2011核酸研究新趨勢:
-非編碼核酸研究
-次世代定序儀的應用

威健股份有限公司
簡報人：李彥樑
jacklee64@welgene.com.tw

台北市南港軟體區, F棟12樓 電話: 02-66160001 ext. 223

威健生技 高通量技術專家

Office Area

Service Lab

Agilent Technologies
Authorized Distributor

Agilent Technologies
Certified Service Provider

EPIGENTEK

Biotech Building, Nan-Kang Business Park

DNAnexus

ISO 17025

Quality is Assured

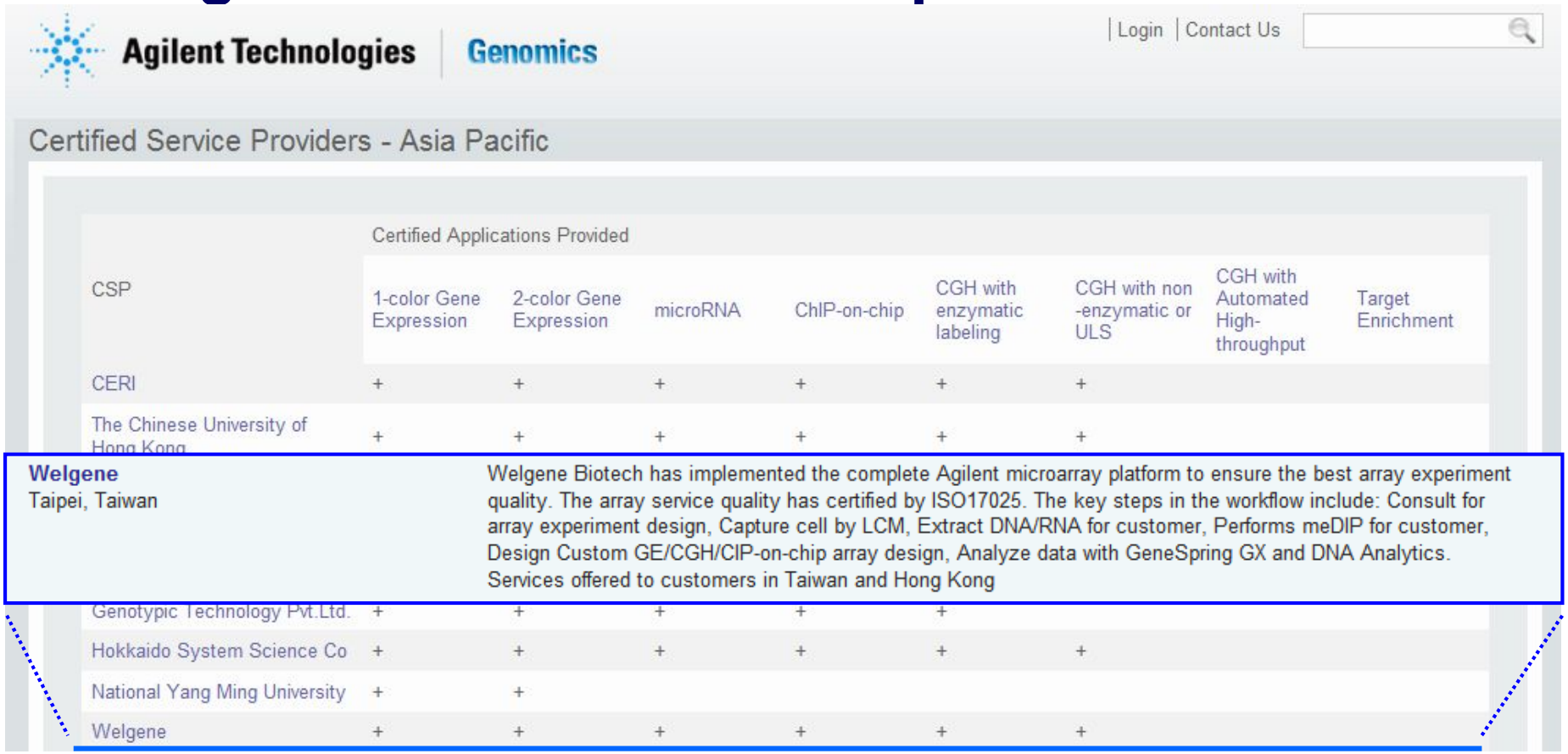
威健股份有限公司
Welgene Biotech. Co.,Ltd.

Welgene Biotech. Co. Ltd
http://www.welgene.com.tw

品質認證

Quality is Assured

- **ISO 17025**
- **Agilent certified service provider**



Agilent Technologies | Genomics

Login | Contact Us

Certified Service Providers - Asia Pacific

CSP	Certified Applications Provided							
	1-color Gene Expression	2-color Gene Expression	microRNA	ChIP-on-chip	CGH with enzymatic labeling	CGH with non -enzymatic or ULS	CGH with Automated High-throughput	Target Enrichment
CERI	+	+	+	+	+	+		
The Chinese University of Hong Kong	+	+	+	+	+	+		
Genotypic Technology Pvt.Ltd.	+	+	+	+	+			
Hokkaido System Science Co	+	+	+	+	+	+		
National Yang Ming University	+	+						
Welgene	+	+	+	+	+	+		

Welgene
Taipei, Taiwan

Welgene Biotech has implemented the complete Agilent microarray platform to ensure the best array experiment quality. The array service quality has certified by ISO17025. The key steps in the workflow include: Consult for array experiment design, Capture cell by LCM, Extract DNA/RNA for customer, Performs meDIP for customer, Design Custom GE/CGH/ChIP-on-chip array design, Analyze data with GeneSpring GX and DNA Analytics. Services offered to customers in Taiwan and Hong Kong

威健股份有限公司
Welgene Biotech. Co.,Ltd.

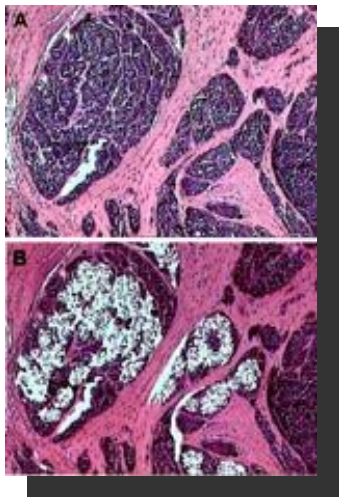
服務內容

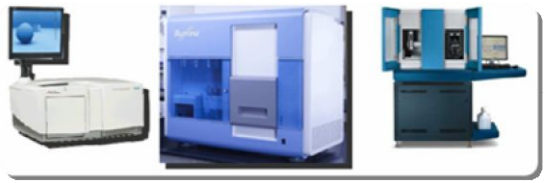
LCM
RNA isolation
Sample preparation

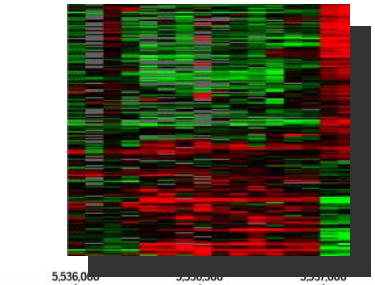
Array / NGS experiment

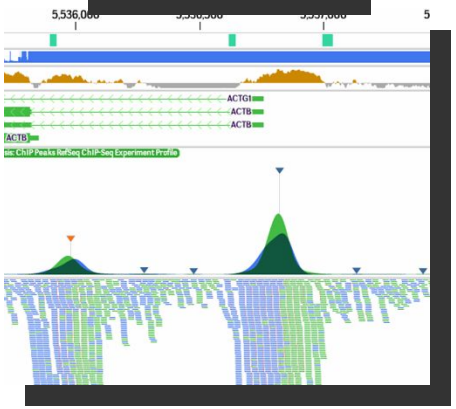
Data analysis (GeneSpring GX /DNAnexus)

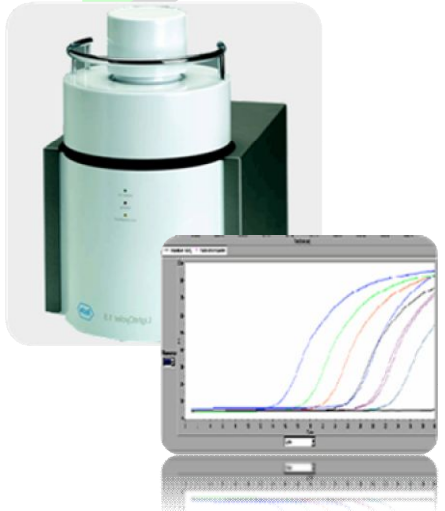
Validation (Q-RT-PCR)











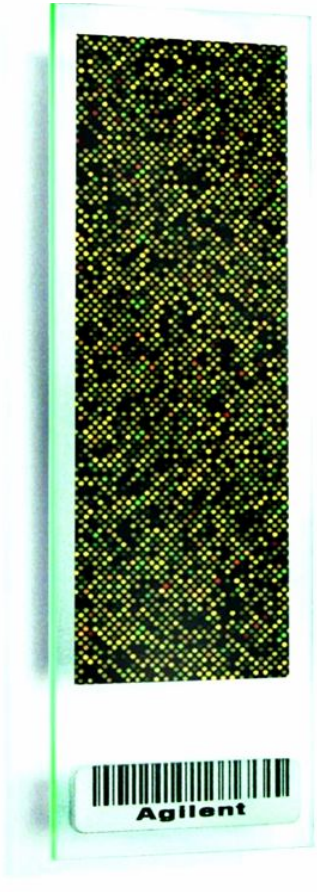
Welgene Biotech. Co. Ltd
http://www.welgene.com.tw

威健股份有限公司
Welgene Biotech. Co.,Ltd.

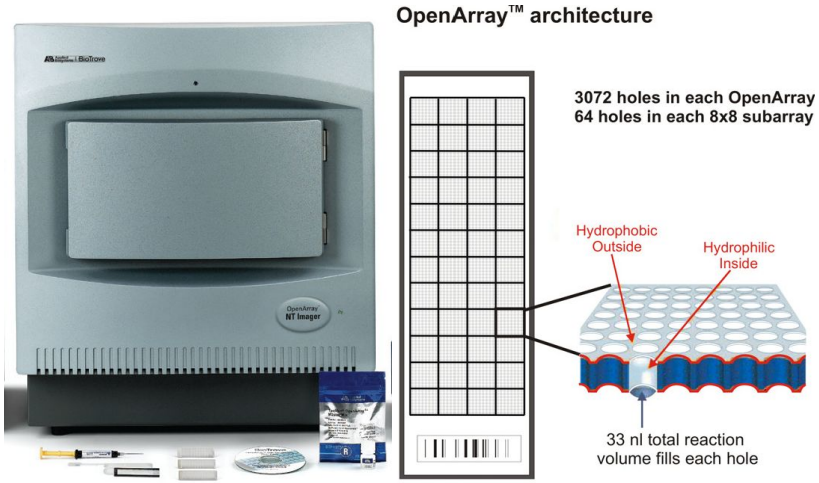
實驗平台

Microarray

Agilent Technologies



SNP QPCR Array



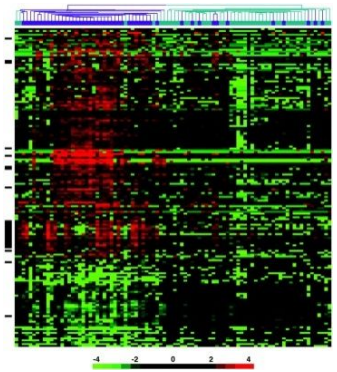
QPCR

NGS

Illumina Solexa (Experiment is performed by Fasteris or BGI)

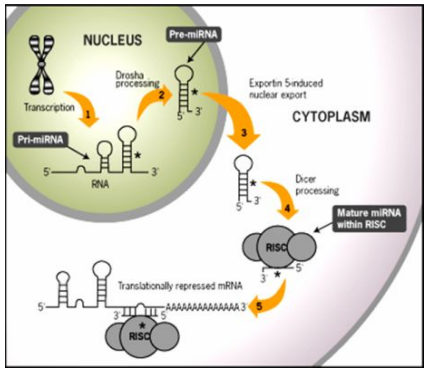
威健股份有限公司
Welgene Biotech. Co.,Ltd.

RNA Array



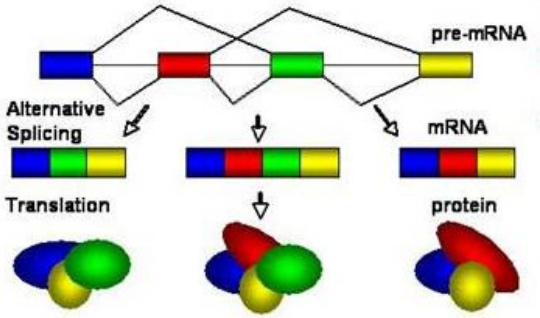
Gene Expression

Human Whole Genome Array 40,000 genes
Mouse Whole Genome Array 41,000 genes
Rat Whole Genome Array 41,000 genes



miRNA expression

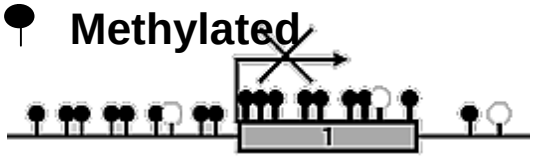
Human rel14 miRNA Array
Mouse rel14 miRNA Array
Rat rel14 miRNA Array



Alternative Splicing

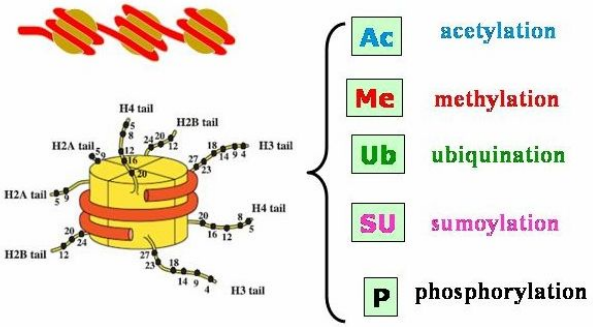
Species	Array Format	Genes Targeted	No. of Exon Probes	Probes from 8x60K Design	Databases Used for Design
Human	4x180K	20,411	174,458	19,128 (56%)	RefSeq Build 36.3, RSNM only
	2x400K	27,696	233,164	26,873 (78%)	RefSeq Build 36.3 Ensembl Release 52 Unigene Build 216 (Apr 2009) GenBank mRNA (Apr 2009)
Mouse	4x180K	23,215	165,984	19,203 (49%)	RefSeq Build 37, RSNM only
	2x400K	33,795	235,714	31,608 (80%)	RefSeq Build 37 Ensembl Release 55 Unigene Build 176 (Apr 2009) GenBank mRNA (Apr 2009) RIKEN 3
Rat	4x180K	20,483	160,141	23,444 (50%)	RefSeq Build 36.2, RSNM only
	2x400K	26,276	214,270	31,948 (72%)	RefSeq Build 36.2 Ensembl Release 55 Unigene Build 177 (Oct 2008) GenBank mRNA (Jan 2009)

DNA Array

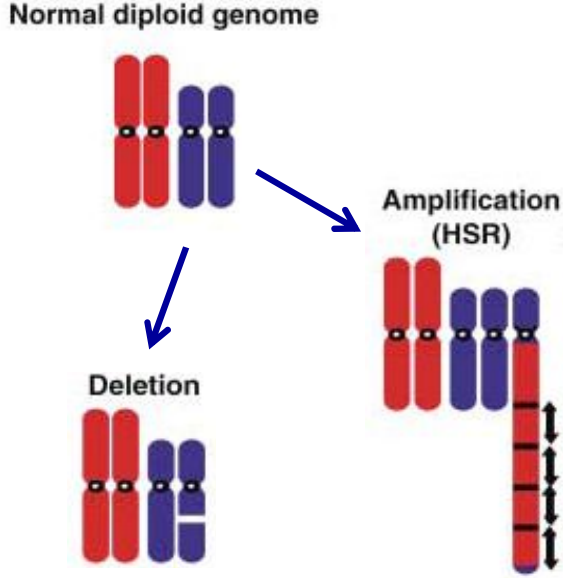


Gene inactive

DNA Methylation



Histone Modification



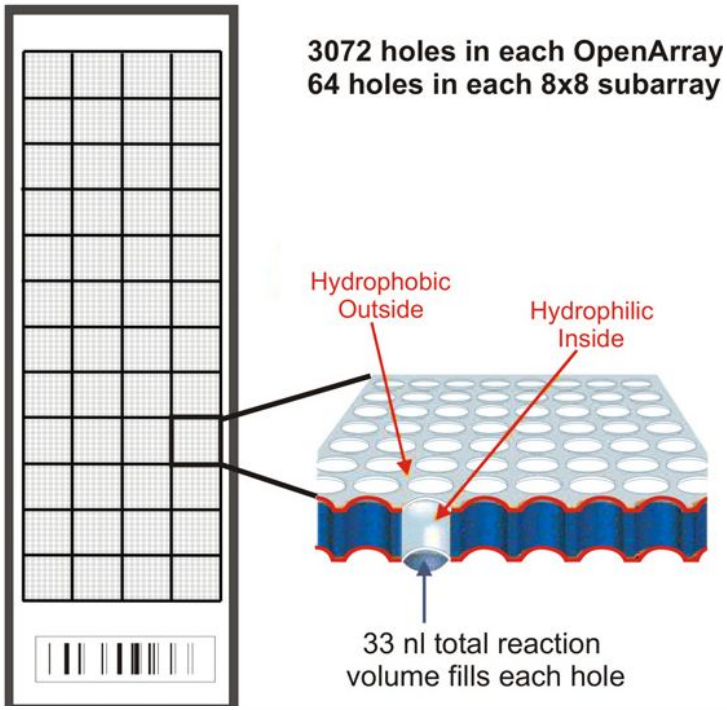
Copy Number Variation (CNV)

威健股份有限公司
Welgene Biotech. Co.,Ltd.

SNP Q-PCR Array System

基因分型儀器平臺（TaqMan® assay）

OpenArray™ architecture



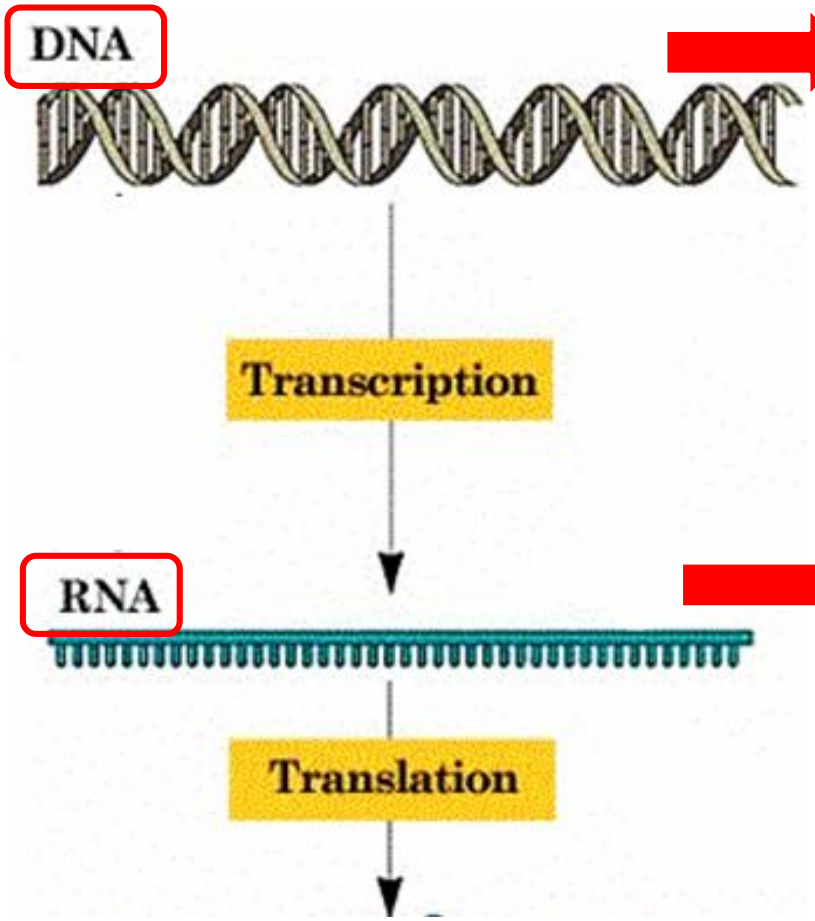
可分析不同 SNP 數	Openarray plate 訂購單位	每單位可分析總樣品數
16	10	1278
32	10	846
64	20	874

威健股份有限公司
Welgene Biotech. Co.,Ltd.

服務成果

- **Endocrinology, 2011** ;152(2) Insulin Up-Regulates Heme Oxygenase-1 Expression in 3T3-L1 Adipocytes via PI3-Kinase- and PKC-Dependent Pathways and Heme Oxygenase-1–Associated MicroRNA Downregulation.
- **Peptides. 2010** Jul;31(7):1262-72. Epub 2010 Apr 10. Antimicrobial peptide of an anti-lipopolysaccharide factor modulates of the inflammatory response in RAW264.7 cells.
- **Exp Cell Res. 2010** Apr 15;316(7):1119-26. Identification of the potential target genes of microRNA-146a induced by PMA treatment in human microvascular endothelial cells.
- **Vet Microbiol. 2010** Jan 25 A time-course study of gene responses of chicken granulosa cells to Salmonella Enteritidis infection.
- **J Neurotrauma. 2009** Dec;26(12):2345-53.Profiling muscle-specific microRNA expression after peripheral denervation and reinnervation in a rat model.
- **BMC Med Genomics. 2009** Aug 8;2:51 Hepatic inflammation mediated by hepatitis C virus core protein is ameliorated by blocking complement activation.
- **J Control Release. 2009** Nov 3;139(3):197-204. Epub 2009 Jul 18. Interactions between octaarginine and U-937 human macrophages: global gene expression profiling, superoxide anion content, and cytokine production.
- **BMC Bioinformatics 2009** ;17(10):85 Statistical identification of gene association by CID in application of constructing ER regulatory network.
- **Nucleic Acids Research, 2009** ;37(10):e77 Labeled microRNA pull-down assay system: an experimental approach for high-throughput identification of microRNA-target mRNAs.
- Breast Cancer: Basic and Clinical Research. 2008. Comparison and Identification of Estrogen-Receptor Related Gene Expression Profiles in Breast Cancer of Different Ethnic Origins.
- BMC Genomics. 2008 ;29(9):109 Altered expression patterns of lipid metabolism genes in an animal model of HCV core-related, nonobese, modest hepatic steatosis.
- J Hum Genet. 2008;53:136–43 Gene dosage change of TPTE and BAGE2 and breakpoint analysis in Robertsonian Down syndrome.
- Oncogene. 2007 ;26(57):7859-71. Molecular Signatures of Metaplastic Carcinoma of Breast by Large-Scale Transcriptional Profiling: Identification of Genes Potentially Related to Epithelial-Mesenchymal Transition.
- Oncol Rep. 2006 ;15(4):919-26. A homologue of the Drosophila headcase protein is a novel tumor marker for early-stage colorectal cancer.
- Fertility and Sterility 2006 ;86(6):1650-58, Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue.
- Ann Surg Oncol. 2006 ;13(11):1474-84. A gene expression profile for vascular invasion can predict the recurrence after resection of hepatocellular carcinoma: a microarray approach.
- Hepatogastroenterology. 2006 ;53(70):484-90. Distinct gene expression profiles in gastric epithelial cells induced by different clinical isolates of Helicobacter pylori--implication of bacteria and host interaction in gastric carcinogenesis.

