

# MethylTarget项目总结报告



|  |  |
| --- | --- |
| **项目编号：** | **16B0330A** |
| **项目名称：** | **上海复旦王久存－王明华甲基化目的区域测序** |
| **客户单位：** | **复旦大学现代人类学中心** |
| **报告日期：** | **2016-08-31** |

上海天昊生物科技有限公司

**目录**

**1、项目概况**

**2、服务介绍与说明**

**3、建库测序**

**4、数据分析**

**5、附件**

**6、实验人员**

**7、权责声明**

**附录I: 上海天昊遗传分析中心联系方式**

**附录II:上海天昊生物科技有限公司简介**

1. **项目概况**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **备注** | |
| **项目编号** | 16B0330A |  | |
| **项目类型** | MethylTarget |  | |
| **样本数目** | 210 |  | |
| **样本类型** |  |  | |
| **目标区域信息** | 约50个基因 | 附件1目标基因信息表 | |
| **服务项目** | 16B0330A ---基因组DNA 抽提 | DNA 抽提报告 | |
| 16B0330A ---MethylTarget测序 | Illumina Hiseq，2×150bp | |
| 16B0330A ---MethylTarget测序标准生物信息分析 | 数据量以及数据质量信息汇总 | √ |
| Bisulfite处理不完全情况统计 | √ |
| 甲基化比例分析 | √ |
| 单倍型分析 | √ |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| **个性化分析项目** |  |  | |
| **完成日期** |  |  | |

1. **服务介绍与说明**
2. 背景介绍

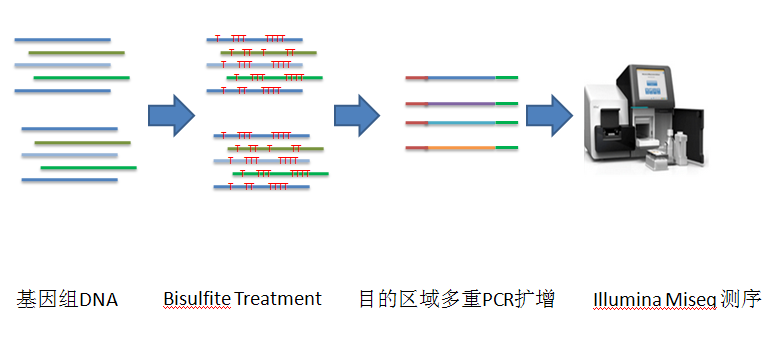
甲基化是一种非常重要的表观遗传学标记信息，参与了包括基因调控、生物发育、疾病等方方面面，并且在大多数癌症中广泛发现甲基化异常。获得检测范围内所有C位点的甲基化水平数据，对于表观遗传学的时空特异性研究具有重要意义。MethylTarget以二代测序高通量测序平台为基础，结合Bisulfite处理和生物信息数据分析进行低成本、高效率、高准确度的DNA甲基化水平图谱绘制，为癌症、自身免疫等疾病的诊疗提供了有力的帮助。

1. 样本要求
2. 样品类型：基因组DNA，溶解在H2O或TE (pH 8.0)中。
3. 样品纯度：OD 260/280值应在1.7～2.0 之间，260/230 值在1.8以上。
4. 样品浓度：最低浓度不低于20ng/µL。
5. 样品总量：每个样品总量1µg，满足约48个目的片段测序。
6. 数据量要求

Hiseq 2×150，所有样本目的区域的平均有效测序深度的平均值>1000X，90%样本的平均有效测序深度>500X，对于平均有效测序深度大于500X的样本， 90%以上的测序片段测序深度大于10X。

1. 服务流程

样品接收－样品DNA抽提与质检－建库测序－生物信息分析－项目总结报告

1. **建库测序**
2. 建库测序流程
3. 
4. 试剂耗材

3.1、 主要试剂：

|  |  |
| --- | --- |
| **试剂名称** | **试剂供应商** |
| MiSeq Reagent Kits v3 | Illumina, San Diego, CA, USA |
| EZ DNA Methylation-Gold Kit | ZYMO, CA,USA |
| TIANGEN Gel Extraction kit | TIANGEN, Beijing，China |
| 10× Reaction buffer | TaKaRa, Dalian, China |
| HotStarTaq polymerase | TaKaRa, Dalian, China |

