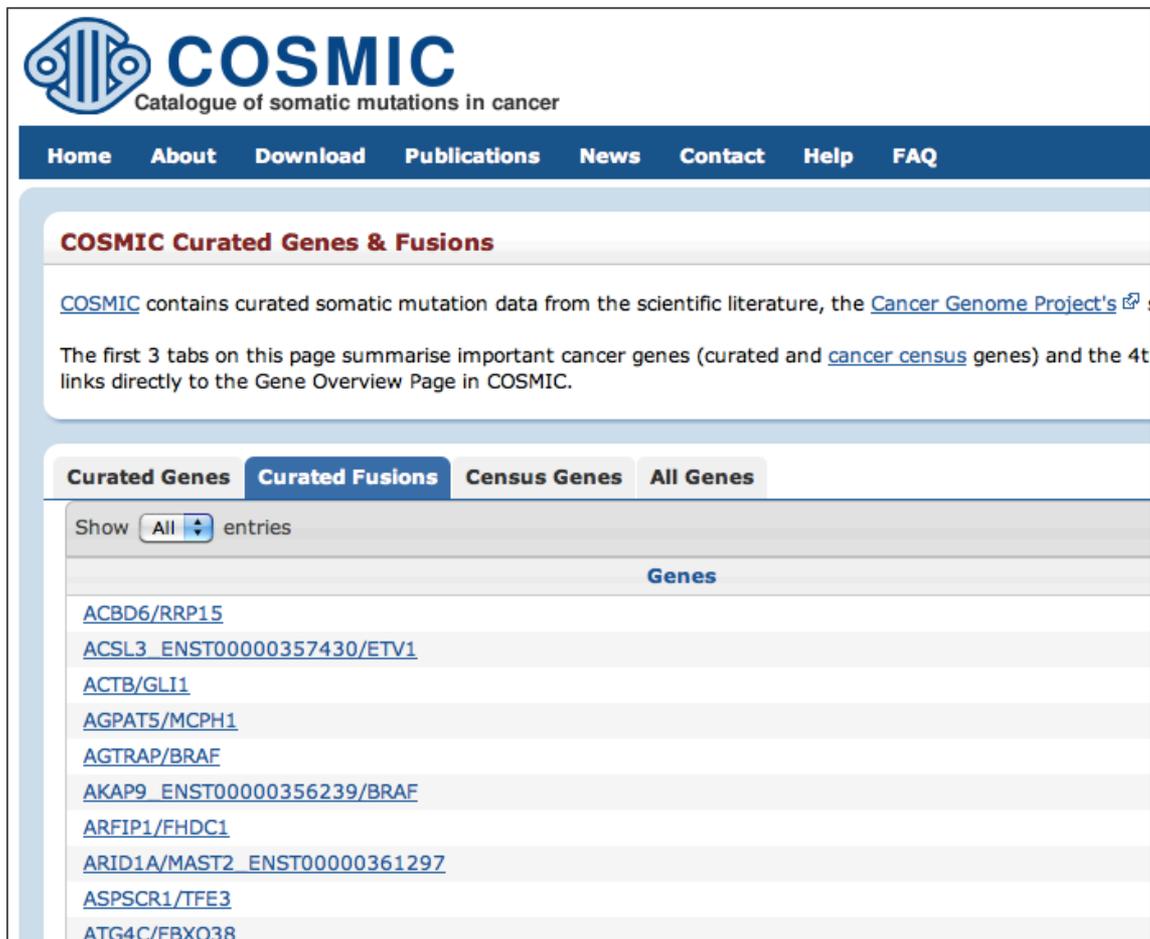
Shanker Kalyana-Sundaram,^{1,2,6,7} Chandan Kumar-Sinha,^{1,2,7} Sunita Shankar,^{1,2} Dan R. Robinson,^{1,2} Yi-Mi Wu,^{1,2} Xuhong Cao,^{1,3} Irfan A. Asangani,^{1,2} Vishal Kothari,¹ John R. Prensner,^{1,2} Robert J. Lonigro,^{1,2} Matthew K. Iyer,¹ Terrence Barrette,^{1,2} Achiraman Shanmugam,⁶ Saravana M. Dhanasekaran,^{1,2} Nallasivam Palanisamy,^{1,2} and Arul M. Chinnaiyan^{1,2,3,4,5,*}