各式晶片分析



晶片應用

SNP+aCGH (LOH Analysis)
aCGH (copy number variation)
MeDIP-chip (DNA Methylation)
ChIP-chip

Gene+LincRNA array (Expression)
Exon array (Splicing Analysis)
miRNA Expression

Non-coding RNA (ncRNA) Overview

A non-coding RNA (ncRNA) is a functional RNA molecule that is not translated into a protein. …

Non-coding RNA genes include:
1.Highly abundant and functionally important large RNAs (tRNA and rRNA)
2. Small RNAs (snoRNA, microRNA, siRNA and piRNA)
3. Long ncRNAs (such as Xist and HOTAIR)

Annotation of Non-Coding RNAs

tRNA
transfer RNA

Mt-tRNA
transfer RNA located in the mitochondrial genome

rRNA
ribosomal RNA

scRNA
small cytoplasmic RNA

snRNA
small nuclear RNA

snoRNA
small nucleolar RNA

miRNA
microRNA

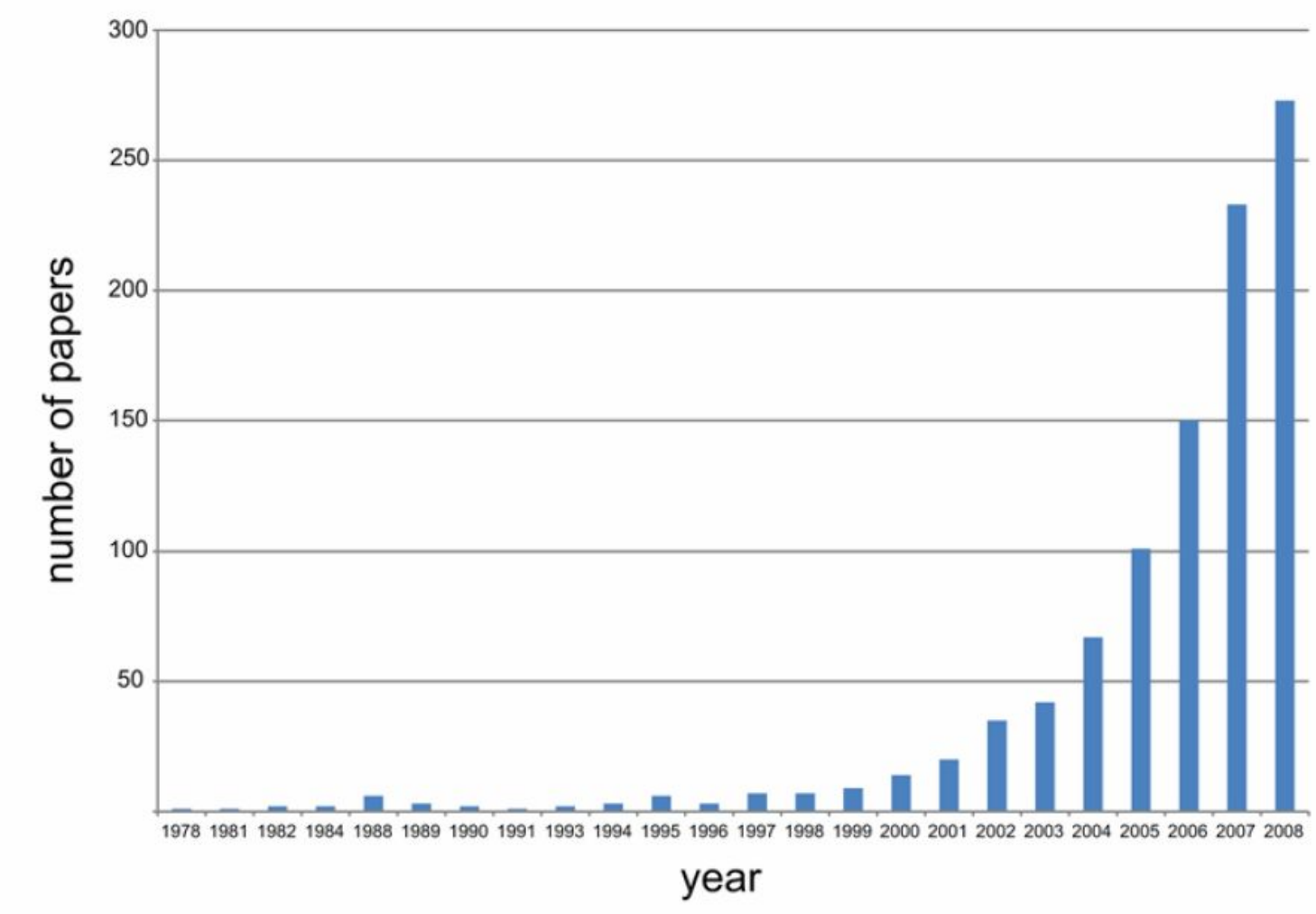
misc_RNA
miscellaneous other RNA

lincRNA
Long non-coding RNAs

http://www.ensembl.org/info/docs/genebuild/ncrna.html

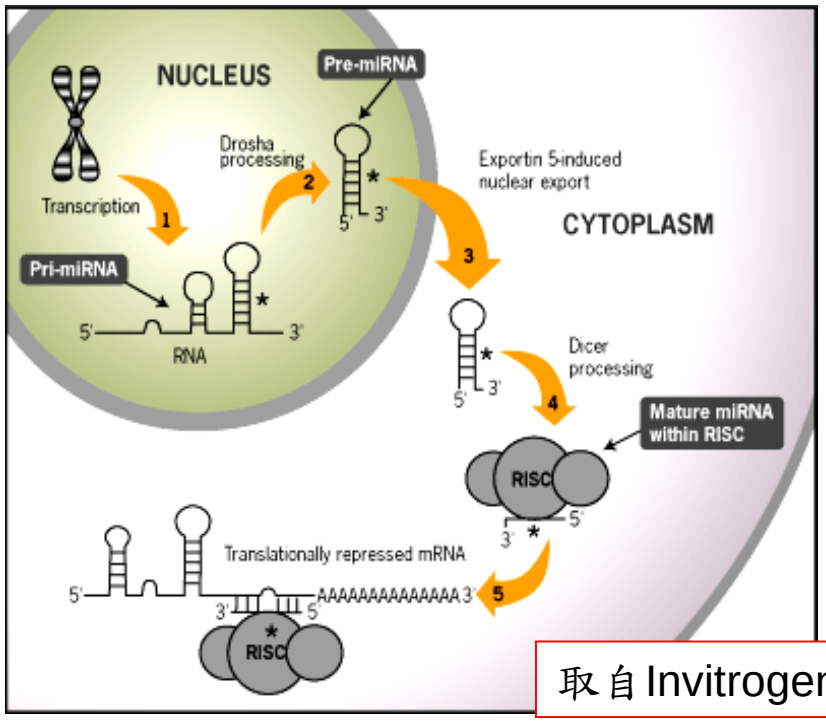
威健股份有限公司
Welgene Biotech. Co.,Ltd.

The recent rise in papers on ncRNAs



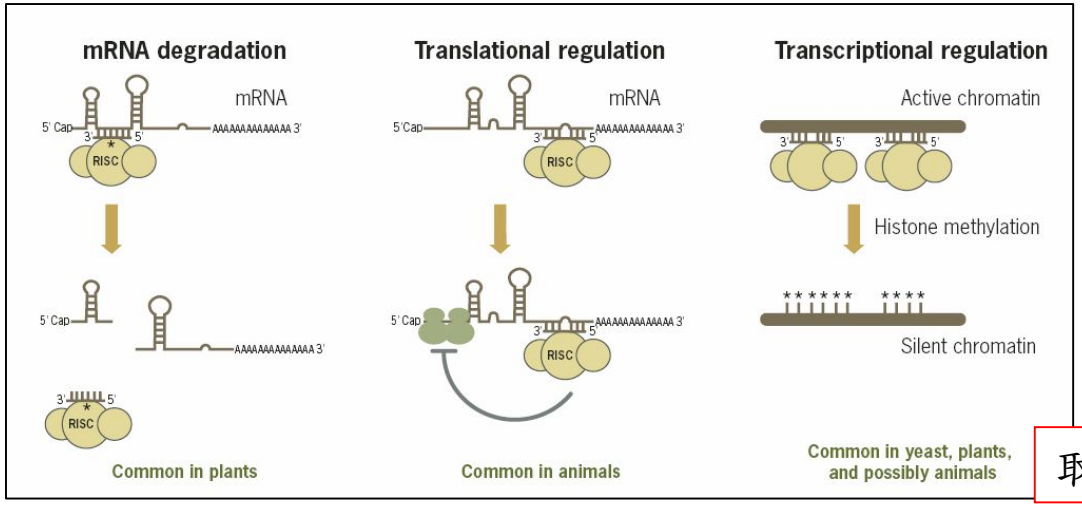
威健股份有限公司
Welgene Biotech. Co.,Ltd.

miRNA



取自 Invitrogen

- 20~24 nt short single strand RNA molecule
- Endogenous non-coding sequence
- Acting as translation inhibitor by degrading mRNA or blocking protein synthesis



取自 ambion

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Agilent miRNA array

V14/15

Sequences miRBase

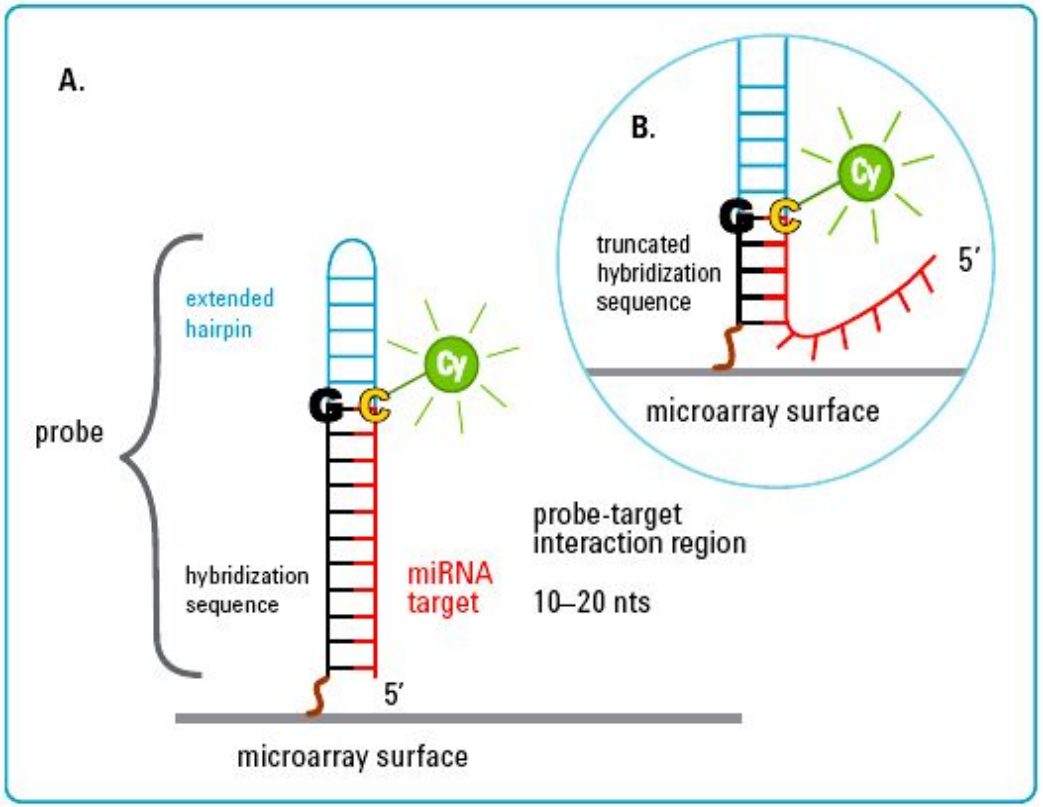


Figure 1. Components of the Agilent miRNA microarray probe design. An unmodified microarray probe (black) is a synthesized sequence that hybridizes to the target miRNA (red). Probes are anchored to the glass slide surface by a stilt (brown). **A.** Inclusion of a G residue (black) to the 5' end of the hybridization sequence complements the 3' end C residue (yellow) introduced in labeling. This additional G-C pair in the probe-target interaction region stabilizes targeted miRNAs relative to homologous RNAs. Additionally, all probes contain a 5' hairpin (blue), abutting the probe-target region, to increase target and size miRNA specificity. **B.** Destabilization of probes that are too stable. For probes requiring it, reduction of probe-target base-pairing is achieved through sequential elimination of base pairing from the 5' end of the miRNA.

- **Human rel14 miRNA Array**

  887 human miRNA

- **Mouse rel15 miRNA Array**

  690+ mouse-relative miRNA

- **Rat rel15 miRNA Array**

  388+ rat miRNA

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Journal of Oncology
Volume 2010, Article ID 950201, 7 pages
doi:10.1155/2010/950201

miRNA 與癌症

Review Article

MicroRNAs as Novel Biomarkers for Breast Cancer

H. M. Heneghan, N. Miller, A. J. Lowery, K. J. Sweeney, and M. J. Kerin

Department of Surgery, Clinical Science Institute, National University of Ireland, Galway, Ireland

Correspondence should be addressed to H. M. Heneghan, helenheneghan@hotmail.com

Received 15 March 2009; Accepted 8 May 2009

Academic Editor: Ben Davidson

ABLE 1: MiRNAs with altered expression in malignancy.

Tissue/tumor type	Increased expression	Decreased expression
Breast [26, 27]	miR-21, miR-29b-2	miR-125b, miR-145 miR-10b, miR-155, miR-17-5p, miR-27b
Ovarian [28, 29]	miR-141, miR-200(a-c), miR-221	let-7f, miR-140, miR-145, miR199a, miR-424
Endometrial [30–32]	miR-103, miR-107, miR-185, miR-205, miR-210, miR-449	miR-99b, miR-152, miR-193, miR-204, miR-221, let-7i
Glioblastoma [4, 25]	miR-221, miR-21	miR-181a, miR-181b, miR-181c
Chronic lymphocytic leukaemia [24]		miR-15, miR-16
Lymphoma [4, 11]	miR-155, miR-17-92cluster	miR-15a
Colorectal [4, 11, 25]	miR-10a, miR-17-92 cluster, miR-20a, miR-24-1, miR-29b-2, miR-31	miR-143, miR-145, let-7
Thyroid [4, 25]	miR-221, miR-222, miR-146, miR-181b, miR-197, miR-346	
Hepatocellular [4, 25]	miR-18, miR-224	miR-199a, miR-195, miR-200a, miR-125a
Testicular [11]	miR-372, miR-373	
Pancreatic [4, 11, 25]	miR-221, miR-376a, miR301, miR-21, miR-24-2, miR-100, miR-103-1,2, miR-107, miR-125b-1	miR-375
Cholangiocarcinoma [25]	miR-21, miR-141, miR-200b	
Prostate [11]	let-7d, miR-195, miR-203	miR-128a
Gastric [4, 11, 25]	miR-223, miR-21, miR-103-2	miR-218-2
Lung [4, 11, 25]	mir-17-92 cluster, miR-17-5p	let-7 family

Cell Death and Differentiation (2010) 17, 209–214

www.nature.com/cdd

Review

miR-10 in development and cancer

AH Lund*,1

The microRNA (miRNA) miR-10 family has attracted attention because of its conservation and the position of the miR-10 genes within the *Hox* clusters of developmental regulators. In several species, miR-10 is coexpressed with a set of *Hox* genes and has been found to regulate the translation of *Hox* transcripts. In addition, members of the miR-10 family are de-regulated in several cancer forms. Aside from acting in translational repression, miR-10 was recently found to bind a group of transcripts containing a terminal oligo-pyrimidine (TOP) motif and to induce their translation, thereby adding a new function to the miRNA repertoire.
Cell Death and Differentiation (2010) **17**, 209–214; doi:10.1038/cdd.2009.58; published online 22 May 2009

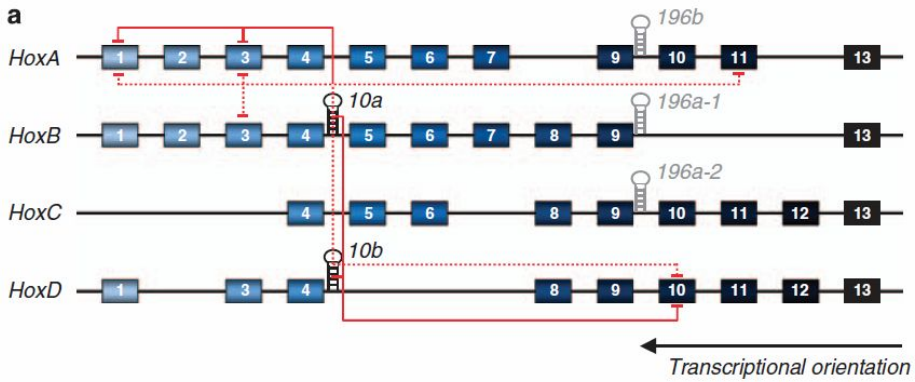


Table 1 Deregulation of miR-10 family members in cancer

Cancer type	miR-10 family member	Regulation	References
Glioblastoma	miR-10a, miR-10b	Upregulation	Gaur *et al.*[26], Silber *et al.*[33], Ciafre *et al.*[34]
Hepatocellular carcinomas	miR-10a, miR-100	Upregulation	Varnholt *et al.*[35]
Pancreatic cancer	miR-10b, miR-99 and miR-100	Upregulation	Bloomston *et al.*[36]
Colon cancer	miR-10a	Upregulation	Volinia *et al.*[37]
Breast cancer (metastasis)	miR-10b	Upregulation	Ma *et al.*[1]
Acute myeloid leukaemia with *NPM1* mutations	miR-10a, miR-10b, miR-100	Upregulation	Garzon *et al.*[29]
B-cell chronic lymphocytic leukaemia	miR-10b	Upregulation	Calin *et al.*[28]
Melanoma	miR-10a	Gene amplification	Zhang *et al.*[23]
Breast cancer	miR-10a	Gene amplification	Zhang *et al.*[23]
Chronic myeloid leukaemia	miR-10a	Downregulation	Agirre *et al.*[22]
Acute myeloid leukaemia	miR-10a	Downregulation	Jongen-Lavrencic *et al.*[27]

Am J Transl Res 2010;2(2):170-180
www.ajtr.org /AJTR1003003

Review Article

miR-145-mediated suppression of cell growth, invasion and metastasis

Mohit Sachdeva, Yin-Yuan Mo

Department of Medical Microbiology, Immunology and Cell Biology, Southern Illinois University School of Medicine, Springfield, IL, USA.

Received March 17, 2010; accepted March 22, 2010, available online March 25, 2010

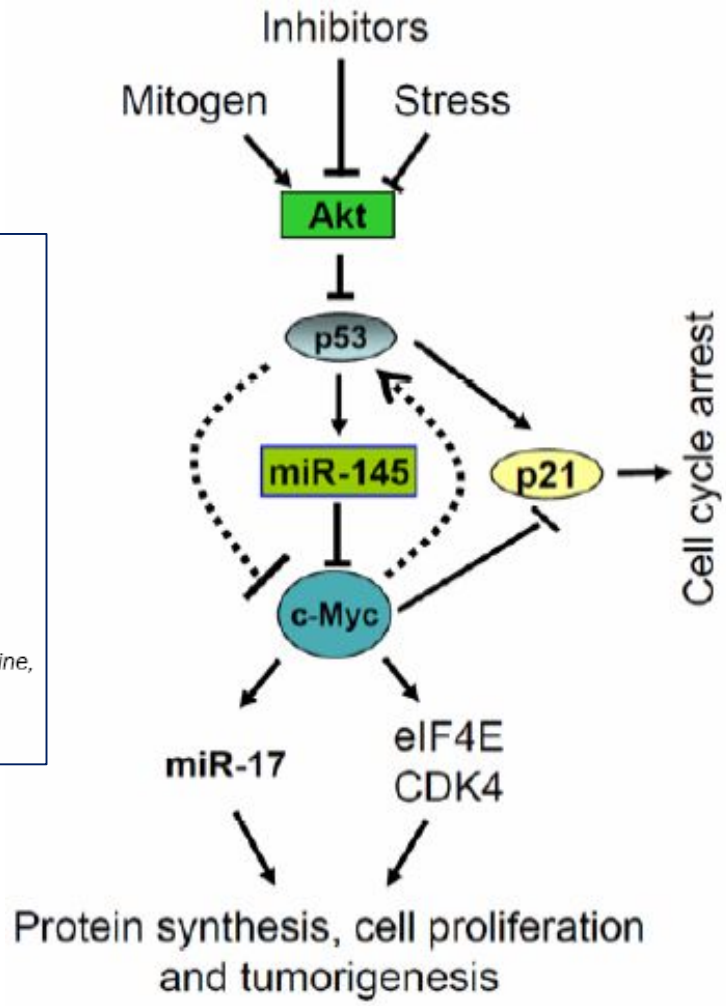
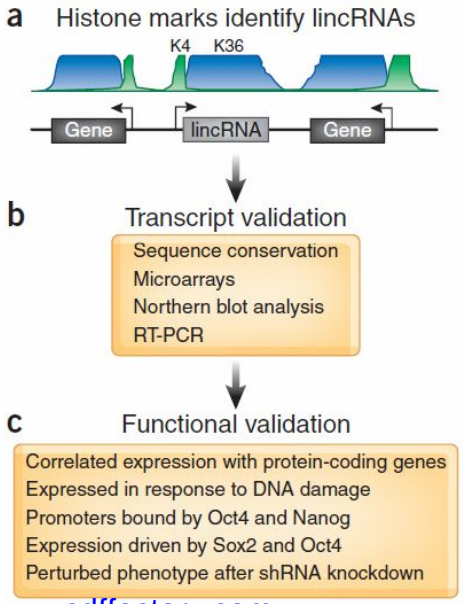


Figure 1. A simplified working model for miR-145 and its function as a tumor suppressor.

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Large Interngenic Non-coding RNA (LincRNA)

- > 200 nt, was considered as “transcriptional noise”
- Many large ones have polyA
- Also called as “Long intervening noncoding RNA” or “Large non-coding RNA (LncRNA)”
- Long noncoding RNAs with little or no protein-coding capacity
- Evolutionarily conserved in mammalian genomes and thus presumably function in diverse biological processes
- Recent discovery of more than a thousand genes known as large intergenic non-coding RNAs



NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 4 APRIL 2009

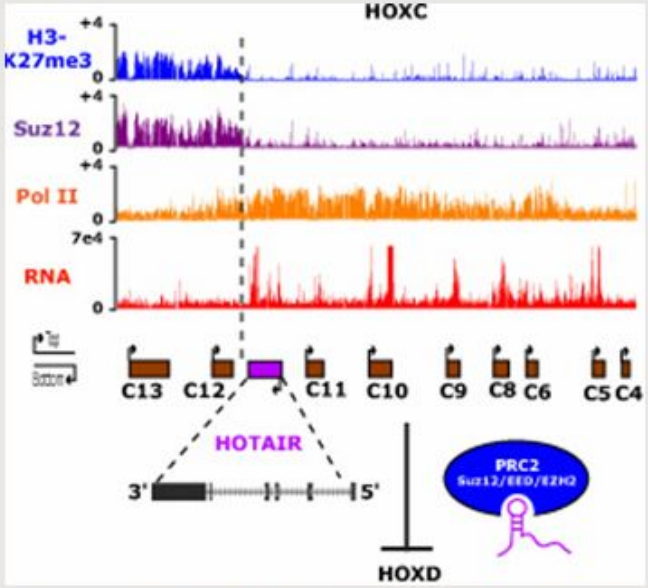
Figure 1 Identification and validation of long intervening noncoding RNAs. (**a**) Patterns of H3K4me3 (trimethylated lysine 4 of histone H3, green) adjacent to H3K36me3 (trimethylated lysine 36 of histone H3, blue) mark transcription units and identify unannotated transcripts positioned between protein-coding genes. (**b**) Large multi-exonic non-protein-coding RNA transcripts are detected by a variety of methods. (**c**) Experimental and computational validation suggests that lincRNAs are biologically functional. shRNA, short hairpin RNA.

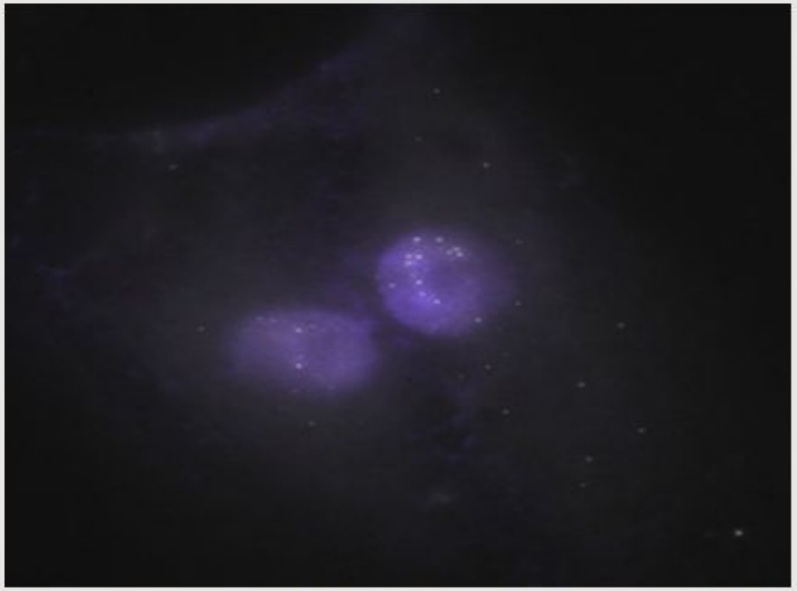
Dr. John Rinn

Assistant professor of pathology at Harvard Medical School and Beth Israel Deaconess Medical Center and an associate member of the Broad Institute.

威健股份有限公司
Welgene Biotech. Co.,Ltd.

LincRNAs and Chromatin Structure





Recently a lincRNA termed HOTAIR (Rinn et al, 07) was discovered that is encoded antisense to the HOXC cluster at the exact juncture of a 40 Kb domain of heterochromatin and a 60 Kb domain of euchromatin. However, HOTAIR doesn? serve to regulate this boundary; Remarkably HOTAIR affects the global epigenetic state of the HOXD cluster located on a separate chromosome. HOTAIR binds the Polycomb Repressive Complex 2 (PRC2) and is required for PRC2 occupancy and histone H3 lysine-27 trimethylation of HOXD locus. Thus, transcription of large ncRNA may associate with chromatin remodeling complexes and guide them to their sites of action. We are employing a combined informatic, biochemical and immunofluorescent approaches to identify other lincRNAs associated with chromatin remodeling complexes. Together our studies aim to understand the roles of lincRNAs in portioning the genome into proper domains of euchromatin and heterochromatin.

HOTAIR Became Real

hot air

即時發音 看清

【俚】空話；大話

Cell. 2007 June 29; 129(7): 1311–1323.

Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs

John L. Rinn[1], **Michael Kertesz**[2,5], **Jordon K. Wang**[1,5], **Sharon L. Squazzo**[4], **Xiao Xu**[1], **Samantha A. Brugmann**[3], **Henry Goodnough**[3], **Jill A. Helms**[3], **Peggy J. Farnham**[4], **Eran Segal**[2], and **Howard Y. Chang**[1]

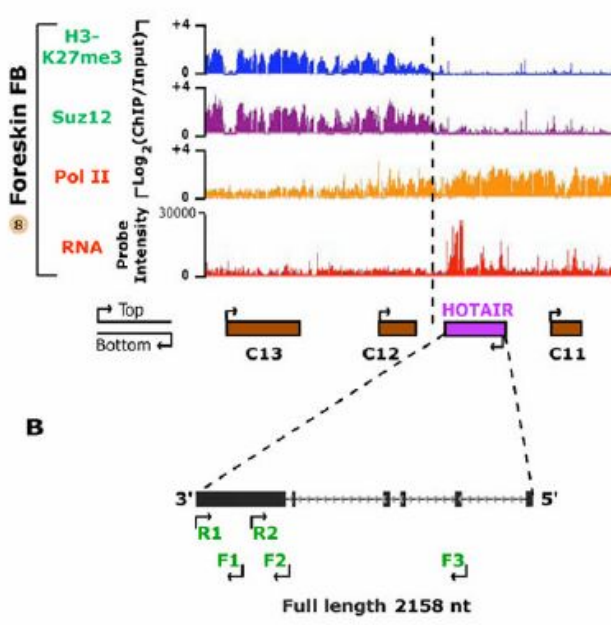
1 *Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA*

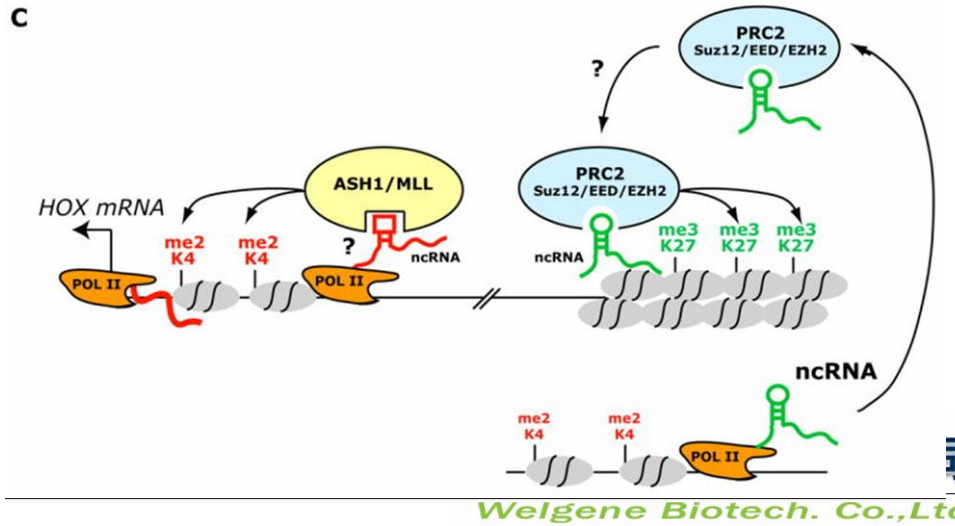
3 *Department of Surgery, Stanford University School of Medicine, Stanford, CA 94305, USA*

2 *Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel*

4 *Department of Pharmacology and Genome Center, University of California-Davis, 95616*

We identified a 2.2 kilobase ncRNA residing in the HOXC locus, termed HOTAIR, which represses transcription *in trans across 40 kilobases of* the HOXD locus. HOTAIR interacts with Polycomb Repressive Complex 2 (PRC2) and is required for PRC2 occupancy and histone H3 lysine-27 trimethylation of HOXD locus. Thus, transcription of ncRNA may demarcate chromosomal domains of gene silencing at a distance; these results have broad implications for gene regulation in development and disease states.





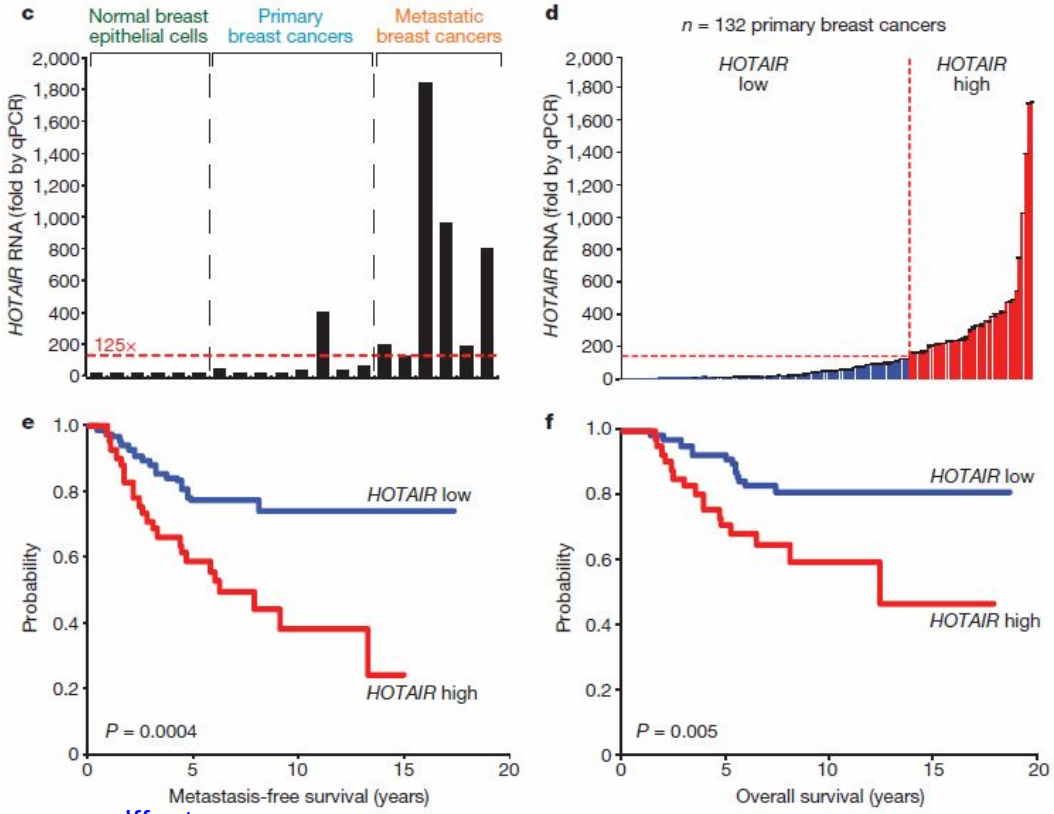
Welgene Biotech. Co.,Ltd.

Vol 464|15 April 2010|doi:10.1038/nature08975 nature

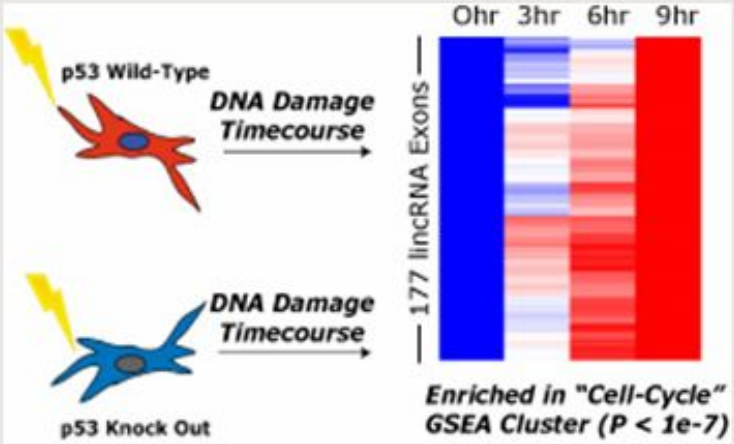
LETTERS

Long non-coding RNA *HOTAIR* reprograms chromatin state to promote cancer metastasis

Rajnish A. Gupta[1], Nilay Shah[4], Kevin C. Wang[1], Jeewon Kim[2], Hugo M. Horlings[6], David J. Wong[1], Miao-Chih Tsai[1], Tiffany Hung[1], Pedram Argani[5], John L. Rinn[7], Yulei Wang[8], Pius Brzoska[8], Benjamin Kong[8], Rui Li[3], Robert B. West[3], Marc J. van de Vijver[6], Saraswati Sukumar[4] & Howard Y. Chang[1]



LincRNAs in Cancer



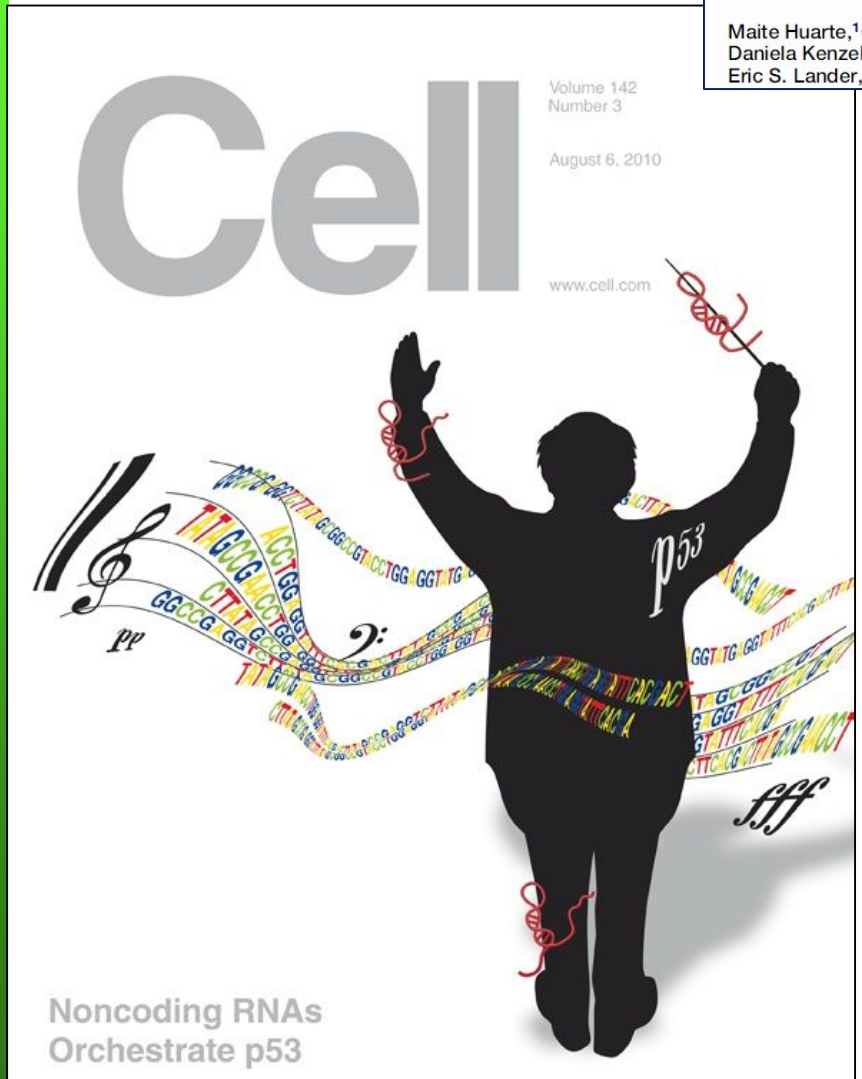
The largest signature our guilt by association method revealed was comprised of hundreds of lincRNAs that all share functional associations with cell-cycle regulation, proliferation, RNA binding proteins and chromatin remodeling complexes. These results strongly suggested misregulation of lincRNAs would result in tumor formation by loss of epigenetic regulation of cell-cycle regulation genes, oncogenes and/or tumor suppressors. We have further identified candidate ? nco-lincRNAs??and ? umor-suppressor lincRNAs??based on associations with key cancer pathways as well as known oncogenic and tumor-suppressor pathways, respectively. In addition, we will look for misregulation of lincRNAs that are strongly associated with cell-cycle regulation, proliferation, RNA binding proteins and chromatin remodeling complexes. By profiling lincRNAs across numerous cancer types we have rapidly identify lincRNAs that could play key roles tumor initiation and/or progression that will be explored for their functional mechanism. In summary, lincRNAs could herald a new paradigm in our understanding of cellular transformation and/or metastasis. Defining the roles of these RNA molecules in cancer could open up new avenues for better diagnostics.

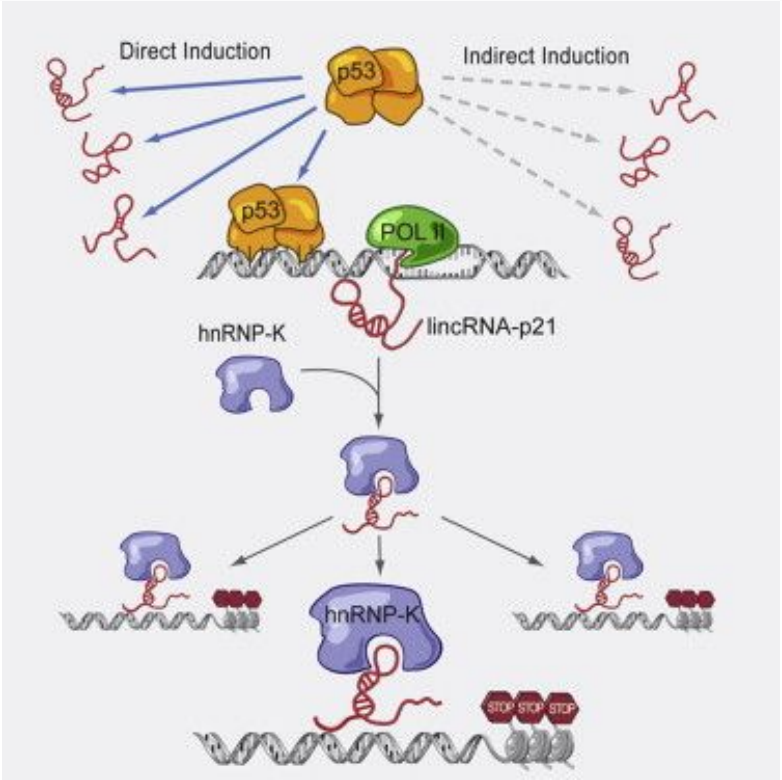
WELGENE

Cell

A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response

Maite Huarte,[1,2,*] Mitchell Guttman,[1,3] David Feldser,[3,4] Manuel Garber,[1] Magdalena J. Koziol,[1,2] Daniela Kenzelmann-Broz,[5,6] Ahmad M. Khalil,[1,2] Or Zuk,[1] Ido Amit,[1] Michal Rabani,[1] Laura D. Attardi,[5,6] Aviv Regev,[1,3] Eric S. Lander,[1,3,7] Tyler Jacks,[3,4] and John L. Rinn[1,2,*]





LincRNA-p21 mediates global gene repression and apoptosis in the p53 pathway

Human Molecular Genetics, 2010 R1–R10
doi:10.1093/hmg/ddq353

Large non-coding RNAs: missing links in cancer?

Maite Huarte[1,2] and John L. Rinn[1,2,*]

[1]The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA and [2]Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, 02215, USA

Received July 17, 2010; Revised August 6, 2010; Accepted August 16, 2010

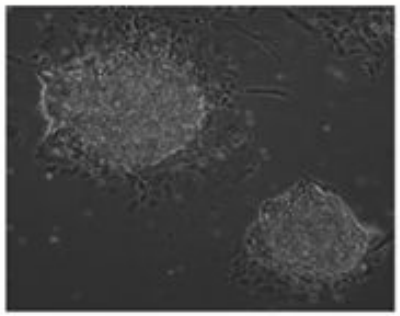
Table 1. Examples of large ncRNAs potentially oncogenic

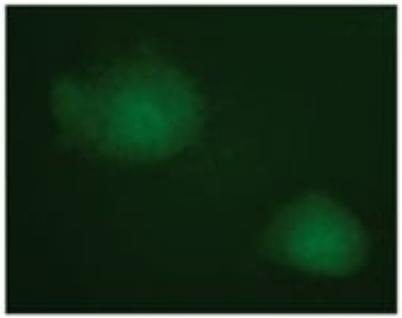
RNA	Organism	Size	Genomic location	Expression	Functional characteristics	Mechanism	References
p21 NAT/ Bx332409	*Homo sapiens*	≥423 bp	Antisense *cdkn1a/ p21*	EST sequenced from neuroblastoma	Antisense of *Cdnk1a/p21*. Negative regulation of *Cdnk1a/p21* by epigenetic. silencing	Requirement of Ago1 but not Ago2 for epigenetic silencing of *Cdkn1a/p21*	(44)
HOTAIR	*H. sapiens*	2.2 kb	Intergenic *HoxC* locus	Distal fibroblasts/metastatic breast tumors	Gene silencing in *trans*. Metastasis	Interaction with PRC2 and LSD1 complexes and targeting to repressed genes	(22,34)
MALAT-1	*H. sapiens*	8.7 kb	Intergenic Chr11	Overexpression in lung adenocarcinoma, breast, pancreas, colon, prostate and hepatocellular carcinomas	Metastasis	Induction of *GAGE6* proto-oncogene transcription by inhibition of PSF repressor	(35,49–51)
VL30-1	*Mus musculus*	4.9 kb	Retroelement ncRNA		Ras-mediated transformation of mouse fibroblasts	Induction of *Rab23* proto-oncogene transcription by inhibition of PSF repressor	(47,60–62)
ANRIL	*H. sapiens*	2.2 kb	Antisense of *INK4n/ARF/ INK4a* and *p15/ CDKN2B*	Upregulated in prostate cancer	Gene silencing of *INK4a/ARF/ INK4a* and *p15/CDKN2B*	Interaction with CBX7 component of PRC1 complex	(42,43,63–65)
H19	*H. sapiens*	2.3 kb	Imprinted *H19-Igf2* locus in Chr11	Bladder, human lung and breast carcinoma, choriocarcinoma	Control of imprinting. Oncogenic or tumor suppressor. Containing microRNA miR-675	Unknown	(66–70)
CUDR	*H. sapiens*	2.2 kb	Intergenic Chr19	Overexpressed in a doxorubicin-resistant human squamous carcinoma subline	May regulate drug sensitivity and promote cellular transformation through resistance to apoptosis	Unknown	(71)
Zeb2/Sip1 NAT	*H. sapiens*	680 bp	Antisense of *Zeb2/ Sip1*	Overexpressed in human tumors with low E-cadherin expression	Inhibition of *E-cadherin* through induction of Zeb2 protein levels	Inhibition of splicing of *Zeb2* first exon containing IRES sequence	(46)
SRA-1	*H. sapiens*	875 bp	Alternative splicing of *SRA* gene, loss of coding frame	In breast cancer cells, different balance of coding/non-coding splicing isoforms	Co-activator of steroid receptors and other transcription factors as MyoD. Increased levels of non-coding isoform associated with metastasis	Interaction in ribonucleoprotein complexes with several positive regulators, including SRC-1, p68 and p72, Pus1p and Pus3p, as well as negative regulators, such Sharp and SLIRP to be recruited to promoters of regulated genes	(72–79)
PCGEM1	*H. sapiens*	1.6 kb	Intergenic, Chr2	Upregulated in prostate cancer in African-American patients	Inhibition of apoptosis, promotion of cell growth	Unknown	(80–84)
UCA1	*H. sapiens*	1.4 kb	Intergenic, Chr19	Embryonic development and bladder cancer associated	Increases proliferation, migration, invasion and drug resistance of human bladder cell line	Unknown	(85)

Table 2. Examples of potential tumor-suppressor large ncRNAs

RNA	Organism	Size	Genomic location	Expression	Functional characteristics	Mechanism	References
GAS5	*H. sapiens*, *M. musculus*	Multiple splicing isoforms 0.6–1.8 kb (human)	Intergenic. Some introns encode snoRNAs	Downregulated in breast cancer, upregualted by induced growth arrest	Induces growth arrest and apoptosis	Blocks transcriptional induction by GR by binding to GR DNA binding domain and competing with DNA GRE	(52–54,86–91)
lincRNA-p21	*M. musculus*	3.1 kb	Intergenic, upstream of *p21/cdkn1a*	Induced by p53 upon DNA damage or oncogenic stress	Global gene repression in the p53 transcriptional response inducing cellular apoptosis	Physical interaction with hnRNP-K, targeting genes for transcriptional repression	(36)
MEG3	*H. sapiens*	Several isoforms 1.6–1.8 kb	Intergenic, Chr 14	Maternally expressed imprinted gene. Hypermethylated in myeloid leukemia, multiple myeloma and pituitary tumors	Activates p53 activity on specific genes by increasing p53 protein levels by suppressing MDM2 levels. Also inhibits cell proliferation independently of p53	Unknown	(92–98)
ncRNA CCND1/ cyclin D1	*H. sapiens*	≥200–300 nt	Transcribed from 5' end of *cyclin D1* gene	Induced by DNA damage	Induces repression of *cyclin D1* gene	Binding to TLS protein induces TLS allosteric change, allowing interaction with *cyclin D1* gene, inhibiting CBP/p300 HAT activity and resulting in *cyclin D1* repression	(56)

LincRNAs in ES Pluripotency





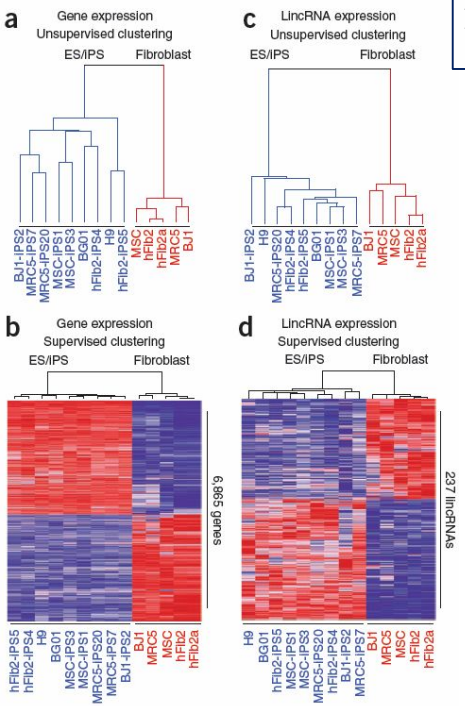
Our ? uilt by association??informatics approach revealed a striking class of lincRNAs that appear to be involved in Embryonic Stem (ES) cell pluripotency. These lincRNAs are not only exclusively expressed in ES cells but are also directly transcribed by key transcription factors (Sox2/Oct4/Nanog) that are known to regulate ES pluripotency. However, the ultimate proof that these lincRNAs are required for ES pluripotency would be to perform loss-of-function by depleting lincRNA transcripts and monitoring the resulting phenotypic changes (i.e. loss of self-renewal and/or pluripotency). To this end, we along with the Broad RNAi Platform are performing high-throughput loss of function experiments using shRNAs targeting ES lincRNAs, and monitoring the resulting perturbations to ES pluripotency.