- ▶ 13種類の癌を対象に 293 サンプルを解析
- ▶ これまでの課題：高い配列の類似性のために偽遺伝子と遺伝子の発現を分けて解析することは困難
- ▶ 本論文：RNA Seqにより、偽遺伝子を区別して発現を解析
 - 癌種特異的に発現している偽遺伝子の発現を観察
 - 偽遺伝子の発現は転写産物全体に重要な影響を与え、細胞間の違いや癌の進行に重要な役割を果たしていると考え

癌の融合遺伝子

COSMICデータベースの *Curated Fusions* で 208 の登録 (2014年9月10日時点)

<http://cancer.sanger.ac.uk/cancergenome/projects/classic>

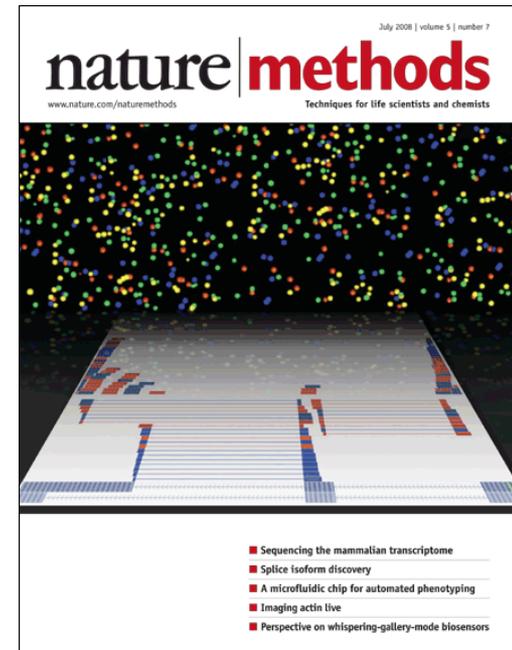


The screenshot shows the COSMIC website interface. At the top left is the COSMIC logo with the tagline "Catalogue of somatic mutations in cancer". A navigation bar contains links for Home, About, Download, Publications, News, Contact, Help, and FAQ. The main heading is "COSMIC Curated Genes & Fusions". Below this, a paragraph explains that COSMIC contains curated somatic mutation data from scientific literature and the Cancer Genome Project's data. It notes that the first three tabs summarize important cancer genes, and the fourth tab links to the Gene Overview Page. The "Curated Fusions" tab is selected, showing a list of gene fusion entries. A dropdown menu is set to "All" entries. The list includes:

Genes
ACBD6/RRP15
ACSL3_ENST00000357430/ETV1
ACTB/GLI1
AGPAT5/MCPH1
AGTRAP/BRAF
AKAP9_ENST00000356239/BRAF
ARFIP1/FHDC1
ARID1A/MAST2_ENST00000361297
ASPSCR1/TFE3
ATG4C/FBXO38

RNA Seq で可能なこと

- ✓ スプライスバリエントごとの発現解析
- ✓ 融合遺伝子の検出
- ✓ アリルごとの発現解析
- ✓ 新規転写産物の発見、解析
- ✓ Non coding RNA の解析
- ✓ 超微量からの発現解析
- ✓ 幅広いダイナミックレンジ



July 2008 - Vol 5 No 7
Nature Methods 5 (2008)
The beginning of the end for microarrays?

本日の内容

- ▶ 今なぜRNA Seq?
 - マイクロアレイの原理と限界
 - RNA Seqの原理と長所
- ▶ NextSeq 500が実現する簡単、高速、安価なRNA Seq
 - NextSeq 500紹介
 - RNA Seqワークフロー
 - RNA Seqコスト
- ▶ 更に詳しく知りたい方へ