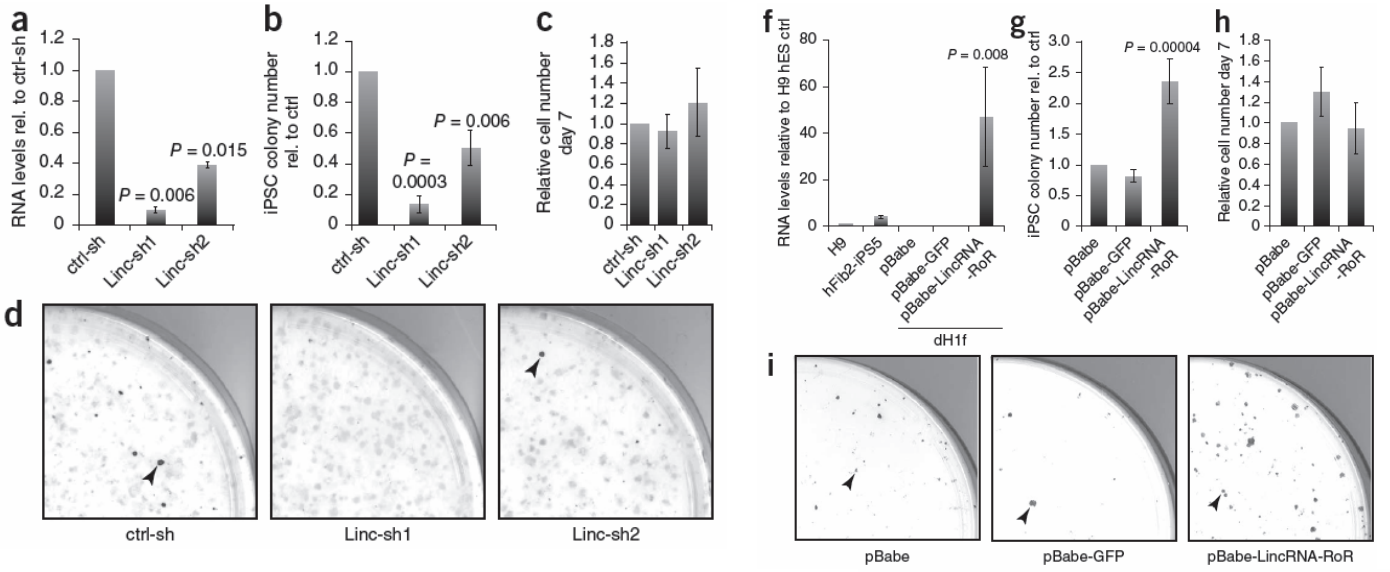
LETTERS

nature genetics

Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells

Sabine Loewer[1–4], Moran N Cabili[5,6], Mitchell Guttman[5,7], Yuin-Han Loh[1–4], Kelly Thomas[5,8], In Hyun Park[1–4,12], Manuel Garber[5], Matthew Curran[1,3], Tamer Onder[1–4], Suneet Agarwal[1–3], Philip D Manos[1,3,4], Sumon Datta[1,3,4], Eric S Lander[5–7], Thorsten M Schlaeger[1,3,4], George Q Daley[1–4,8,9] & John L Rinn[5,10,11]





威健股份有限公司
Welgene Biotech. Co.,Ltd.

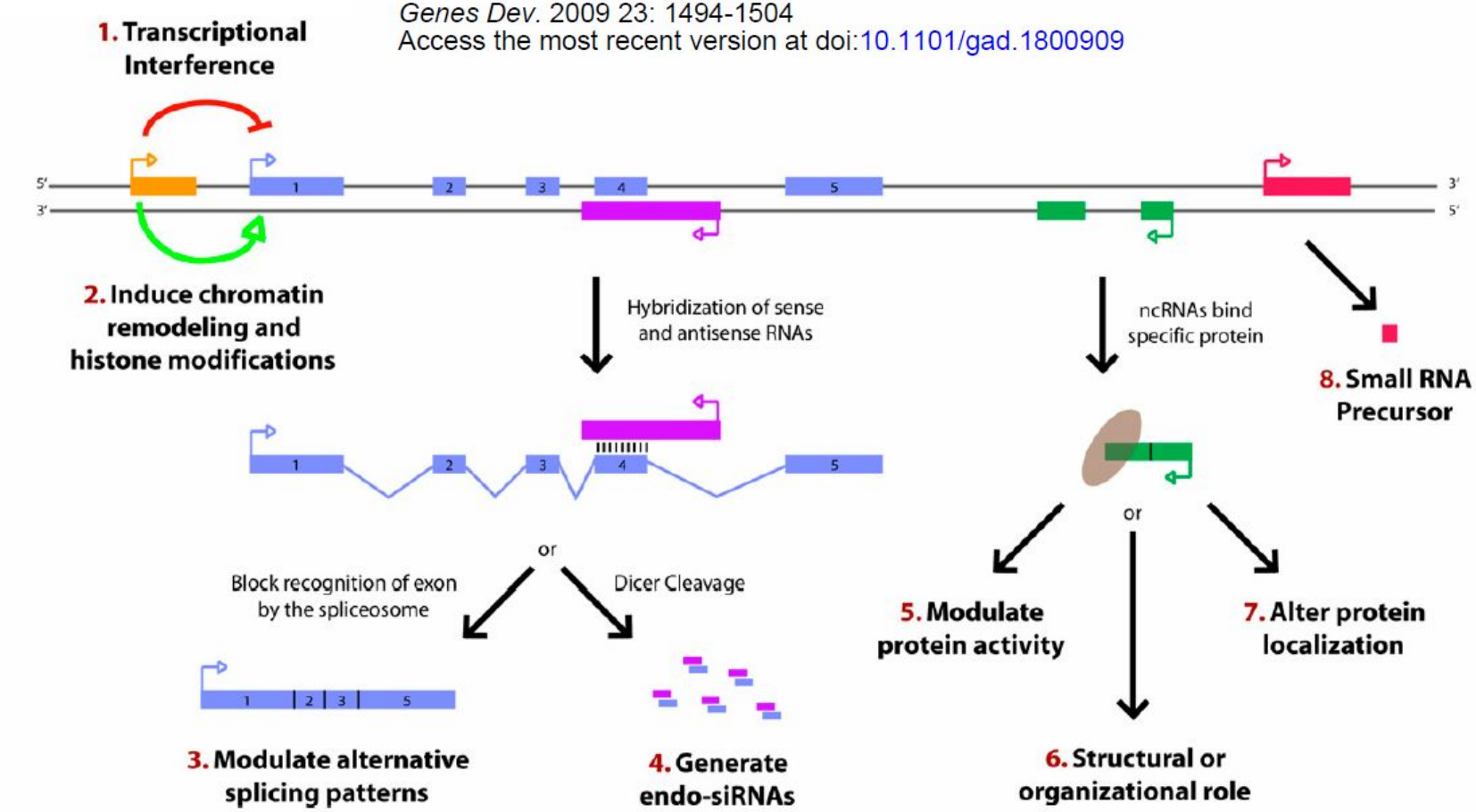
WELGENE

CSH PRESS Genes & Development

Long noncoding RNAs: functional surprises from the RNA world

Jeremy E. Wilusz, Hongjae Sunwoo and David L. Spector

Genes Dev. 2009 23: 1494-1504
Access the most recent version at doi:10.1101/gad.1800909



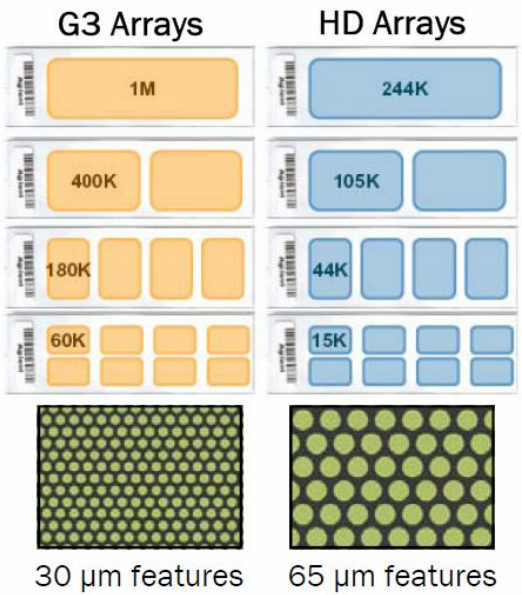
威健股份有限公司
Welgene Biotech. Co.,Ltd.

Agilent Technologies Gene + LincRNA Array

NEW

SurePrint G3 8x60K Microarrays

Species	Array Format	Entrez Gene Targets	% New Probes	lincRNA Targets	Database used for design
Human	**SurePrint G3 8x60K**	**27,958**	**45.6%**	**7,419**	*Refseq Build 36.3; Ensembl Release 52; Unigene Build 216 (Apr 2009); GenBank mRNA (Apr 2009)*
	SurePrint HD 4x44K v2	27,958	—-	Ø	
Mouse	**SurePrint G3 8x60K**	**34,017**	**63.4%**	**4,623**	*Refseq Build 37; Ensembl Release 55; Unigene Build 176 (Apr 2009); GenBank mRNA (Apr 2009)*
	SurePrint HD 4x44K v2	34,017	—-	Ø	
Rat	**SurePrint G3 8x60K**	**26,930**	**61.8%**	**Ø**	*Refseq Build 36.2; Ensembl Release 55; Unigene Build 177 (Oct 2008); GenBank mRNA (Jan 2009)*
	SurePrint HD 4x44K v2	26,930	—-	Ø	



威健股份有限公司
Welgene Biotech. Co.,Ltd.

Next Generation Sequencing

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Next Generation Sequencing (NGS)

Sanger Sequencing 一次讀 <100 條序列, 每條 ~1,000bp

- 750-1000 bp/read
- 96 reads/run
- 2-200 kb paired-end
- ~1 Mb/day

NGS (massively parallel sequencing) 一次讀 10^8條序列, 每條 ~100bp

- 50-100 bp/read (>500bp/read)
- 100,000,000 reads/run (1,000,000 reads)
- 10-25 Gb/day

Illumina Solexa

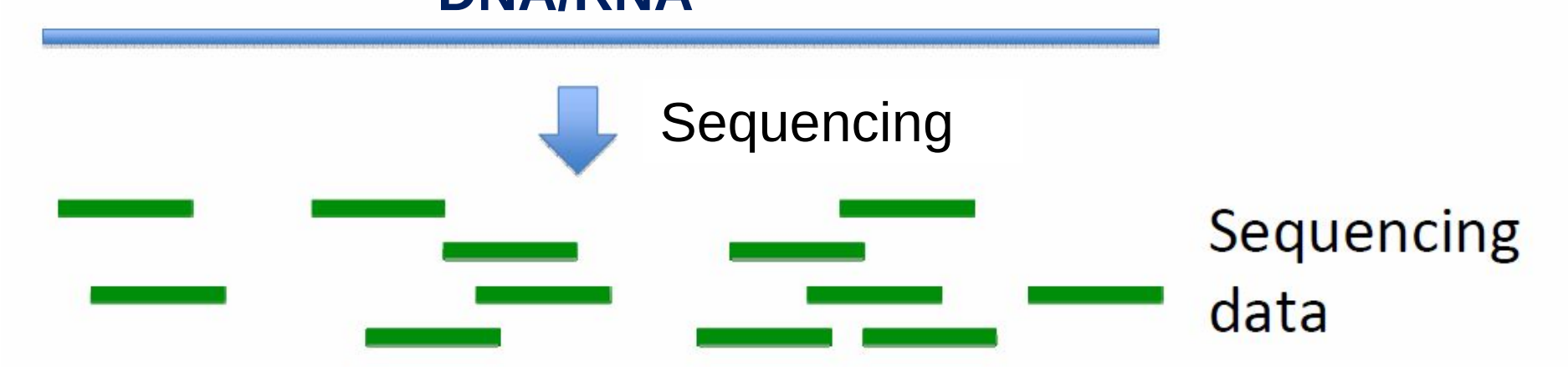
AB SOLiD

Roche 454

	Roche (454): Titanium series reagents (run on FLX Genome Sequencner)	Illumina: Genome Anlayser II	Applied Biosystems: SOLiD™
Sequencing chemistry	pyrosequencing	polymerase-based sequencing-by-synthesis	Ligation-based sequencing
Amplification approach	Emulsion PCR	Bridge amplification	Emulsion PCR

威健股份有限公司
Welgene Biotech. Co.,Ltd.

DNA/RNA

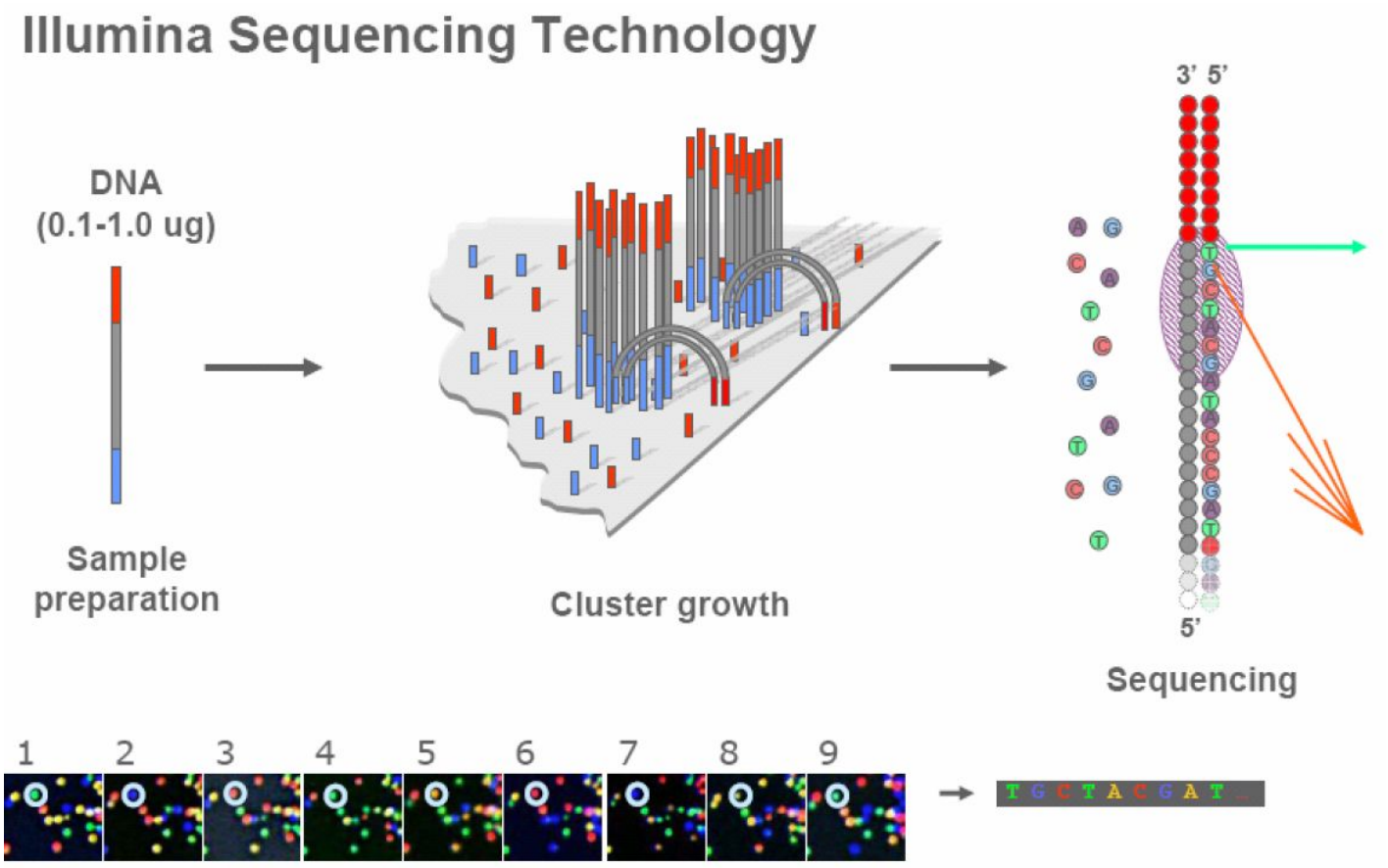


一次讀 10^8條序列, 每條 ~100bp

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Current (2nd) NGS Platform - Solexa

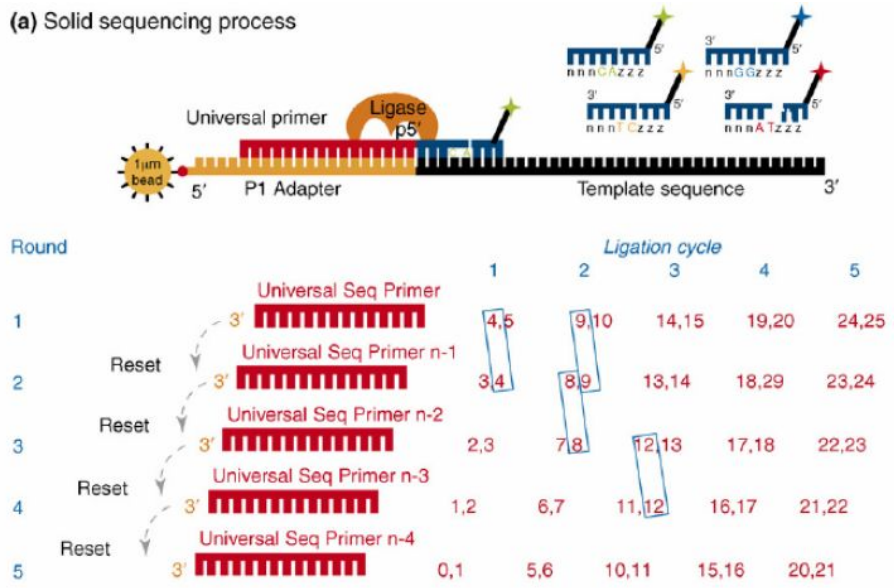
Illumina Solexa



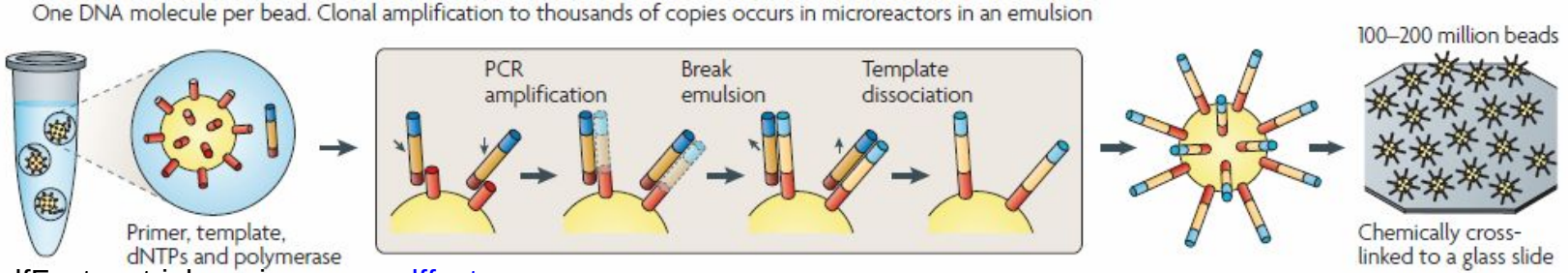
威健股份有限公司
Welgene Biotech. Co.,Ltd.

Current (2nd) NGS Platform - SOLiD

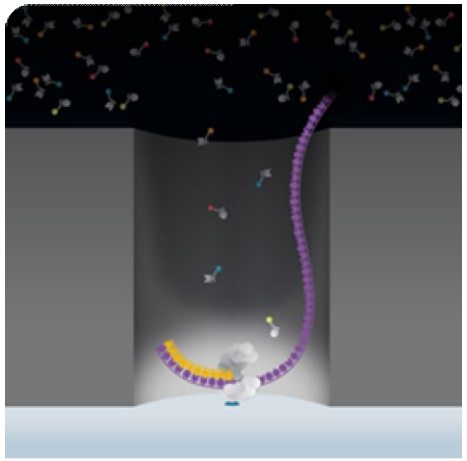
AB SOLiD

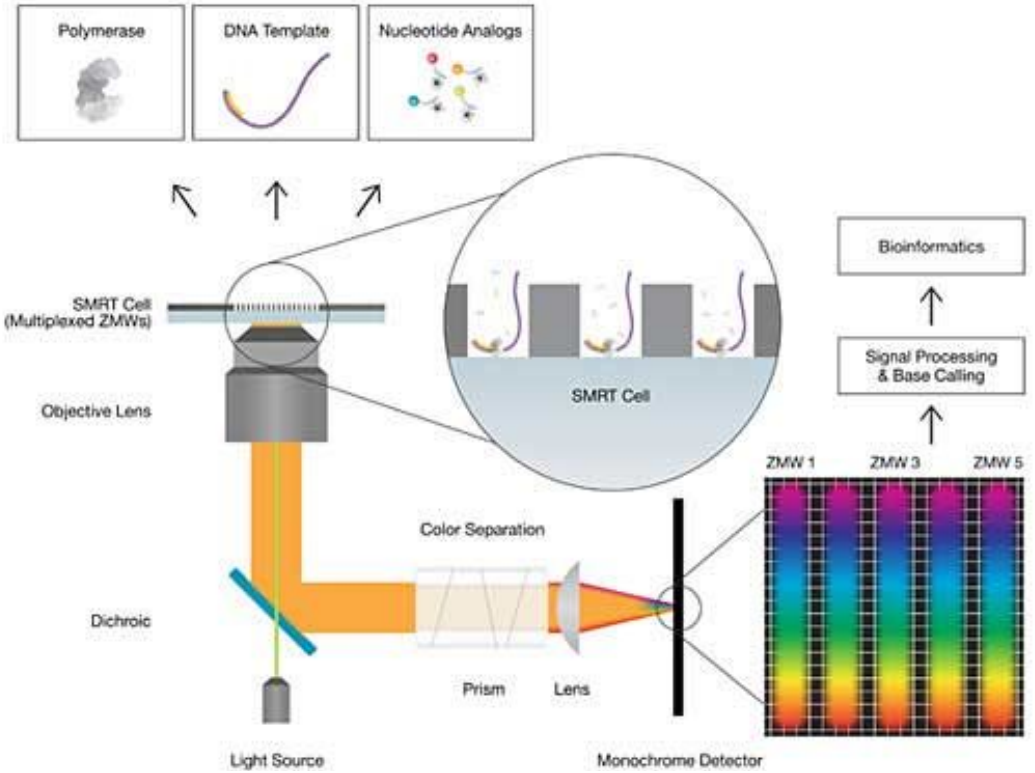


One DNA molecule per bead. Clonal amplification to thousands of copies occurs in microreactors in an emulsion

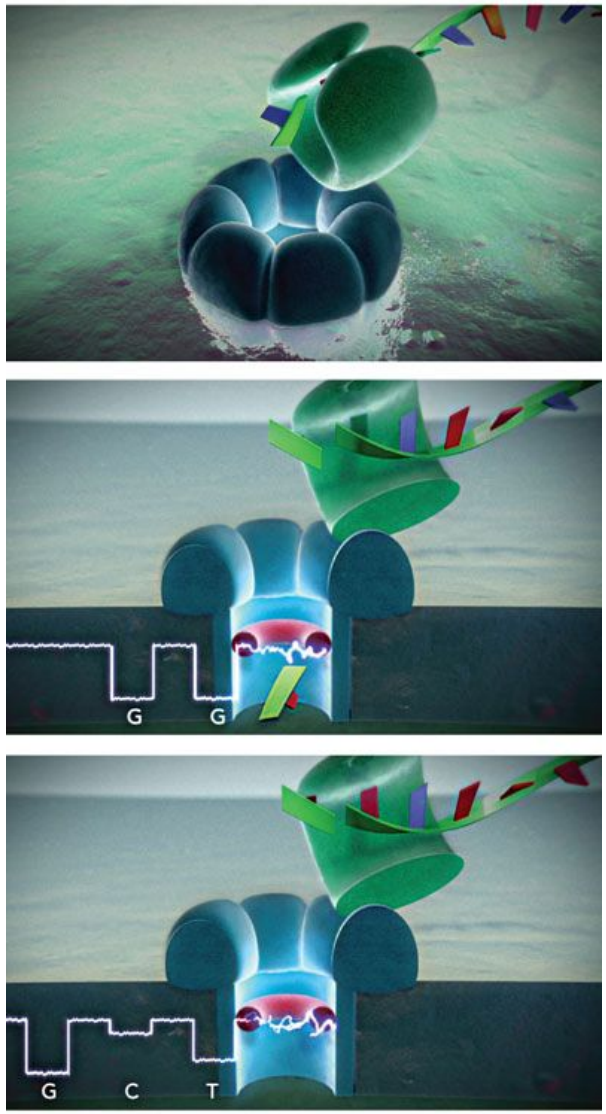


Next (3rd) NGS Technology-PacBio





Next (3rd) NGS Technology-NanoPore



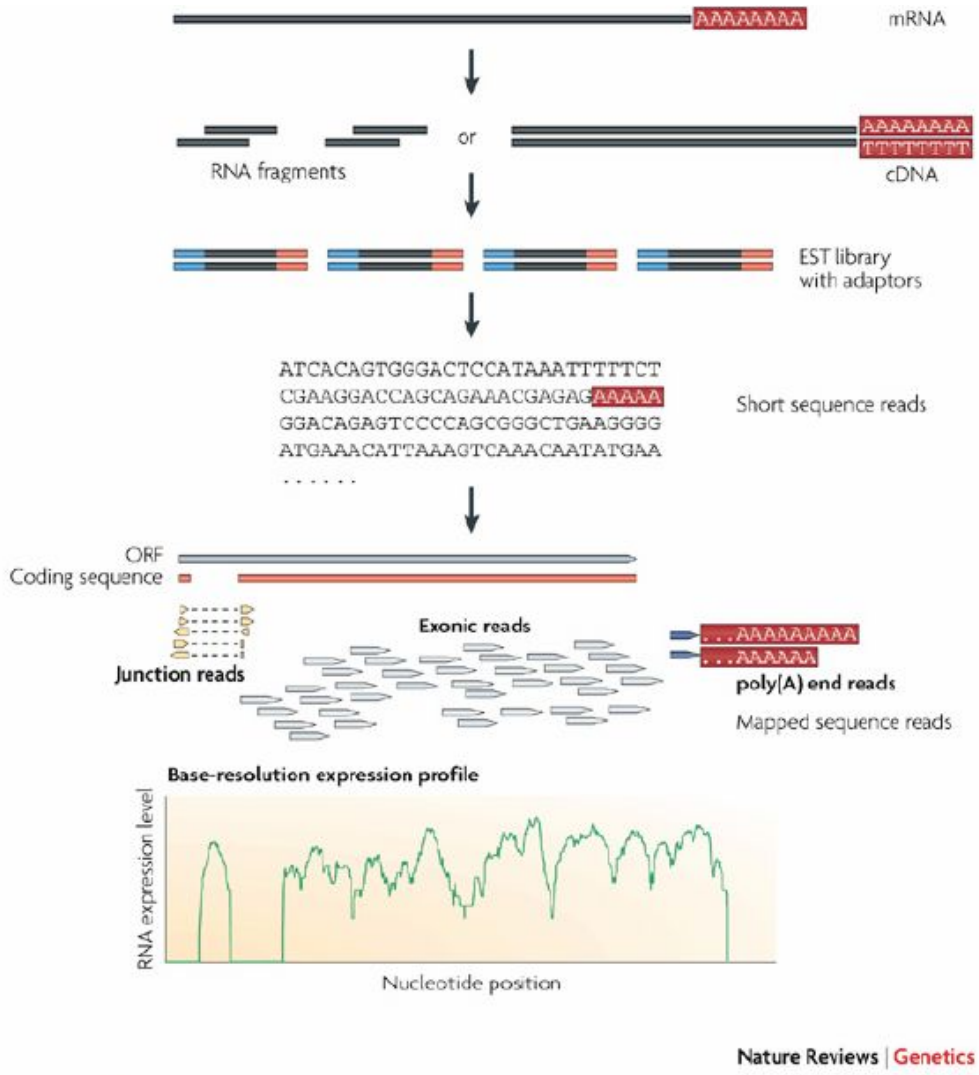
威健股份有限公司
Welgene Biotech. Co.,Ltd.

Applications

- **RNA-seq**: expression quantification, alternative splicing
- **ChIP-seq**: Identification of TF binding sites and broad regional interactions (e.g. histone modifications)
- **Exom-seq**: Exom region enrichment and sequencing
- **Small RNA-seq:** miRNA expression quantification
- **Whole genome sequencing**

威健股份有限公司
Welgene Biotech. Co.,Ltd.

RNA-seq



How RNA-seq works

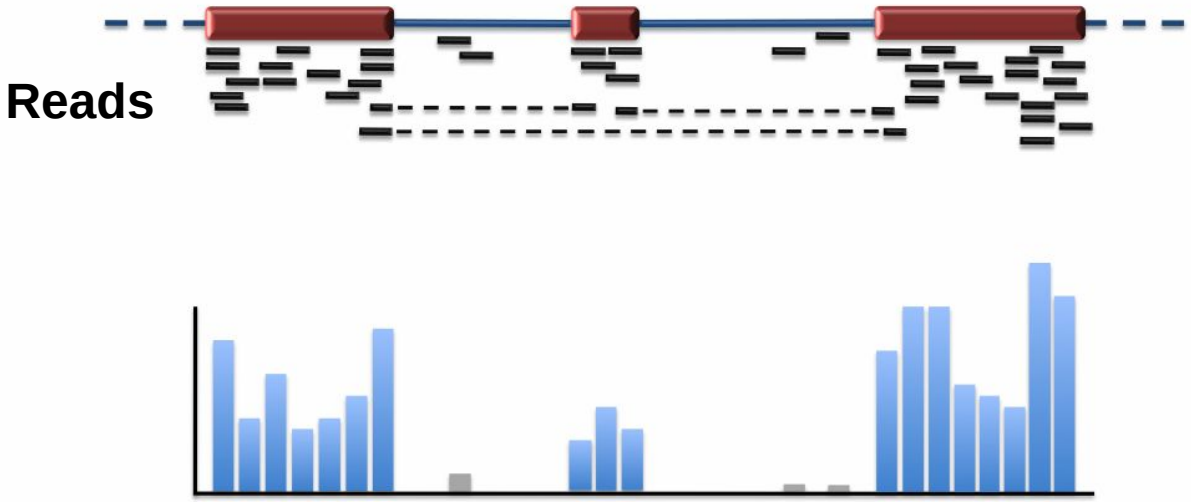
Sample preparation

Next generation sequencing (NGS)

Data analysis:
- ✓Mapping reads
- ✓Visualization (Gbrowser)
- ✓De novo assembly
- ✓Quantification

Figure from Wang et. al, **RNA-Seq: a revolutionary tool for transcriptomics,** Nat. Rev. Genetics 10, 57-63, 2009).

Reference RNA sequence



28509 rows

Search | Export results | RPKM Threshold

Rank	Gene	Location	RPKM ▲	Read count
14	rpsa	chr6: 8,274,537-8,419,743	2,262.1	246,929.0
13	mt-co3	chrM: 9,735-10,520	2,373.3	191,382.0
12	mt-co1	chrM: 6,425-7,975	2,738.6	435,776.0
11	ef1a	chr19: 43,128,766-43,132,081	2,883.5	516,210.0
10	mt-co2	chrM: 8,130-8,820	3,550.2	251,679.0
9	rplp0	chr5: 1,613,717-1,626,549	3,871.7	430,972.0
8	5_8S_rRNA	chrU: 169,855,501-169,855,656	3,946.4	63,160.2
7	zgc:193613	chr16: 22,497,632-22,498,922	3,972.3	270,598.0
6			5,188.0	2,238,130.0

RNA-seq獲得比Array更深入的資訊

- **Exon level expression / Splicing event**
- **Novel exon / intron discovery**
- **Point mutation**
- **Allelic specific expression**
- **Gene fusion (?)**

威健股份有限公司
Welgene Biotech. Co.,Ltd.

WELGENE

OPEN ACCESS Freely available online

PLoS one

Tumor Transcriptome Sequencing Reveals Allelic Expression Imbalances Associated with Copy Number Alterations

Brian B. Tuch[1☯], Rebecca R. Laborde[2☯], Xing Xu[1], Jian Gu[3], Christina B. Chung[1], Cinna K. Monighetti[1], Sarah J. Stanley[1], Kerry D. Olsen[4], Jan L. Kasperbauer[4], Eric J. Moore[4], Adam J. Broomer[1], Ruoying Tan[1], Pius M. Brzoska[1], Matthew W. Muller[1], Asim S. Siddiqui[1], Yan W. Asmann[5], Yongming Sun[1], Scott Kuersten[3], Melissa A. Barker[1], Francisco M. De La Vega[1]*, David I. Smith[2]*

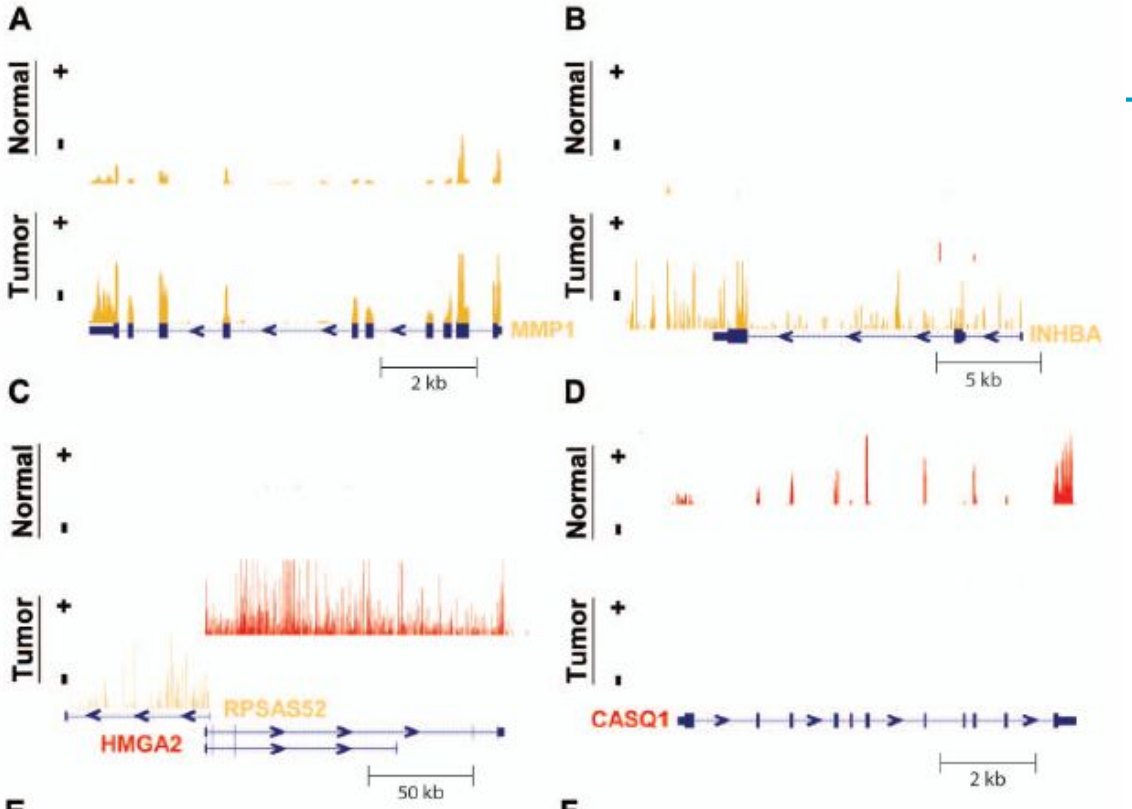
1 Life Technologies Inc., Foster City, California, United States of America, **2** Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, United States of America, **3** Life Technologies Inc., Austin, Texas, United States of America, **4** Department of Otorhinolaryngology, Mayo Clinic, Rochester, Minnesota, United States of America, **5** Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, United States of America

Abstract

Due to growing throughput and shrinking cost, massively parallel sequencing is rapidly becoming an attractive alternative to microarrays for the genome-wide study of gene expression and copy number alterations in primary tumors. The sequencing of transcripts (RNA-Seq) should offer several advantages over microarray-based methods, including the ability to detect somatic mutations and accurately measure allele-specific expression. To investigate these advantages we have applied a novel, strand-specific RNA-Seq method to tumors and matched normal tissue from three patients with oral squamous cell carcinomas. Additionally, to better understand the genomic determinants of the gene expression changes observed, we have sequenced the tumor and normal genomes of one of these patients. We demonstrate here that our RNA-Seq method accurately measures allelic imbalance and that measurement on the genome-wide scale yields novel insights into cancer etiology. As expected, the set of genes differentially expressed in the tumors is enriched for cell adhesion and differentiation functions, but, unexpectedly, the set of allelically imbalanced genes is also enriched for these same cancer-related functions. By comparing the transcriptomic perturbations observed in one patient to his underlying normal and tumor genomes, we find that allelic imbalance in the tumor is associated with copy number mutations and that copy number mutations are, in turn, strongly associated with changes in transcript abundance. These results support a model in which allele-specific deletions and duplications drive allele-specif

Whole Transcriptome Sequencing

	Patient 8				Patient 33				Patient 51			
	Normal		Tumor		Normal		Tumor		Normal		Tumor	
Total reads	129 M		188 M		229 M		256 M		227 M		199 M	
Unaligned	39 M	30%	63 M	34%	48 M	21%	69 M	27%	64 M	28%	84 M	43%
Aligned to filter database	48 M	37%	76 M	41%	118 M	52%	121 M	47%	75 M	33%	48 M	24%
Aligned to rDNA (in filter database)	46 M	36%	69 M	36%	111 M	48%	115 M	45%	66 M	29%	41 M	21%
Aligned to genome	43 M	33%	48 M	26%	63 M	27%	66 M	26%	88 M	39%	66 M	33%
Uniquely aligned to genome	25 M	19%	21 M	11%	39 M	17%	38 M	15%	56 M	25%	40 M	20%
Uniquely aligned to genome within a RefSeq exon	8 M	6%	7 M	4%	16 M	7%	14 M	5%	22 M	10%	15 M	8%
RefSeq exons with at least one alignment	133,207		102,540		136,617		127,464		150,333		143,026	
RefSeq exons with 2+ non-clonal alignments	114,560		80,905		115,275		108,984		136,674		128,546	
Coverage of RefSeq exons	5.6x		5.1x		11.5x		10.3x		15.9x		11.0x	

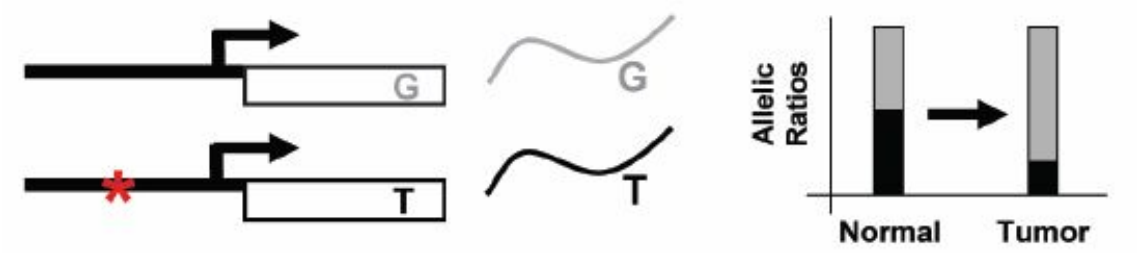


E

GO Category Description	Count in Category	Count in Overlap	P-value
proteinaceous extracellular matrix	171	23	9.5E-08
cell adhesion	392	28	2.6E-05
digestion	20	7	5.4E-05
metalloendopeptidase activity	55	10	1.6E-04
collagen catabolic process	10	5	1.9E-04
cell motility	213	17	5.3E-04
organ development	634	40	8.9E-06
cell-cell signaling	202	15	5.6E-03
nervous system development	340	26	1.2E-05
intrinsic to plasma membrane	496	40	2.6E-09
transmembrane transporter activity	365	26	7.4E-05
negative regulation of follicle-stimulating hormone secretion	3	3	8.5E-04
steroid dehydrogenase activity	11	4	5.7E-03
potassium ion binding	33	6	8.3E-03

F

GO Category Description	Count in Category	Count in Overlap	P-value
extracellular region part	305	31	6.0E-11
proteinaceous extracellular matrix	171	19	1.1E-07
cell adhesion	392	34	5.4E-09
cell-matrix junction	21	5	4.6E-03
keratinocyte differentiation	23	6	1.0E-03
epidermis development	85	11	1.0E-03
muscle contraction	89	35	9.8E-28
striated muscle contraction	23	17	1.7E-19
striated muscle cell development	5	4	1.2E-04
myofibril	32	21	1.9E-22
sarcoplasmic reticulum	15	10	1.1E-10
calcium ion binding	475	40	4.6E-10
actin cytoskeleton	157	30	4.8E-14
response to wounding	193	15	3.9E-03
oxygen binding	9	4	1.6E-03

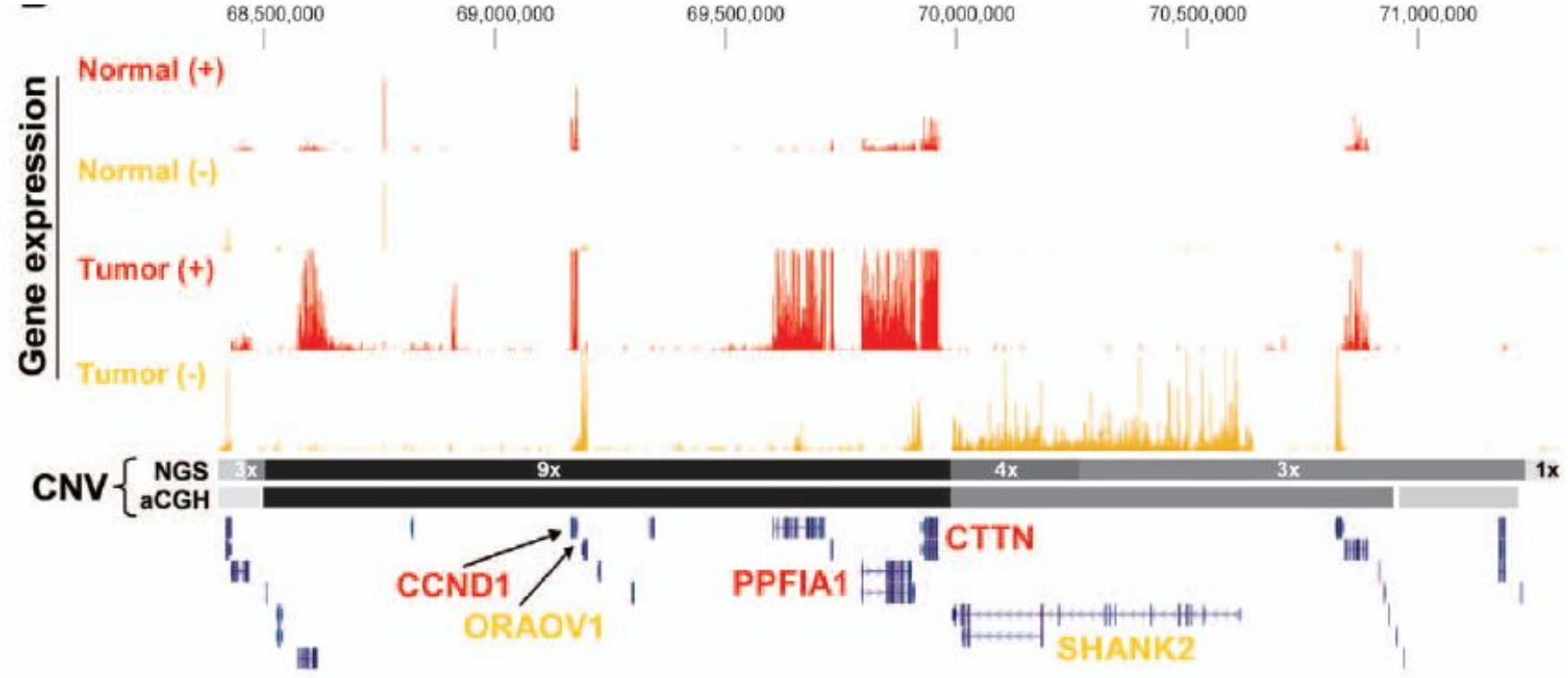


D

	Allelic Imbalance?			Gene Expression Tumor vs Normal			Allelic Imbalance Details			Adhesion/ECM	
	8	33	51	8	33	51	8	33	51		
CD44	●	●	●	-0.9	0.5	1.3	Syn, I	K->R	I/NSS	●	Cd44 molecule
DSC3	●	●	●	0.6	0.5	1.1	3	3, K->R, S->T	3 (3)	●	desmocollin 3 isoform Dsc3a preproprotein
DST	●	●	●	-1.3	0.1	0.0	L->F	Syn, I/NSS	G->R	●	dystonin isoform 1e precursor
MALAT1	●	●	●	0.3	-0.6	0.3	NC	NC	NC		metastasis associated lung adenocarcinoma 1
PERP	●	●	●	0.7	0.0	1.7	3 (2)	Syn, R->P	3	●	PERP, TP53 apoptosis effector
ALDH3A1	●	●		-0.9	1.4	-2.4	3, Syn	D->G			aldehyde dehydrogenase 3A1
CCND1	●	●		3.7	-0.1	1.2	3 (2)	3			cyclin D1
CTNND1	●	●		1.2	0.0	1.0	I, 3	I, 3		●	catenin, delta 1 isoform 1A
DSP	●	●		-0.1	-0.3	1.6	I->F, Syn	Syn		●	desmoplakin isoform I
FAT2		●	●	0.1	0.3	1.9		Syn	S->L, I, I->M	●	FAT tumor suppressor homolog 2
GJB2		●	●	0.6	-0.7	1.2		W->C	I, 3	●	gap junction protein, beta 2
H19	●	●		-0.6	-2.1	-2.0	NC (6)	NC			H19
HLA-A	●		●	-0.3	0.3	0.5	Syn, F->L	3			major histocompatibility complex, class I, A
HSPG2		●	●	0.1	0.4	-1.1		R->H	Syn	●	heparan sulfate proteoglycan 2
KRT1	●		●	0.8	-3.0	1.5	3		Syn		keratin complex 2, basic, gene 1
KRT6A	●	●		-1.8	-0.1	0.9	3 (2), I (3)	I, 3			keratin 6A
KRT14	●		●	-0.5	0.4	0.9	Syn, I/NSS		I/NSS		keratin 14
MAP4		●	●	-1.5	-0.1	-1.0		3	3		microtubule-associated protein 4
NCRNA00084	●	●		0.4	0.1	-0.3	NC (5)	NC (4)			NCRNA00084
PITX1	●	●		-0.9	-0.7	-2.5	Syn, I	Syn			paired-like homeodomain 1
PKP1	●		●	0.0	0.6	1.5	3, I, I/NSS		I	●	plakophilin 1
RUNX1	●		●	0.6	1.0	-0.4	I		I		runt-related transcription factor 1
SKP1		●	●	0.2	-0.6	-0.4		5	5		S-phase kinase-associated protein 1