今後の遺伝子発現解析を大きく変える NextSeq 500

RNA Seqに最適な次世代シーケンサー



プッシュボタンで操作する簡易性



サンプルから答えまでのワークフロー

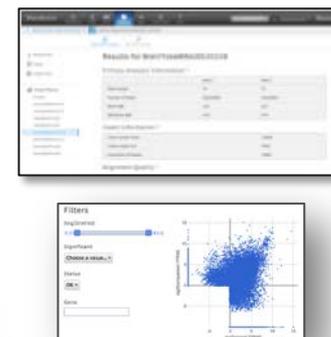
NeoPrep

NextSeq 500

BaseSpace

PCでデータを
閲覧

自動化装置
近日リリース予定



ライブラリー
調製

シーケンス

一次解析

二次解析

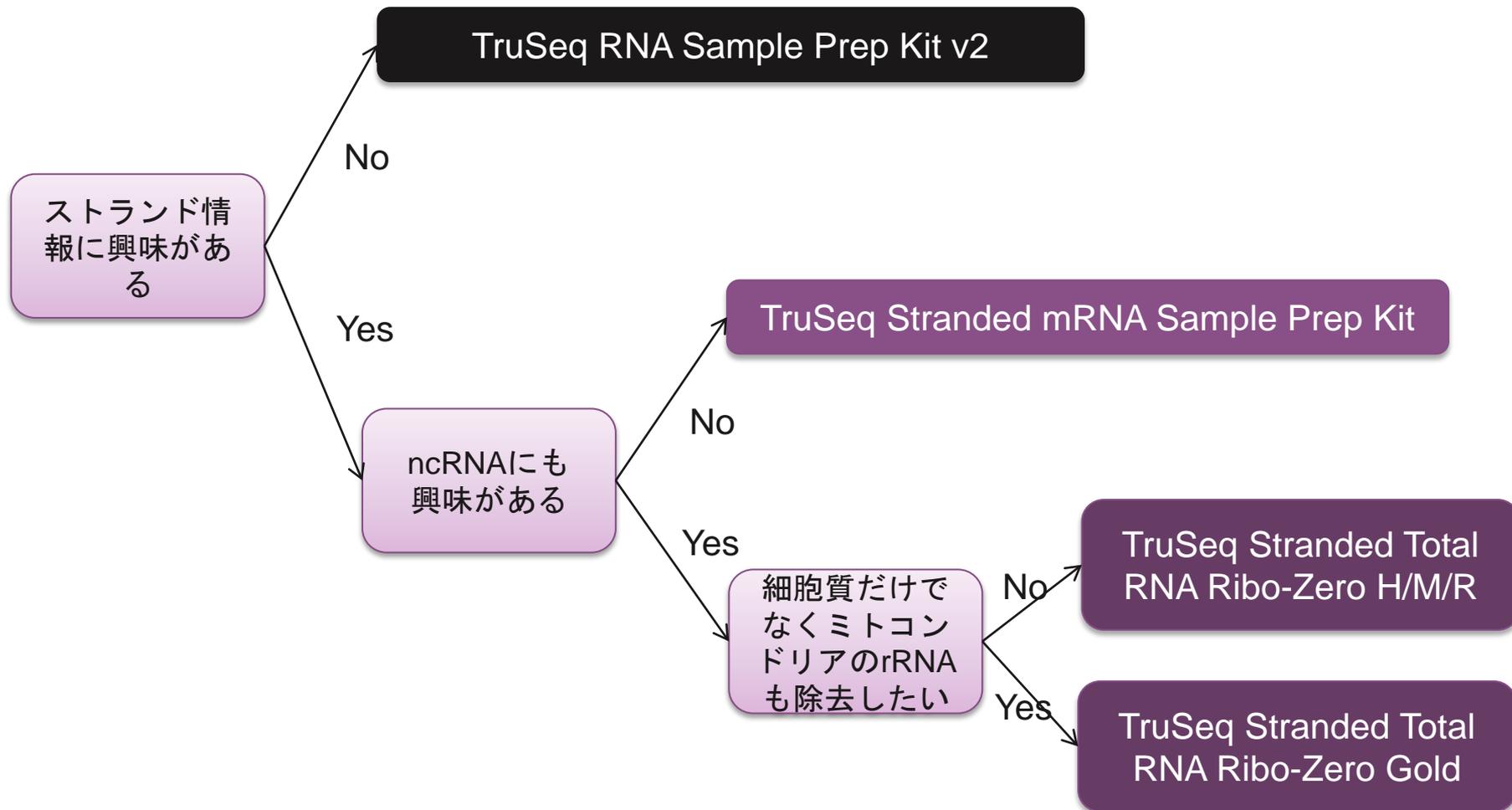
RNA Seq
時間
(ハンズオン時間)