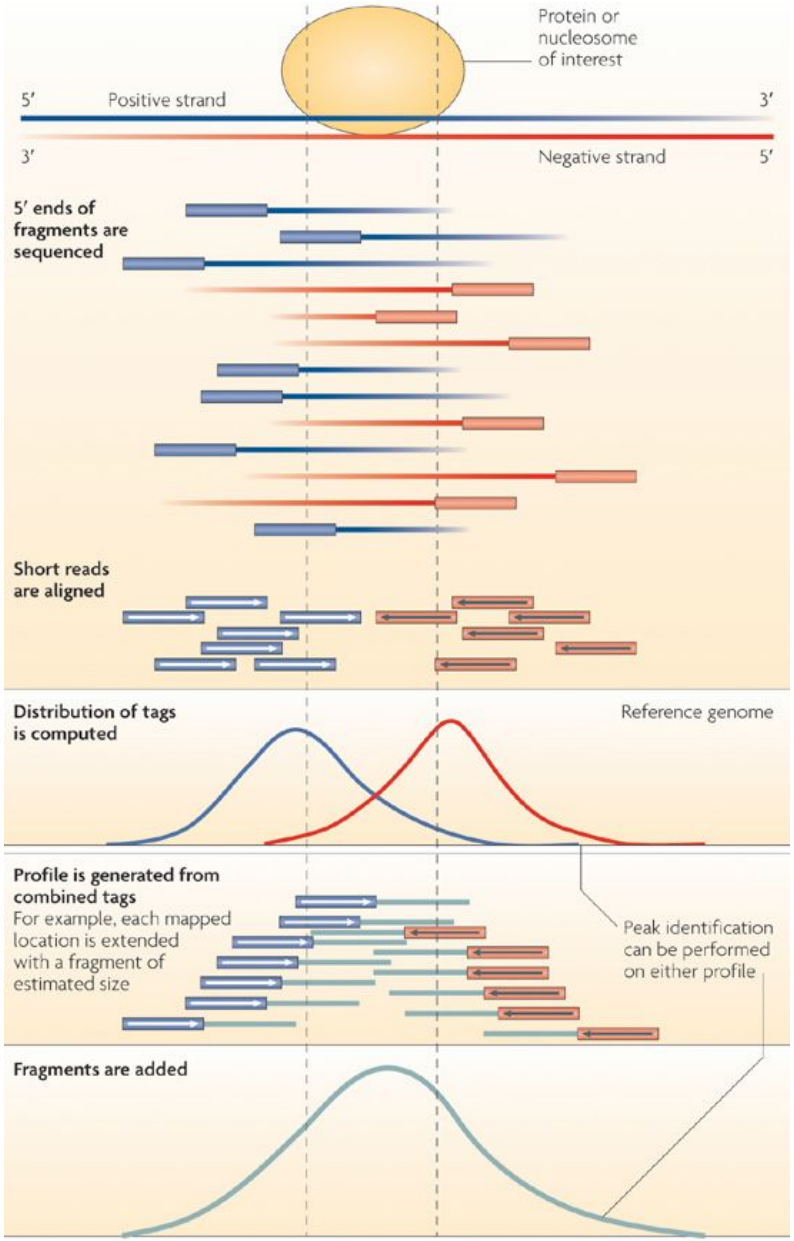
impacted. Also noted is whether or not the gene is involved in cell adhesion or is a component of the ECM. Key: 3 = 3' UTR, I = Intronic, NSS = Near Splice Site, Syn = Synonymous, → = Non-Synonymous.

Mate-Pair Genome Sequencing **vs** Whole Transcriptome Sequencing



威健股份有限公司
Welgene Biotech. Co.,Ltd.

ChIP-seq



Park, P.J., Nat. Genetics (2009) 10:669-680

Nature Reviews | Genetics

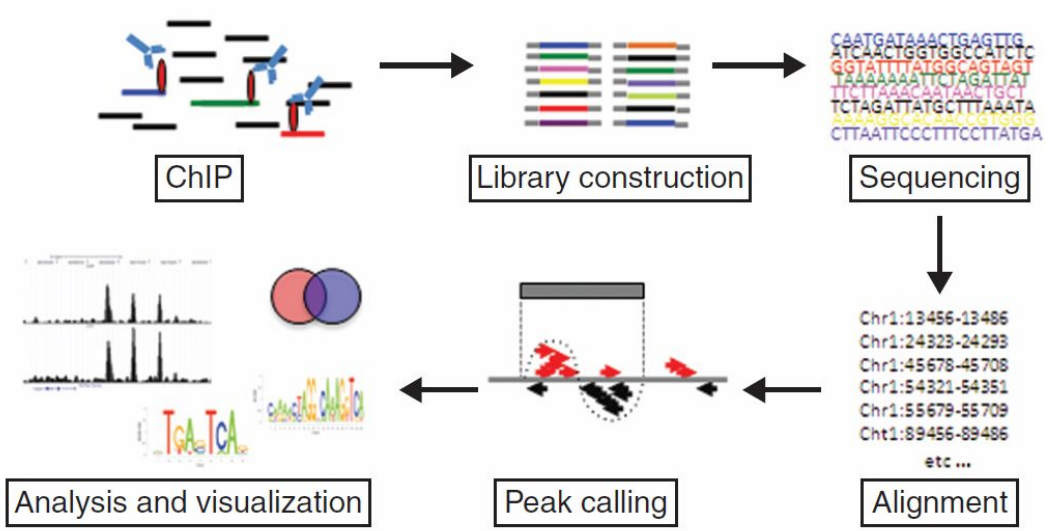


Figure 1. Flow scheme of the central steps in the ChIP-seq procedure.

威健股份有限公司
Welgene Biotech. Co.,Ltd.

ChIP-seq 特色與應用

特色

- IP DNA只要 10-20ng即可定序
- 解析度高
- 敏感度高

應用

- Transcription factor Study
- DNA Methylation (MeDIP)
- Histone Modification

威健股份有限公司
Welgene Biotech. Co.,Ltd.

BMC Genomics

BioMed Central

Methodology article

Open Access

Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing

Philippe Lefrançois[1], Ghia M Euskirchen[1], Raymond K Auerbach[2], Joel Rozowsky[3], Theodore Gibson[3], Christopher M Yellman[1], Mark Gerstein[2,3] and Michael Snyder*[1,2,3]

Address: [1]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA, [2]Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06520, USA and [3]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

Email: Philippe Lefrançois - philippe.lefrancois@yale.edu; Ghia M Euskirchen - ghia.euskirchen@yale.edu; Raymond K Auerbach - raymond.auerbach@yale.edu; Joel Rozowsky - joel.rozowsky@yale.edu; Theodore Gibson - theodore.gibson@yale.edu; Christopher M Yellman - christopher.ye

* Corresponding author

Published: 21 January 2009

BMC Genomics 2009, **10**:37 doi:10.1186/

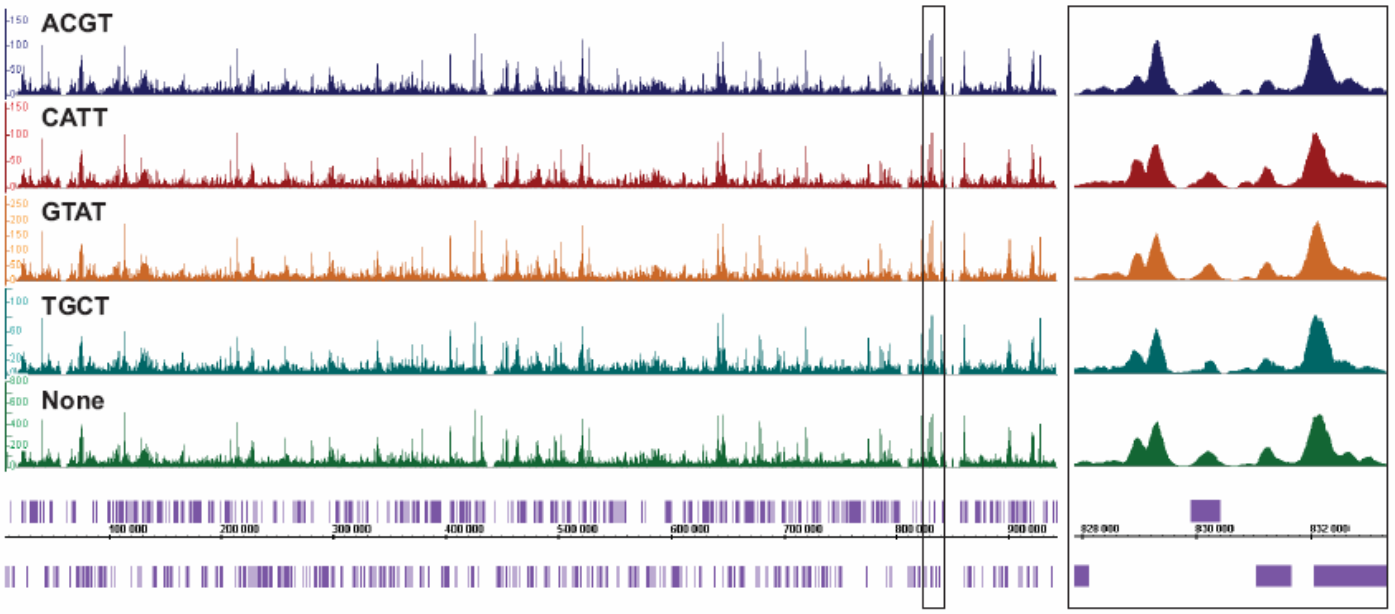
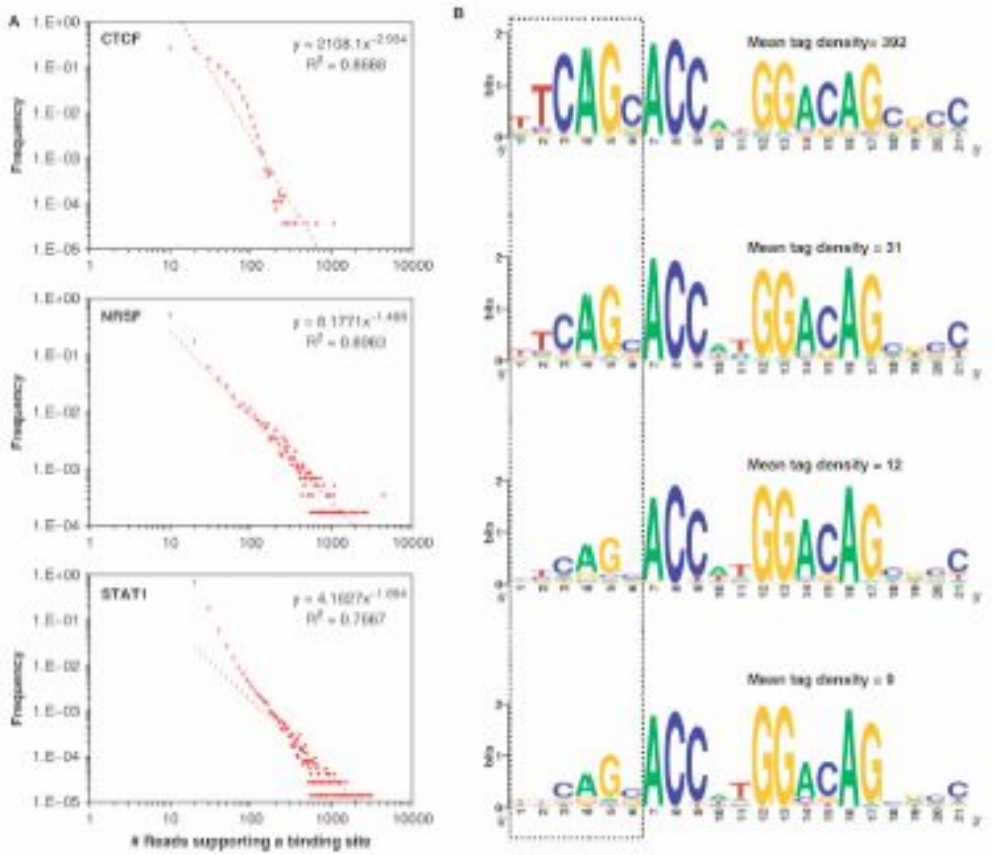


Figure 2
Comparison of input DNA signal tracks among all four barcoded adapters relative to standard Illumina adapters. An input sample was split in five aliquots. Four were barcoded differentially (top four lanes) and one had non-barcoded,

ChIP-seq read density is quantitative measure of binding level



Jothi, R., et. al. (2008) Nucleic Acids Res. 36:5221-31

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Published in final edited form as:
Nature. 2009 February 12; 457(7231): 854–858. doi:10.1038/nature07730.

ChIP-seq accurately predicts tissue-specific activity of enhancers

Axel Visel[1,*], **Matthew J. Blow**[1,2,*], **Zirong Li**[3], **Tao Zhang**[2], **Jennifer A. Akiyama**[1], **Amy Holt**[1], **Ingrid Plajzer-Frick**[1], **Malak Shoukry**[1], **Crystal Wright**[2], **Feng Chen**[2], **Veena Afzal**[1], **Bing Ren**[3], **Edward M. Rubin**[1,2], and **Len A. Pennacchio**[1,2]

[1]Genomics Division, MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

[2]US Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA.

[3]Ludwig Institute for Cancer Research, University of California San Diego (UCSD) School of Med

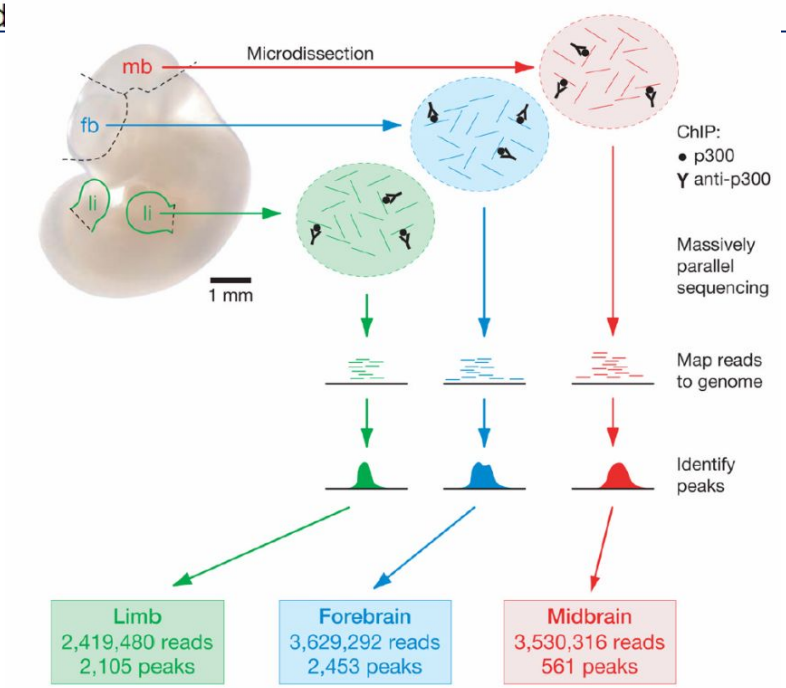


Figure 1. Tissue dissection boundaries, overview of the ChIP-seq approach and summary of p300 results

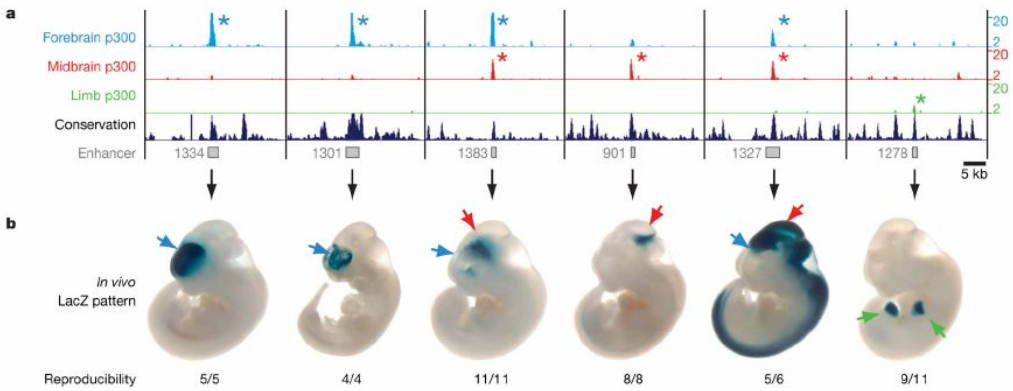
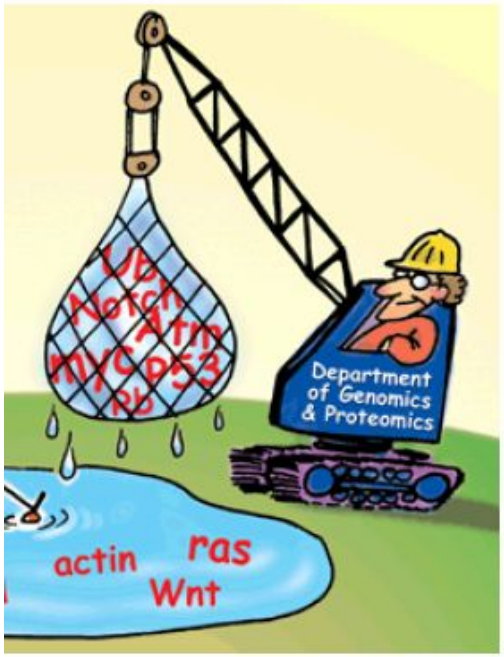


Figure 3. Examples of successful prediction of *in vivo* enhancers by p300 binding in embryonic tissues

威健股份有限公司
Welgene Biotech. Co.,Ltd.

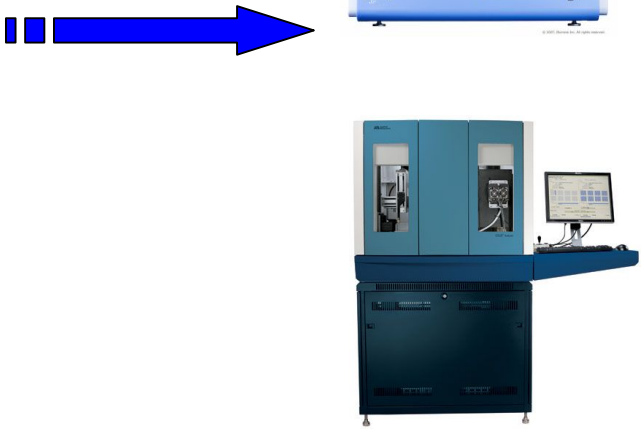
Exon Capture & Sequencing



Newer Methods Appropriate

High Throughput Discovery

SureSelect™ Platform + Next-gen Sequencing



將所有exon (50 M base區域) 抓下後進行定序

威健股份有限公司
Welgene Biotech. Co.,Ltd.

2009, February, Nature Biotechnology!!
Underlying Technology of SureSelect™

ARTICLES

nature biotechnology

Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing

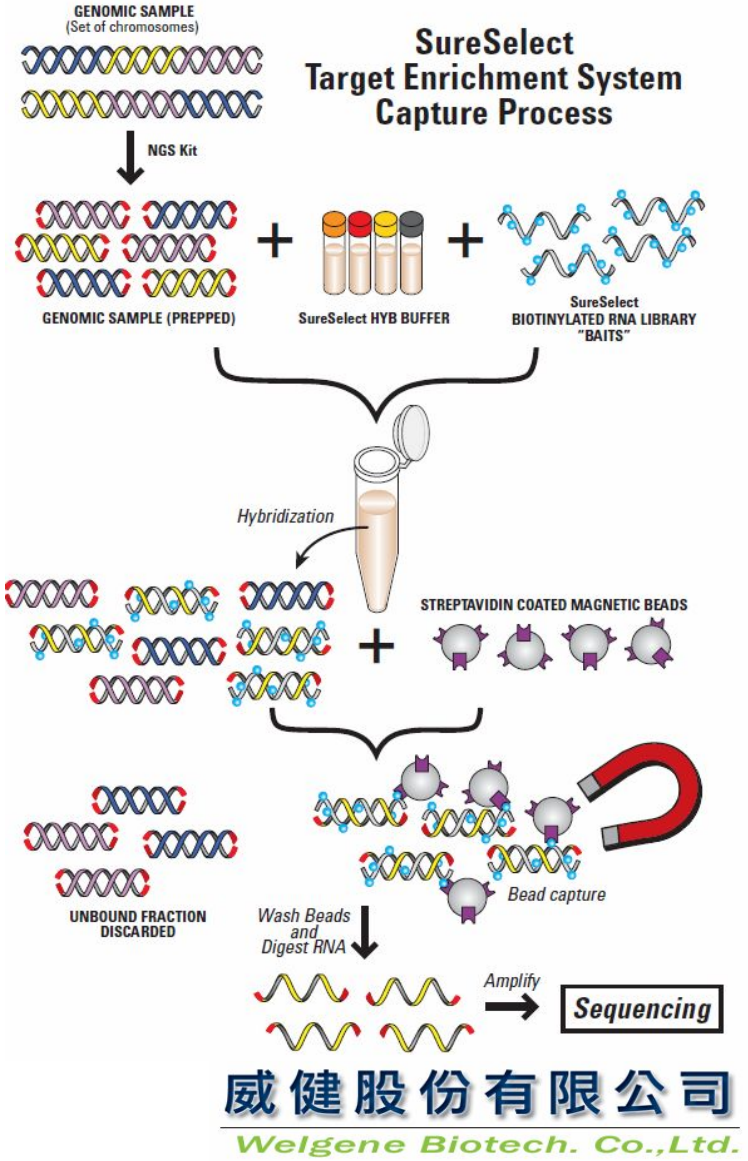
Andreas Gnirke[1], Alexandre Melnikov[1], Jared Maguire[1], Peter Rogov[1], Emily M LeProust[2], William Brockman[1,5], Timothy Fennell[1], Georgia Giannoukos[1], Sheila Fisher[1], Carsten Russ[1], Stacey Gabriel[1], David B Jaffe[1], Eric S Lander[1,3,4] & Chad Nusbaum[1]

Targeting genomic loci by massively parallel sequencing requires new methods to enrich templates to be sequenced. We developed a capture method that uses biotinylated RNA 'baits' to fish targets out of a 'pond' of DNA fragments. The RNA is transcribed from PCR-amplified oligodeoxynucleotides originally synthesized on a microarray, generating sufficient bait for multiple captures at concentrations high enough to drive the hybridization. We tested this method with 170-mer baits that target >15,000 coding exons (2.5 Mb) and four regions (1.7 Mb total) using Illumina sequencing as read-out. About 90% of uniquely aligning bases fell on or near bait sequence; up to 50% lay on exons proper. The uniformity was such that ~60% of target bases in the exonic 'catch', and ~80% in the regional catch, had at least half the mean coverage. One lane of Illumina sequence was sufficient to call high-confidence genotypes for 89% of the targeted exon space.

The development and commercialization of a new generation of increasingly powerful sequencing methodologies and instruments[1–4] have lowered the cost per nucleotide of sequencing data by several orders of magnitude. Within a short time, several individual human

have been tested on target sets complex enough to match the scale of current next-generation sequencing instruments.