8 時間
(手動 5.5 時間)
(自動 30 分)

18 時間
(10 分)

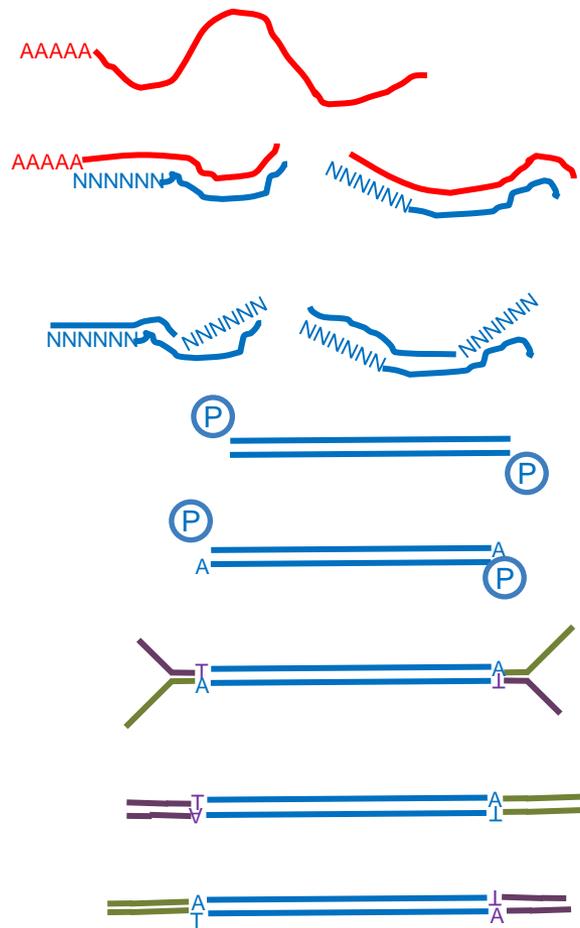
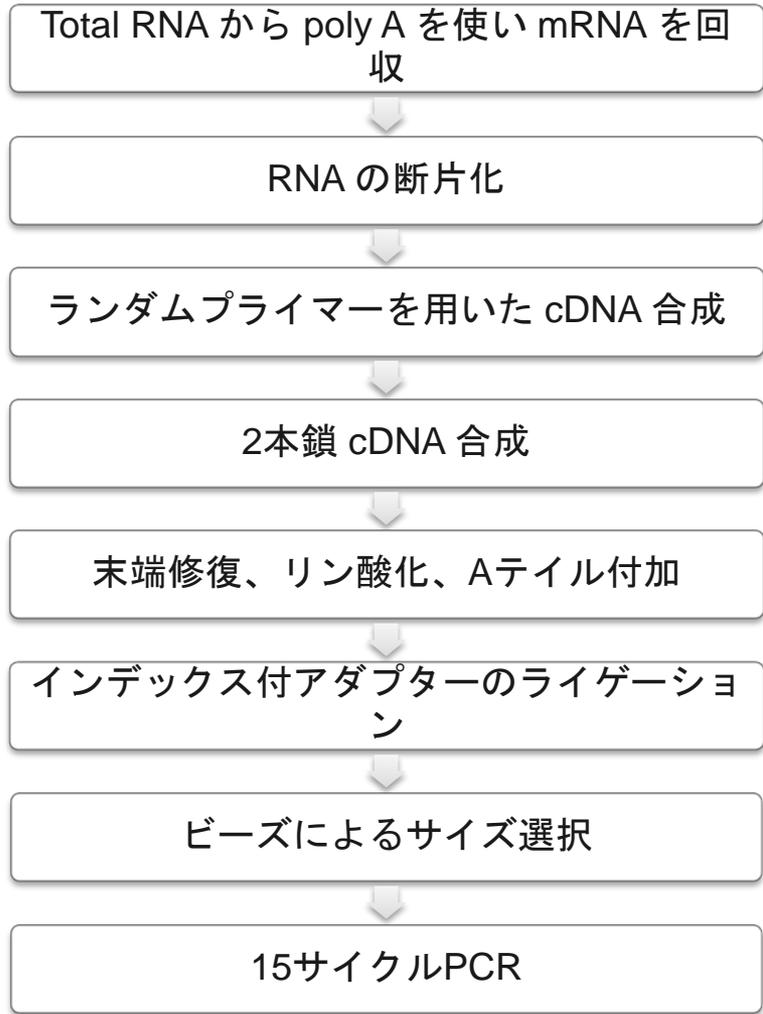
4 時間 / サンプル
(5 分)