The first method, microarray capture[9,12,13], uses hybridization to arrays containing synthetic oligonucleotides that match the target

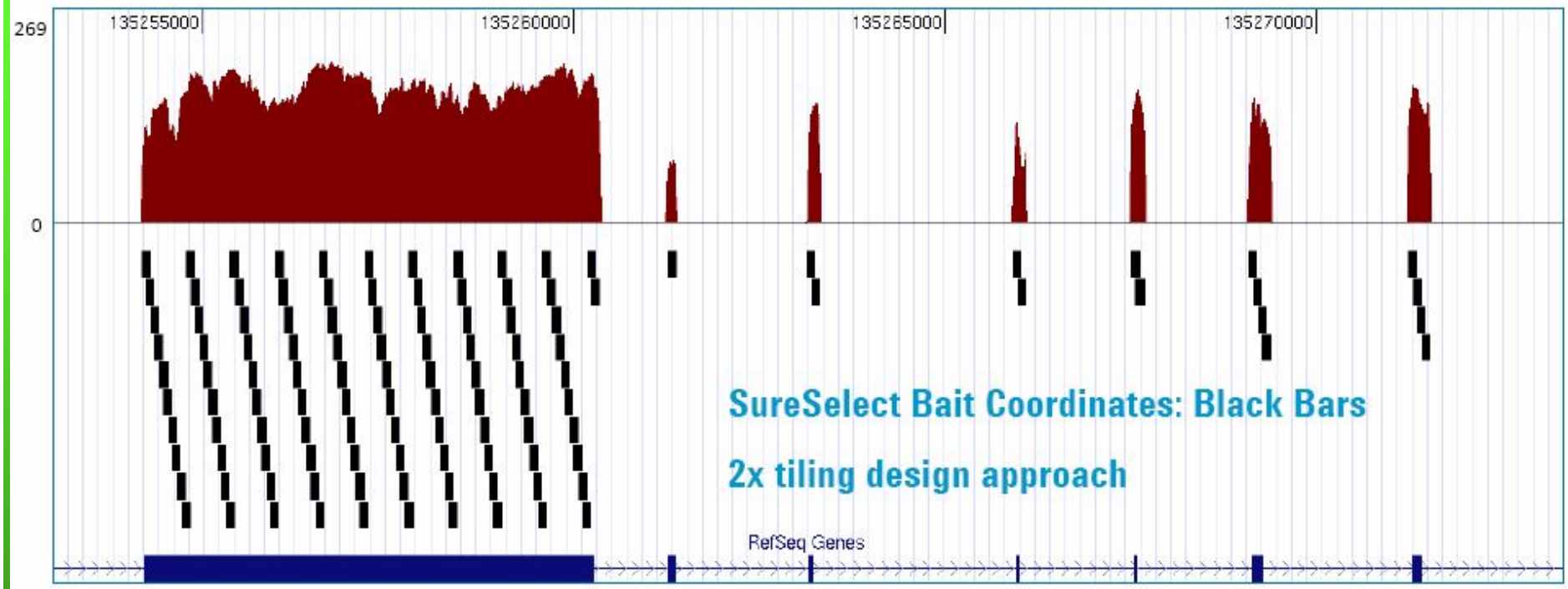


SureSelect Target Enrichment Kits

Product	Target amount (catalog number)	Reactions/kit	Product Definition
Human All Exon v1	38 Mb	5-10,000	CCDS Sept. 2008
Human All Exon v2	44 Mb	5-10,000	CCDS Sept. 2009 + additional RefSeq
Human All Exon 50Mb	50 Mb	5 10,000	GENCODE content – most comprehensive coverage Multiplexable
Human All Exon Plus	38-50 Mb + up to 6.8 Mb of custom content	5-10,000	Add custom content to All Exon catalog content
Kinome	3.2 Mb	5-10,000	All kinases
Mouse All Exon	50 Mb	5-10,000	Mouse Exome
Indexed custom content	<0.2 Mb, 0.2-0.49 Mb, 0.5-1.49 Mb, 1.5- 2.9 Mb 3 – 6.8 Mb	10 – 5,000	

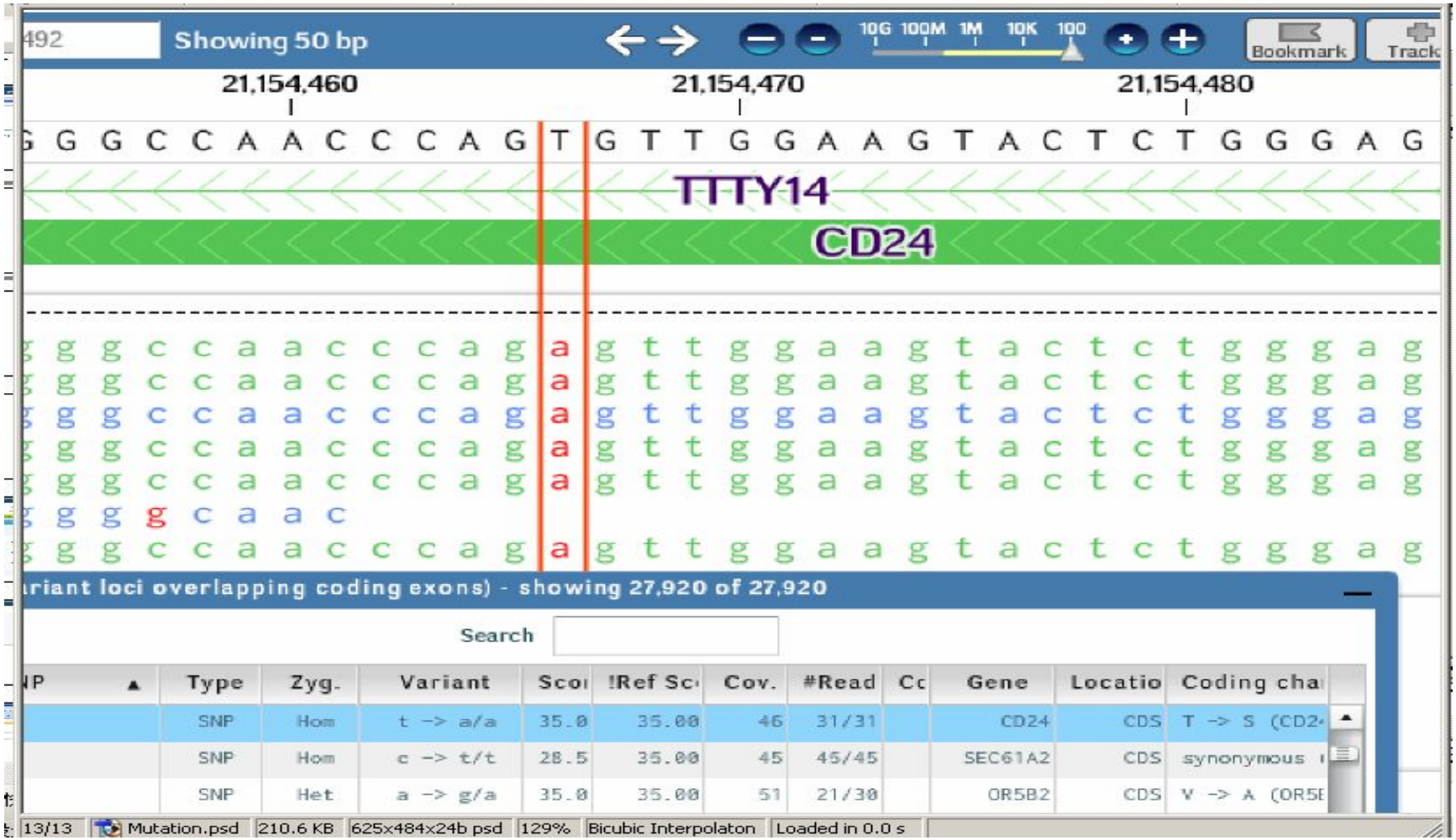
Welgene Biotech. Co.,Ltd.

Agilent SureSelect™ Target Enrichment Efficacy: Sample Coverage Plot



This plot (generated in the UCSC Genome Browser) shows sequence coverage data for a representative portion of the Agilent SureSelect Target Enrichment demo kit, which covers the RefSeq exons on the non-pseudoautosomal portion of Chromosome X using 2x tiling. In red is the log of the sequence coverage. Bait coordinates are denoted in black and the exons that constitute the target region are in blue. Sequence coverage is higher for those exons that are covered by more than one overlapping bait

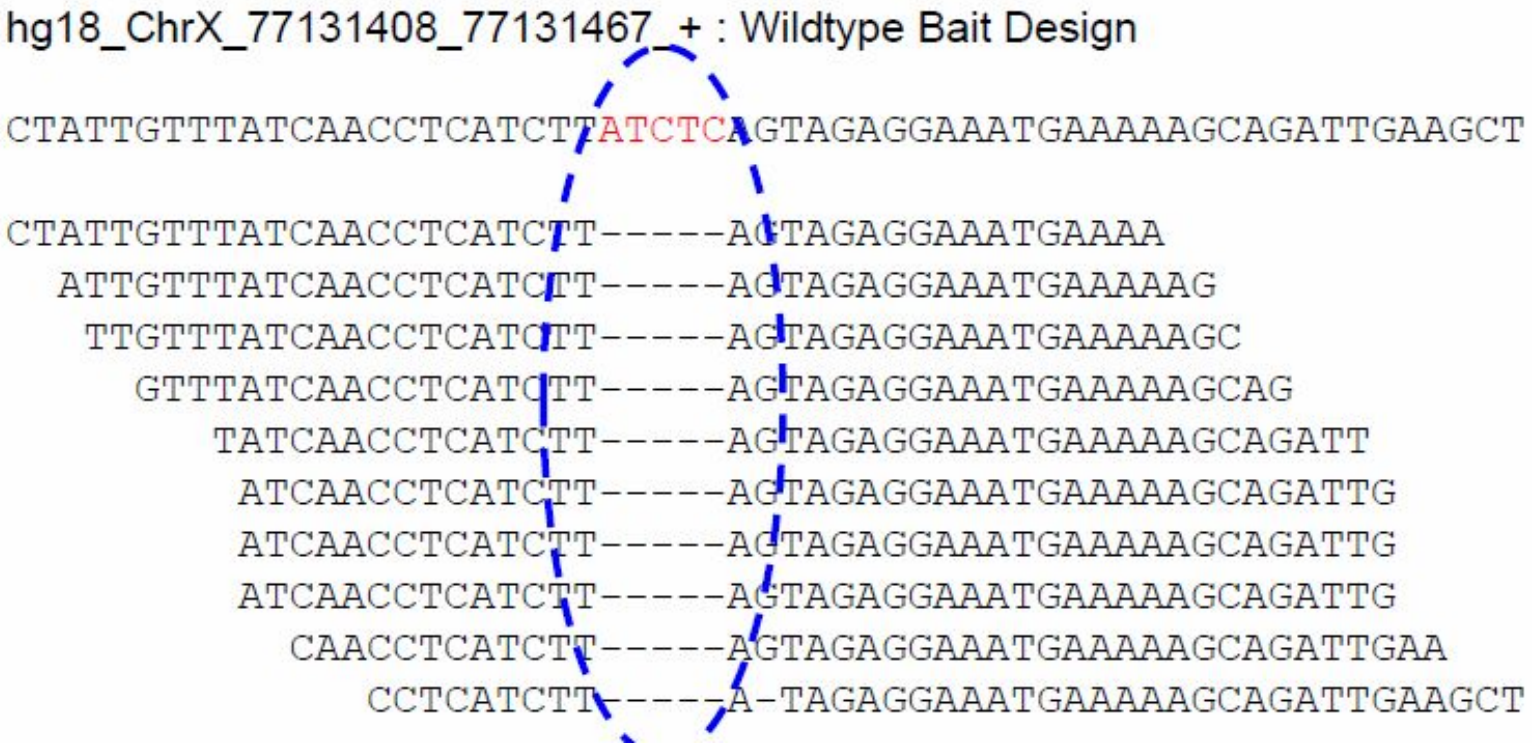
威健股份有限公司
Welgene Biotech. Co.,Ltd.



威健股份有限公司
Welgene Biotech. Co.,Ltd.

Efficient Capture of 5bp deletion on the X-Chromosome: Menke's Syndrome

SureSelect™ Target Enrichment Kit Efficiently Captures 5 bp Mutant
Readout on Illumina GA



威健股份有限公司
Welgene Biotech. Co.,Ltd.

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Small RNA-seq

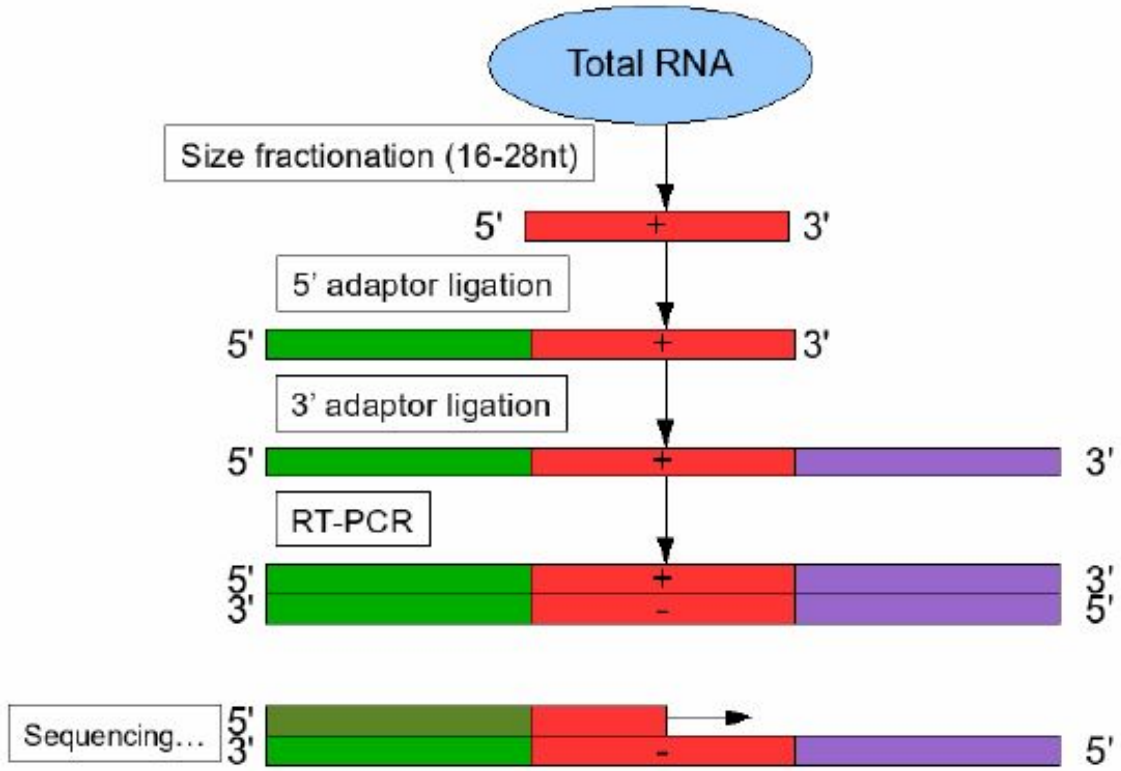


Figure 5-1 Experimental pipeline of small RNA analysis

Vaz et al. BMC Genomics 2010, 11:288
http://www.biomedcentral.com/1471-2164/11/288

BMC Genomics

RESEARCH ARTICLE Open Access

Analysis of microRNA transcriptome by deep sequencing of small RNA libraries of peripheral blood

Ahmad[2], Pratibha Sharma[3], Rashi Gupta[1], Lalit Kumar[3], Ritu Kulshreshtha[2] and

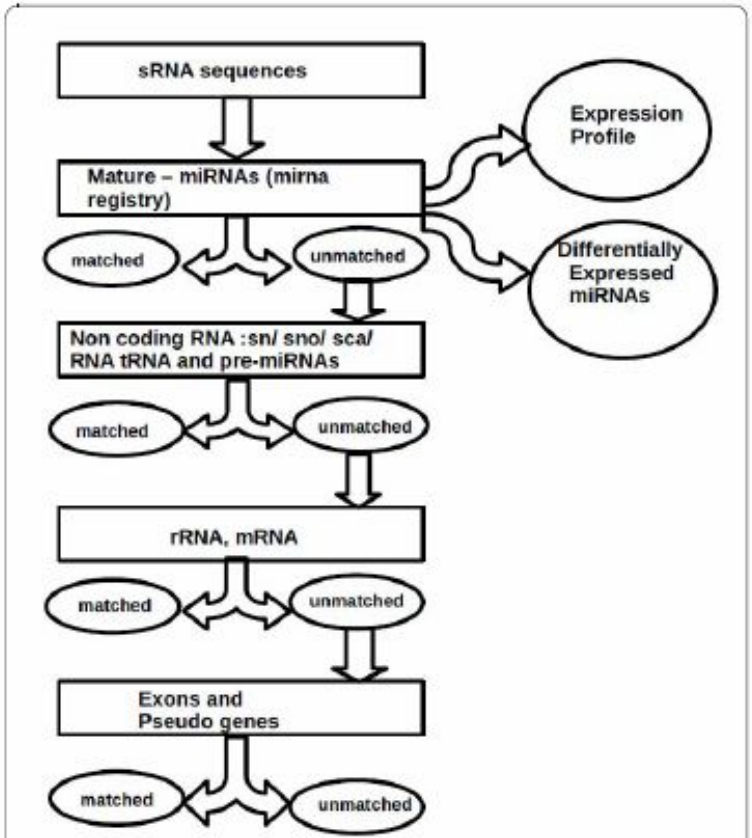
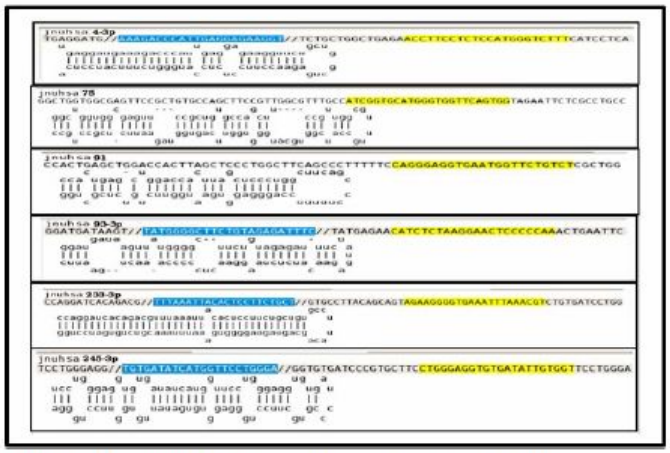


Figure 1 Flowchart describing the elimination pipeline used to filter out the indicated sequences from the library of sRNA sequences. The sequences were matched using an "in house" devel-

A

Novel miRNA ID	Sequence	Length	minor /star seq	CSHMM score	C ID score	miRdeep score	MiPred score	Chromosomal Location
jnuhsa 4-3p	ACCTTCCTCTCCATGGGTCTTT	22	present	0.26	-0.5	-	Real 76.9%	Chr22 NT_011525 (956796..964869) (-)
jnuhsa75	ATCGGTGCATGGGTGGTTCAGT GG	24	-	0.23	-0.6	-	Real 72.20%	Chr1 NT_004487 (115253.61..119256 48) (+)
jnuhsa 91	CAGGGAGGTGAATGGTTCTGTC T	23	-	0.25	-0.61	-	Real 63.0%	Chr10 NT_030859 (576049..579121) (+)
jnuhsa 93-3p	CATCTCTAAGGAACTCCCCCAA	22	present	0.24	-0.6	-	Real 62.50%	Chr1 NT_004610 (5787..9869) (+)
jnuhsa233-3p	AGAAGGGGTGAAATTTAAACGT	22	present	0.33	-0.48	29.4	Real 78.8%	Chr16 NT_010393 (6308044.. 6308529) (+)
jnuhsa 245-3p	CTGGGAGGTGTGATATTGTGGT	22	present	0.24	-0.57	5.4	Real 73.10%	Chr9 NT_019501 (650141..650218) (-)

B



C

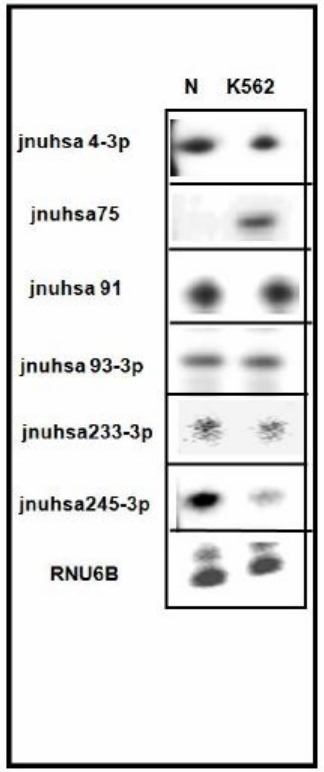


Figure 11 Predicted novel miRNAs. A. A partial list of novel miRNAs predicted from deep sequencing data is displayed along with chromosomal

OPEN ACCESS Freely available online

PLoS one

Characterization of the Melanoma miRNAome by Deep Sequencing

Mitchell S. Stark[1], Sonika Tyagi[1], Derek J. Nancarrow[1], Glen M. Boyle[2], Anthony L. Cook[4], David C. Whiteman[3], Peter G. Parsons[2], Christopher Schmidt[5], Richard A. Sturm[4], Nicholas K. Hayward[1]*

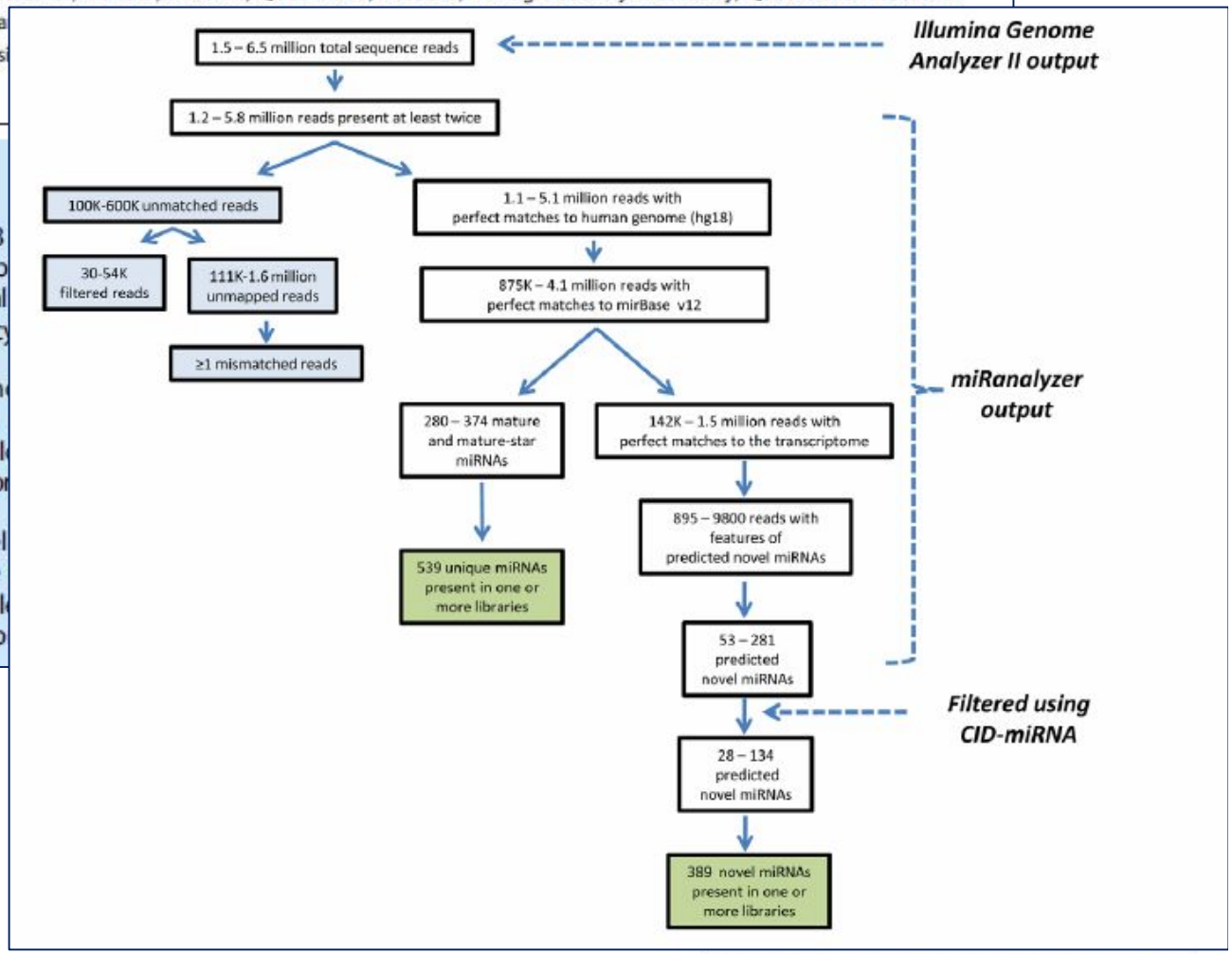
1 Oncogenomics Laboratory, Queensland Institute of Medical Research, Herston, Brisbane, Queensland, Australia, **2** Drug discovery Laboratory, Queensland Institute of Medical Research, Herston, Brisbane, Queensland, Australia, **3** Ca
4 Melanogenix Group, Institute for Molecular Bioscience, Universi
of Medical Research, Herston, Brisbane, Queensland, Australia

Abstract

Background: MicroRNAs (miRNAs) are 18–23
specific manner. Little is known about the rep
therefore undertook a comprehensive anal
melanoblasts, melanocytes, congenital nevoc

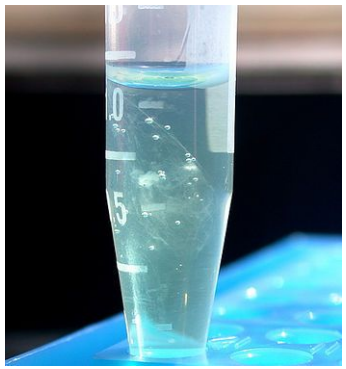
Methodology/Principal Findings: We sequen
massively parallel sequencing approach of a
known mature and mature-star sequences, al
common to 2 or more libraries and 3 were p

Conclusions/Significance: Some of the novel
could be used as biomarkers to assist in the
blood. Follow up studies of the functional rol
shed further light on the development and p



Welgene Biotech. Co.,Ltd.