RNA Seq : キットの選択



TruSeq RNA Sample Prep Kit ワークフロー

1日



データ解析 : BaseSpace で簡単に !

BaseSpace®

The screenshot displays the BaseSpace web interface. At the top, there's a navigation bar with 'Dashboard', 'Prep', 'Runs', 'Projects', 'Apps', 'Public Data', and 'Help'. Below this, a progress bar shows four steps: 1. Biological Samples, 2. Libraries, 3. Pools, and 4. Planned Runs. The main area is titled 'BaseSpace Applications' and features a grid of application icons including 'BWA Enrichment', 'BWA Whole Genome Sequencing v1', and 'Cufflinks Assembly & DE Beta'. A 'TopHat Alignment Beta' application window is open, showing fields for 'App Session Name', 'Save Results To', 'Samples', and 'Reference Genome' (set to 'Homo sapiens/hg19 (RefSeq)'). It also has 'Options' for 'Call Fusions' and 'Trim TruSeq Adapters'. To the right, a scatter plot shows 'log2(generator FPKM)' vs 'log2(control FPKM)'. Below the scatter plot, there are bar charts and a genomic track visualization with labels A through K.

BaseSpace®
ONSITE



主要アプリケーションを搭載

- 全ゲノム解析
- エクソーム解析
- RNA-Seq解析
- 腫瘍/正常細胞解析

データ解析： ウェブブラウザをクリックすることで 行なう

4つのステップで実施

1. **TopHat** をラン
2. TopHat のQCを実施
– 必要に応じてサンプルを排除
3. **Cufflinks** をラン
– コントロールグループの選択
– 比較グループの選択
4. グループ間を視覚化
– 遺伝子発現の相関性

BaseSpace illumina

TopHat Alignment Beta [Details»](#)