Welgene NGS Service Workflow



DNA/RNA from Customer

Bioanalyzer 2100
QC by Welgene

Sequencing

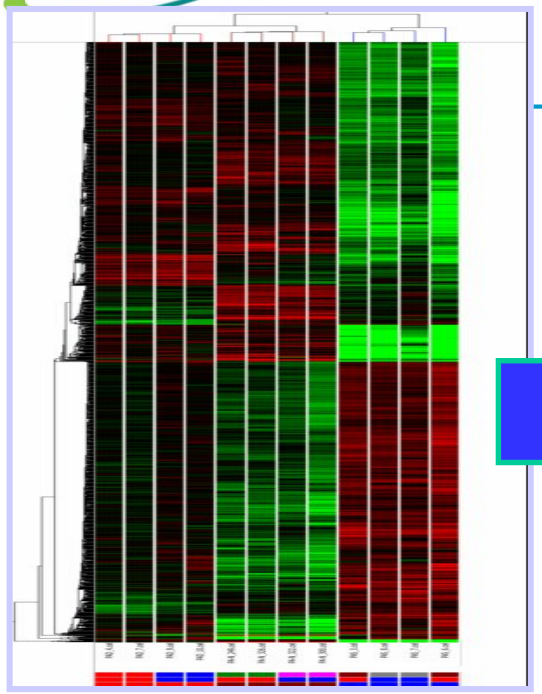
Data Analysis

Data Report

威健股份有限公司
Welgene Biotech. Co.,Ltd.

WELGENE

進階數據分析服務

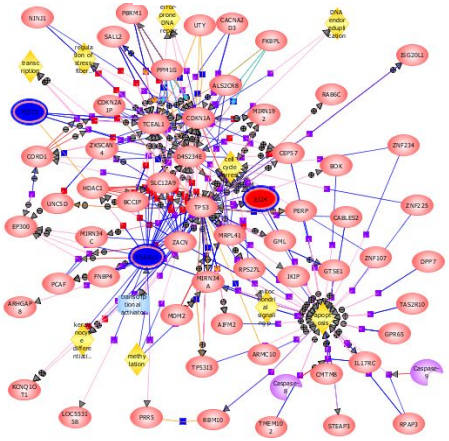


Find Significantly Expressed Genes

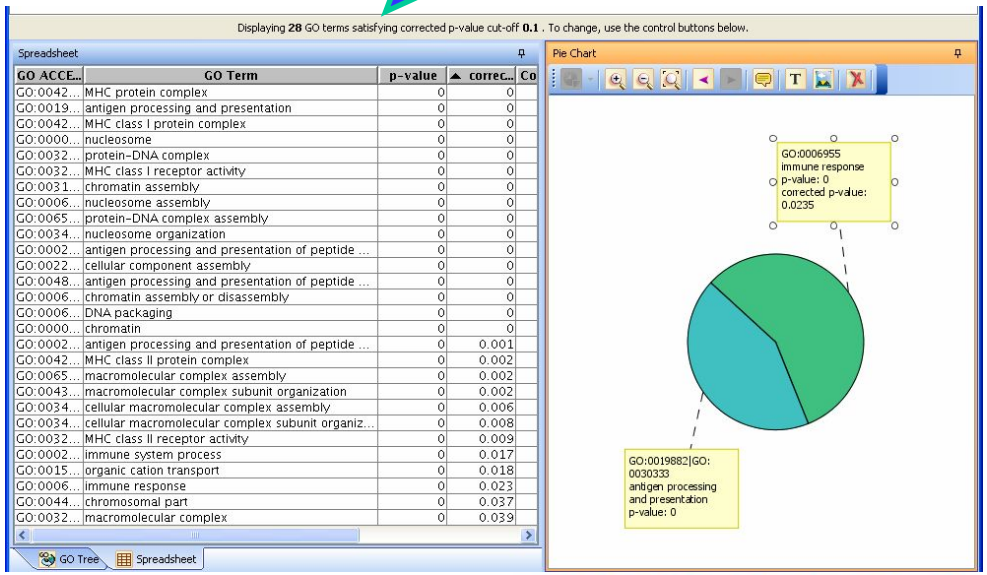
ProbeName	GeneSymbol	Common name	Description
A_23_P64873	DCN	NM_001920	Homo sapiens decorin (DCN), transcript variant A1, mRNA [NM_001920]
A_23_P421401	PDGFRB	NM_002609	Homo sapiens platelet-derived growth factor receptor, beta polypeptide (PDGFRB), mRNA [NM
A_23_P209625	CYP1B1	NM_000104	Homo sapiens cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1), mRNA [NM_
A_23_P211233	COL6A2	NM_001849	Homo sapiens collagen, type VI, alpha 2 (COL6A2), transcript variant 2C2, mRNA [NM_00184
A_23_P164057	MFAP4	NM_002404	Homo sapiens microfibrillar-associated protein 4 (MFAP4), mRNA [NM_002404]
A_23_P142533	COL3A1	NM_000090	Homo sapiens collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominan
A_23_P063	LUM	NM_002345	Homo sapiens lumican (LUM), mRNA [NM_002345]
P68	CPM	NM_001874	Homo sapiens carboxypeptidase M (CPM), transcript variant 1, mRNA [NM_001874]
	SPON2	NM_012445	Homo sapiens spondin 2, extracellular matrix protein (SPON2), mRNA [NM_012445]
	PRRX2	NM_016307	Homo sapiens paired related homeobox 2 (PRRX2), mRNA [NM_016307]
	COL3A1	NM_000090	Homo sapiens collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominan
16	GPM6B	NM_001001996	Homo sapiens glycoprotein M6B (GPM6B), transcript variant 2, mRNA [NM_001001996]
A_23_P00033	PDGFRA	NM_006206	Homo sapiens platelet-derived growth factor receptor, alpha polypeptide (PDGFRA), mRNA [N

Network Analysis

GO Analysis



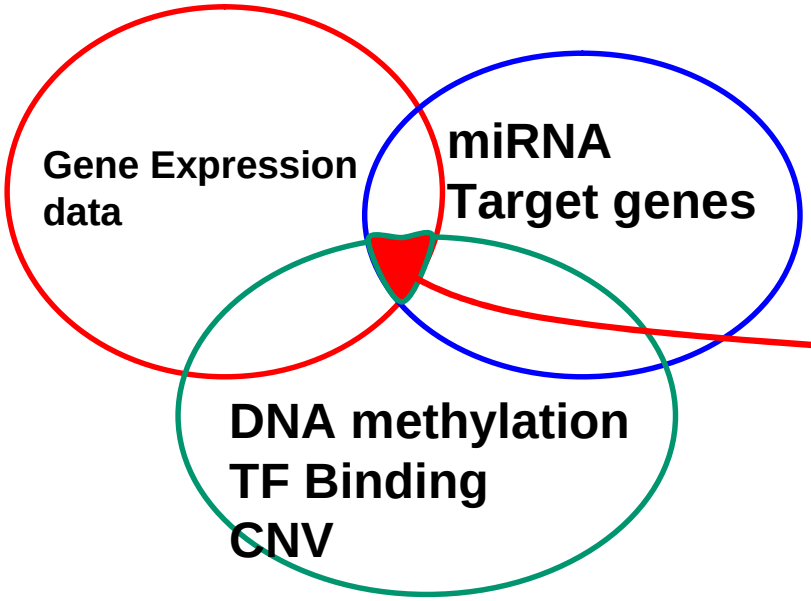
BioPax, NLP, MeSH

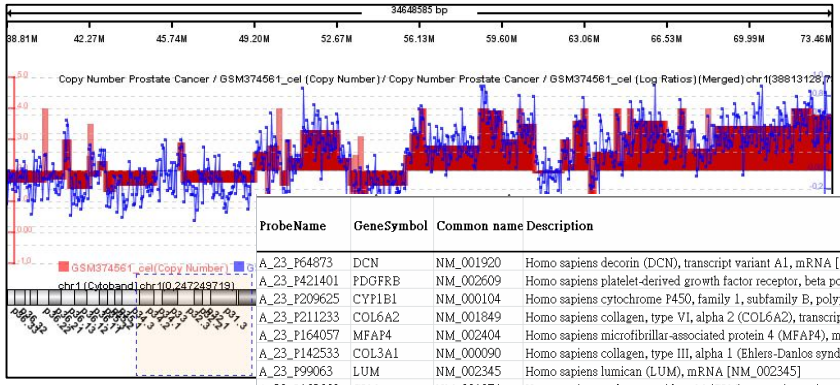


GO ACCE...	GO Term	p-value	correc...	Co
GO:0042...	MHC protein complex	0	0	
GO:0019...	antigen processing and presentation	0	0	
GO:0042...	MHC class I protein complex	0	0	
GO:0000...	nucleosome	0	0	
GO:0032...	protein-DNA complex	0	0	
GO:0032...	MHC class I receptor activity	0	0	
GO:0031...	chromatin assembly	0	0	
GO:0006...	nucleosome assembly	0	0	
GO:0065...	protein-DNA complex assembly	0	0	
GO:0034...	nucleosome organization	0	0	
GO:0002...	antigen processing and presentation of peptide ...	0	0	
GO:0022...	cellular component assembly	0	0	
GO:0048...	antigen processing and presentation of peptide ...	0	0	
GO:0006...	chromatin assembly or disassembly	0	0	
GO:0006...	DNA packaging	0	0	
GO:0000...	chromatin	0	0	
GO:0002...	antigen processing and presentation of peptide ...	0	0.001	
GO:0042...	MHC class II protein complex	0	0.002	
GO:0065...	macromolecular complex assembly	0	0.002	
GO:0043...	macromolecular complex subunit organization	0	0.002	
GO:0034...	cellular macromolecular complex assembly	0	0.006	
GO:0034...	cellular macromolecular complex subunit organiz...	0	0.008	
GO:0032...	MHC class II receptor activity	0	0.009	
GO:0002...	immune system process	0	0.017	
GO:0015...	organic cation transport	0	0.018	
GO:0006...	immune response	0	0.023	
GO:0044...	chromosomal part	0	0.037	
GO:0032...	macromolecular complex	0	0.039	

威健股份有限公司
Welgene Biotech. Co.,Ltd.

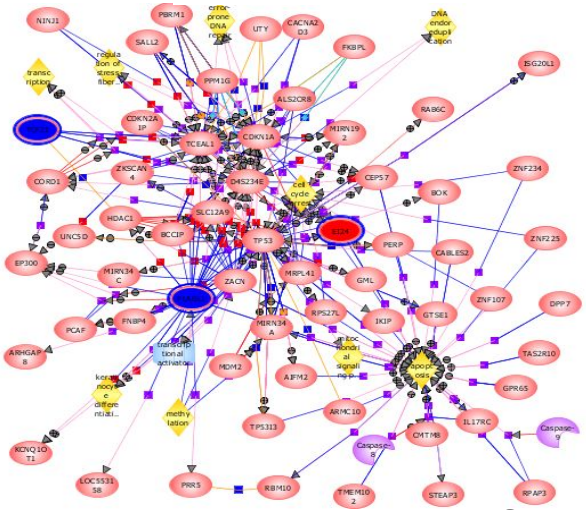
整合性數據分析





ProbeName	GeneSymbol	Common name	Description
A_23_P64873	DCN	NM_001920	Homo sapiens decorin (DCN), transcript variant A1, mRNA [NM_001920]
A_23_P421401	PDGFRB	NM_002609	Homo sapiens platelet-derived growth factor receptor, beta polypeptide (PDGFRB), mRNA [NM
A_23_P209625	CYP1B1	NM_000104	Homo sapiens cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1), mRNA [NM_
A_23_P211233	COL6A2	NM_001849	Homo sapiens collagen, type VI, alpha 2 (COL6A2), transcript variant 2C2, mRNA [NM_00184
A_23_P164057	MFAP4	NM_002404	Homo sapiens microfibrillar-associated protein 4 (MFAP4), mRNA [NM_002404]
A_23_P142533	COL3A1	NM_000090	Homo sapiens collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominan
A_23_P99063	LUM	NM_002345	Homo sapiens lumican (LUM), mRNA [NM_002345]
A_23_P162668	CPM	NM_001874	Homo sapiens carboxypeptidase M (CPM), transcript variant 1, mRNA [NM_001874]
A_23_P121533	SPON2	NM_012445	Homo sapiens spondin 2, extracellular matrix protein (SPON2), mRNA [NM_012445]
A_23_P83298	PRRX2	NM_016307	Homo sapiens paired related homeobox 2 (PRRX2), mRNA [NM_016307]
A_24_P935491	COL3A1	NM_000090	Homo sapiens collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominan
A_23_P378416	GPM6B	NM_001001996	Homo sapiens glycoprotein M6B (GPM6B), transcript variant 2, mRNA [NM_001001996]
A_23_P300033	PDGFRA	NM_006206	Homo sapiens platelet-derived growth factor receptor, alpha polypeptide (PDGFRA), mRNA [N

Find overlapped genes



Build possible gene interaction network

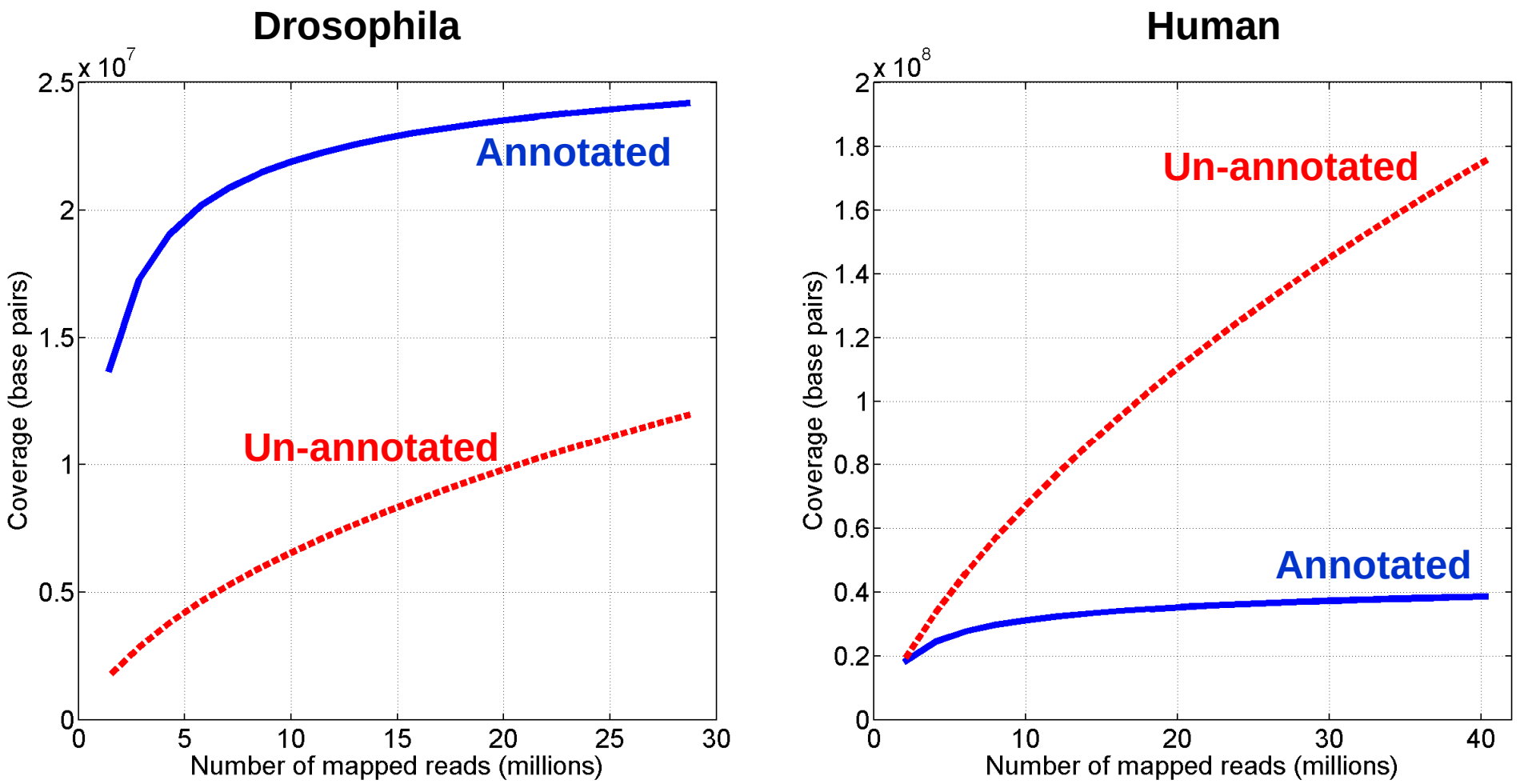
WELGENE

Thank You

Any Question ?

威健股份有限公司
Welgene Biotech. Co.,Ltd.

Coverage vs sequencing depth

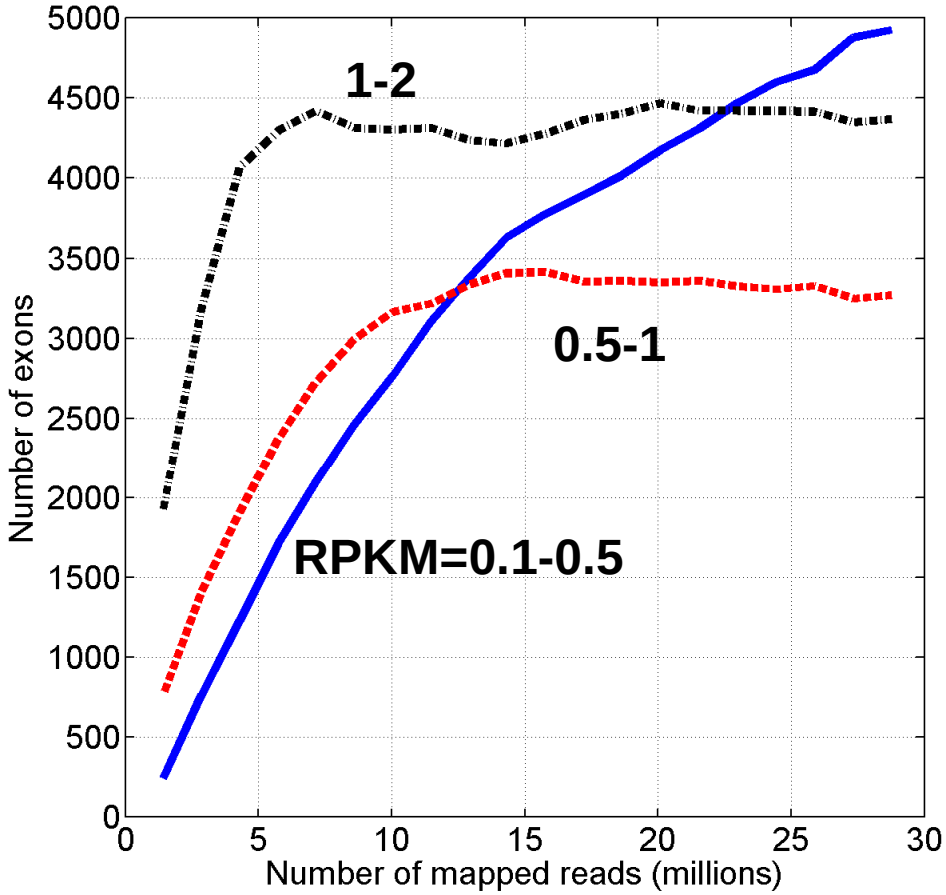


75-mer reads

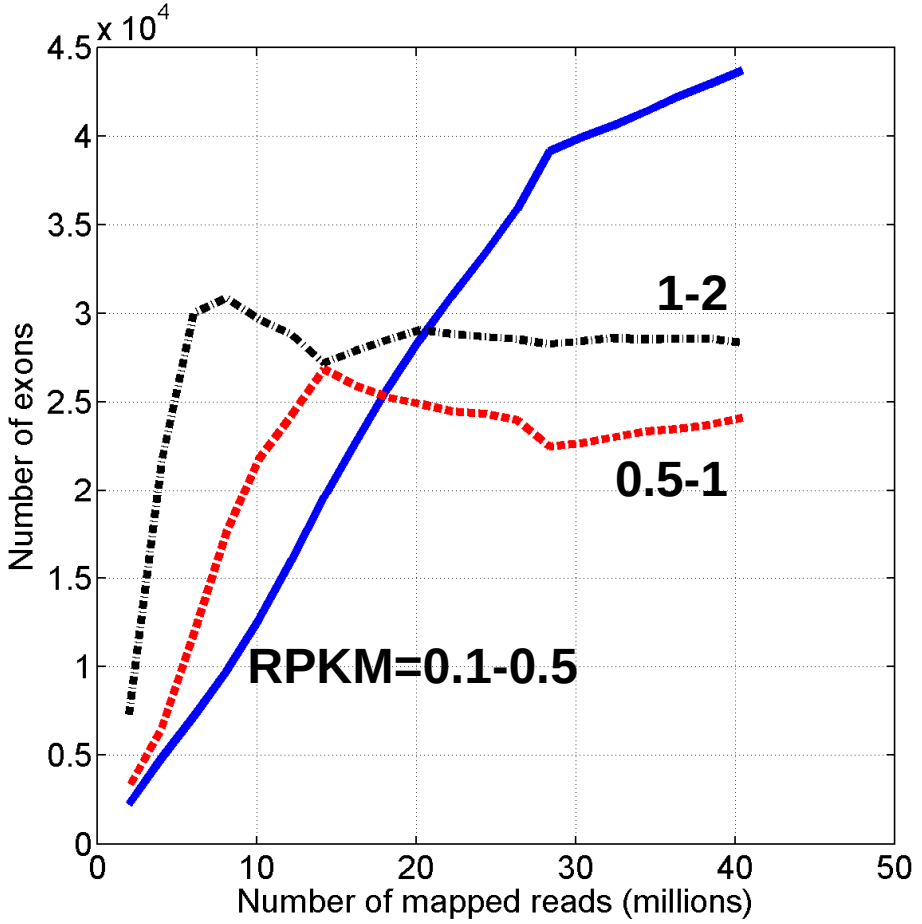
威健股份有限公司
Welgene Biotech. Co.,Ltd.

RPKM vs sequencing depth

Drosophila



Human



RPKM=reads per kilobase per million of mapped reads
Mortazavi et al, 2008: RPKM=1~3 for 1 copy/cell

威健股份有限公司
Welgene Biotech. Co.,Ltd.