A>T Illumina

App Session Name: TopHat Alignment Beta 12/18/2013 9:27:32 ⓘ

1 Save Results To: Select a Project ⓘ

2 Samples: Select Sample ⓘ

Reference Genome: Homo sapiens/hg19 (RefSeq) ⓘ

Options

Call Fusions: ⓘ

Trim TruSeq Adapters: ⓘ

3 Launch

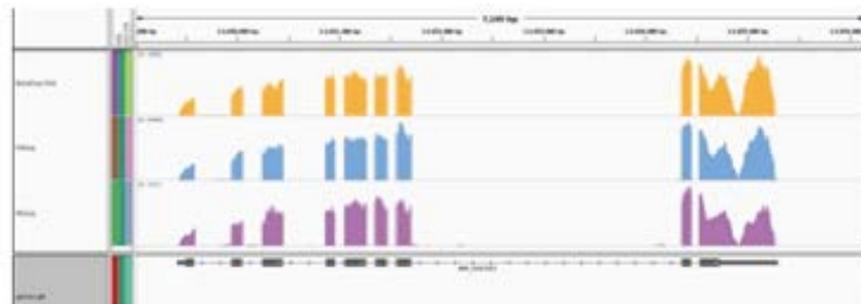
contact us

結果はエクセルとIGVで閲覧可能

tracking_id	class_code	nearest_ref	gene_id	gene_short_name	tss_id	locus	length	coverage	FPKM	FPKM_conf_lo	FPKM_conf_hi	FPKM_status
DDX11L1	-	-	DDX11L1	DDX11L1	TSS14844	chr1:11873-14409	-	-	0.053864	0.0114821	0.103339	OK
WASH7P	-	-	WASH7P	WASH7P	TSS7514	chr1:14361-29370	-	-	7.10972	6.46522	7.7465	OK
FAM138A	-	-	FAM138A	FAM138A	TSS8403	chr1:34610-36081	-	-	0	0	0	OK
OR4F5	-	-	OR4F5	OR4F5	TSS14428	chr1:69090-70008	-	-	0	0	0	OK
LOC729737	-	-	LOC729737	LOC729737	TSS18541	chr1:134772-140566	-	-	5.73767	5.34676	6.05713	OK
LOC100132287	-	-	LOC1001322	LOC1001322	TSS12303	chr1:323891-328581	-	-	0.718542	0.562572	0.877613	OK
LOC100133331	-	-	LOC1001333	LOC1001333	TSS12303	chr1:323891-328581	-	-	6.18633	5.77207	6.61047	OK
OR4F29	-	-	OR4F29	OR4F29	TSS4943	chr1:367658-368597	-	-	0.024205	0	0.0606018	OK
OR4F29	-	-	OR4F29	OR4F29	TSS12680	chr1:621095-622034	-	-	0.024205	0	0.0606018	OK
LOC100133331	-	-	LOC1001333	LOC1001333	TSS13053	chr1:661138-665731	-	-	6.54298	6.03721	7.01381	OK
LOC100288069	-	-	LOC1002880	LOC1002880	TSS8709	chr1:700244-714068	-	-	7.47997	6.14293	7.80318	OK
LINC00115	-	-	LINC00115	LINC00115	TSS18312	chr1:761585-762902	-	-	3.29842	2.57808	3.78791	OK
LOC643837	-	-	LOC643837	LOC643837	TSS1349.TS	chr1:762970-794826	-	-	8.38617	7.95541	8.8435	OK
FAM41C	-	-	FAM41C	FAM41C	TSS20841	chr1:803450-812182	-	-	1.22223	0.922846	1.51213	OK
LOC100130417	-	-	LOC1001304	LOC1001304	TSS20197	chr1:852952-854817	-	-	0	0	0	OK



GAPDH Gene



CALR Gene

TopHat Fusion による融合遺伝子候補のレポート

6. chr11-chr11 rr

B3GAT3	chr11	62387968	GANAB	chr11	62394824	6	149	3	537.38
--------	-------	----------	-------	-------	----------	-------------------	---------------------	---	--------

7. chr9-chr22 rr

ABL1	chr9	133729450	BCR	chr22	23632599	12	4	17	532.63
----------------------	------	-----------	---------------------	-------	----------	--------------------	-------------------	----	--------

RNA Seq による融合遺伝子の検出

Homozygous reference	18,612,849
Heterozygous	58,535
Homozygous variant	1,519
SNV	58,672
Indel	1,397
T _r /T _v	3.00

RNA Seqリードの中に cSNP や InDel を検出

RNA Seq: サンプルあたりのコスト試算

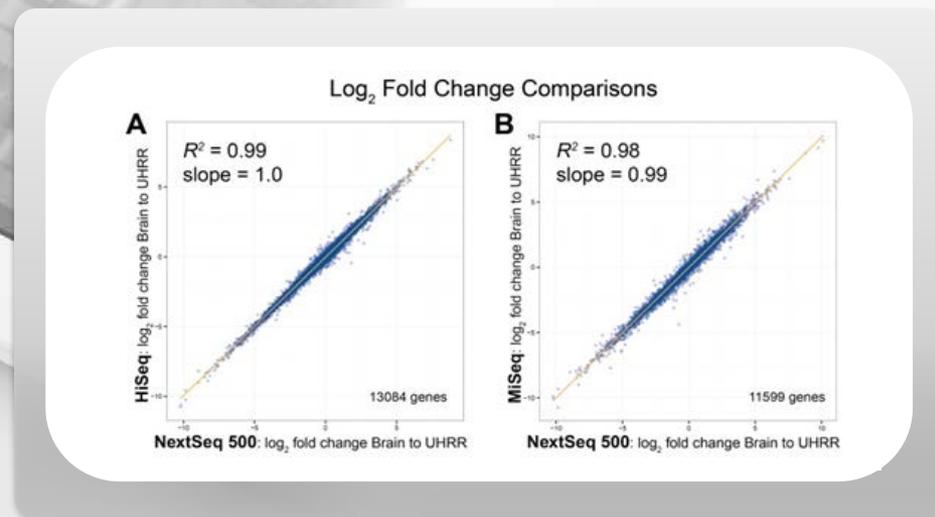
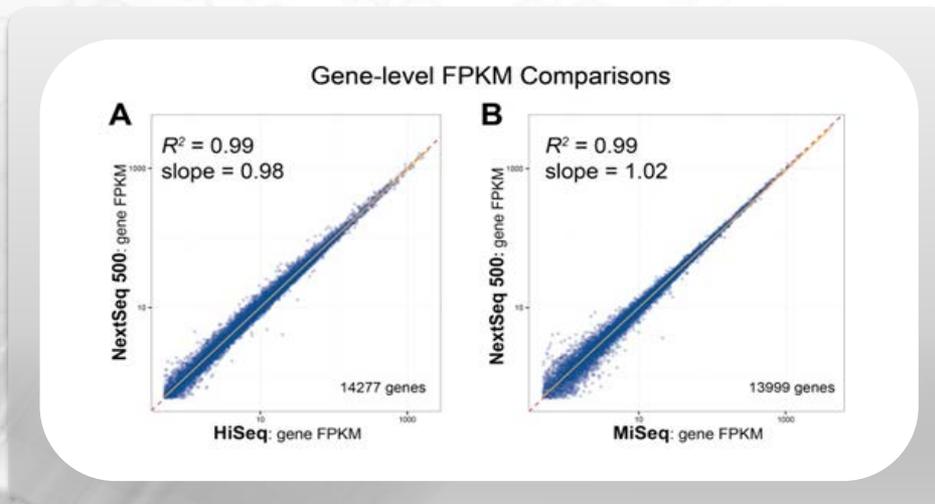
ライブラリー調製およびシーケンスラン試薬

アレイより安価に
解析ができるよう
になりました！

解析目的例	必要リード数	NextSeq サンプルあたりのコスト
遺伝子ごとの発現プロファイル解析 (発現アレイと同等の解析)	1000万	¥ 18,200
転写産物ごとの発現プロファイル解析 (mRNA Seq)	2500万 x2	¥ 36,700
ncRNAも含んだ発現プロファイル解 析 (Total RNA Seq)	5000万 x2	¥ 70,000

高い互換性を示す RNA Seq 結果

HiSeq vs. NextSeq 500



RNA Seq : イルミナ東京ラボでの NextSeq 解析結果

- ▶ タカラバイオ社 市販RNAサンプル 3種 (ヒト心臓、心臓疾患、肝臓)
- ▶ Ambion社 市販RNAサンプル 1種 (ヒト脳)

The screenshot shows the BaseSpace web interface for a project named 'NextSeq_RNA-Seq_Demo'. The interface includes a navigation bar with 'Dashboard', 'Prep', 'Runs', 'Projects', 'Apps', 'Public Data', and 'Help'. The 'Projects' tab is active. Below the navigation bar, there are several action buttons: 'Launch app', 'Download Project', 'Import', 'Share project', 'Get link', 'Edit project', and 'Transfer Owner'. The main content area displays a table of samples. The table has columns for 'Sample Id', 'Date Created', 'Read Length', 'Origin', 'Genome', and 'Size'. There are 8 samples listed, with the first three being 'HumanBrain2x125', 'HumanLiver2x125', and 'HumanHeart2x125'. A red oval is overlaid on the table, containing the text 'BaseSpace プロジェクトとしてデータをシェアしています'.

Sample Id	Date Created	Read Length	Origin	Genome	Size
1 HumanBrain2x125	Jun 18 2014				
2 HumanLiver2x125					
3 HumanHeart2x125					
4 HumanHeartDiseased2x125					
5 HumanLiver					
6 HumanHeartDiseased	Jun 07 2014	2 x 151	NextSeq_RNA-...		17.86 GB
7 HumanHeart	Jun 07 2014	2 x 151	NextSeq_RNA-...		25.13 GB
8 HumanBrain	Jun 07 2014	2 x 151	NextSeq_RNA-...		16.41 GB

本日の内容

- ▶ 今なぜRNA Seq?
 - マイクロアレイの原理と限界
 - RNA Seqの原理と長所
- ▶ NextSeq 500が実現する簡単、高速、安価なRNA Seq
 - NextSeq 500紹介
 - RNA Seqワークフロー
 - RNA Seqコスト
- ▶ 更に詳しく知りたい方へ

イルミナウェビナー

The screenshot shows the Illumina website interface. At the top is the Illumina logo and a navigation menu with items: アプリケーション, システム, インフォマティクス, 臨床研究, 受託サービス, サイエンス, サポート, and カンパニー. The main heading in orange reads '研究者による最新発表から システムの使い方まで'. Below this, text states 'ウェビナーを活用し オンラインで情報を収集'. A scientist in a white lab coat is shown holding a tablet. Below her are four content tiles: 1. 'イルミナセミナー2014 微生物ゲノム解析がもたらす 新たな知見' with a DNA helix icon. 2. 'デスクトップ型システム 新製品 NextSeq 500' with a computer monitor icon. 3. 'イルミナ ラボレーニング' with a pipette icon. 4. 'ウェビナーで最新情報を 研究に活用' with a smaller version of the scientist holding a tablet icon, which is circled in red.

illumina®

アプリケーション システム インフォマティクス 臨床研究 受託サービス サイエンス サポート カンパニー

研究者による最新発表から システムの使い方まで

ウェビナーを活用し
オンラインで情報を収集

イルミナセミナー2014
微生物ゲノム解析がもたらす
新たな知見

デスクトップ型システム
新製品 NextSeq 500

イルミナ
ラボレーニング

ウェビナーで最新情報を
研究に活用

参考文献

▶ マイクロアレイとの比較

Liu, S., Lin, L., Jiang, P., Wang, D. & Xing, Y. A comparison of RNA-Seq and high-density exon array for detecting differential gene expression between closely related species. *Nucleic Acids Res.* 39, 578–588 (2011).

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 18, 1509–1517 (2008).

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628 (2008).

▶ RNA Seq の再現性

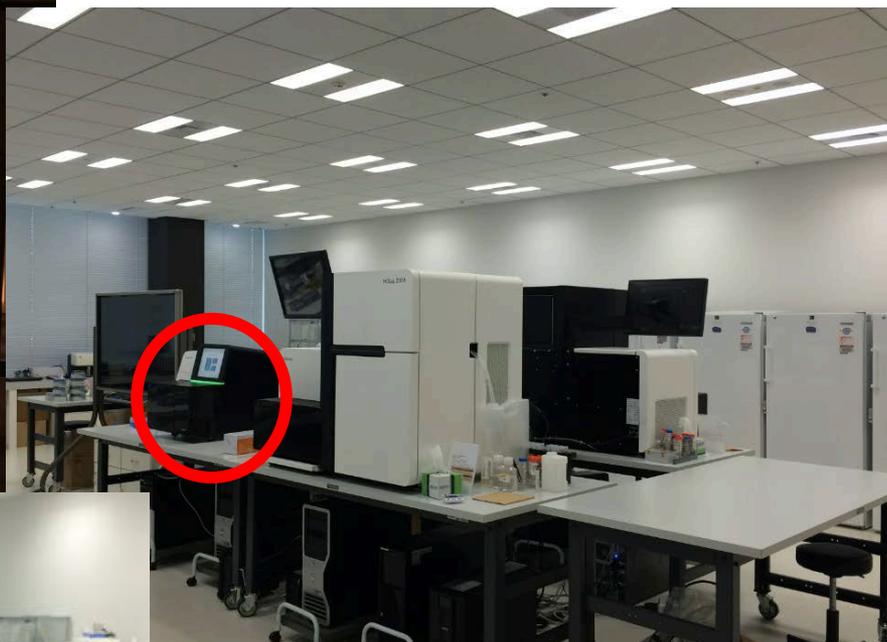
Shi, L. et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 24, 1151–1161 (2006).

SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat. Biotechnol.* doi:10.1038/nbt.2957 (24 August 2014).

't Hoen, P.A. et al. Reproducibility of high-throughput mRNA and small RNA sequencing across laboratories. *Nat. Biotechnol.* 31, 1015–1022 (2013).

NextSeq 500システム：実機はラボにて見学可能

RNA Seq トレーニングも実施中



ご清聴ありがとうございました