3.2、主要器材：

|  |  |
| --- | --- |
| **器材名称** | **器材供应商** |
| AB 2720 Thermal Cycler | Applied Biosystems, Waltham, MA, USA |
| XiangYi H1650-W 离心机 | XiangYi,Hunan,China |
| EP600 Gel electrophoresis | Yubai,Shanghai,China |
| NanoDrop | NanoDrop technologies, Wilmington, DE, USA |
| Invitrogen Qbit Spectrophotometer | Invitrogen, Carlsbad, CA, USA |
| Illumina Miseq | Illumina, San Diego, CA, USA |

1. 实验步骤
2. 亚硫酸氢盐处理

使用EpiTect Bisulfite Kit进行样本处理。

1. 样本目标片段多重PCR反应

使用以下PCR条件：

A 20µl mixture was prepared for each reaction and included:

1x reaction buffer (TAKARA),

2 mM Mg2+,

0.2 mM dNTP,

0.1µM of each primer,

1 U HotStarTaq polymerase (TAKARA) and 2 µl template DNA.

The cycling program was:

95ºC for 2 min; 11 cycles of94ºC for 20 s, 63ºC-0.5 ºC per cycle for 40 s, 72 ºC for 1mins; 24 cycles of94ºC for 20 s, 65 ºC for 30 s, 72 ºC for 1 mins; 72ºC for 2 min.

1. 相同样本的多重PCR反应体系混合
2. 样本添加特异性标签序列

使用以下PCR条件：

A 20µl mixture was prepared for each reaction and included 1x reaction buffer (NEB Q5TM), 0.3 mM dNTP,0.25µM of F primer, 0.25µM of index primer, 1 U Q5TM DNA polymerase (NEB) and 1µl diluted template.

The cycling program was 98ºC for 30s; 11 cycles of98ºC for 10 s, 65ºC for 30 s, 72 ºC for 30 s; 72ºC for 5 min.

1. 定量后上机测序
2. **数据分析**
3. **数据分析流程**

Reads for Sample

Map to Genome（Blast）

Map to Primer（blat）

Filter

Haplotype Calling

Methylation Calling

Reads Recalibrate

Reads Quality

Methylation Report，Haplotype Report，Statistics Report

1. **引物设计与CpGI注释**

### 引物设计

|  |  |
| --- | --- |
| Primer Name | seq |
| chrM\_F | TGTGTGGAAAGTGGTTGTGTAGATATT |
| chrM\_R | AATCACAAATCTATCACCCTATTAACCA |
| ADHFE1\_F | GTTTTGAGTTYGATTGGTTTGAGG |
| ADHFE1\_R | CTACRCRTTACAATTACCTCAACAAATAC |

Primer Name：引物名称

seq：引物序列

* **结果目录：methylation\_info.xlsx**

### CpGI注释

对引物对应的产物进行位置注释。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Product | Chr | Gene | mRNA | Initialcode Position | start | end | strand | Distance to initiation codon |
| TRIM71\_ | 3 | TRIM71 | NM\_001039111 | 32859572 | 32859498 | 32859277 | - | 74 |
| USP44\_ | 12 | USP44 | NM\_032147 | 95928032 | 95943012 | 95942799 | - | -14980 |
| SCOC\_ | 4 | SCOC | NM\_001153552 | 141294690 | 141293869 | 141294152 | + | -821 |

Product：产物名称

Chr：染色体

Gene：基因

mRNA：距离该产物最近的mRNA

Initialcode Position：mRNA编码起始位置

start：产物起始位置

end：产物终止位置

strand：产物方向

Distance to initiation codon：产物距离编码起始位置的距离

* **结果目录：methylation\_info.xlsx**

1. **数据质控**

### QC

对下机fastq原始数据，使用fastqc进行质量检测。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample | R1 | | | | R2 | | | |
| Total Reads | Reads length | Q20 | Q30 | Total Reads | Reads length | Q20 | Q30 |
| CT | 229879 | 35-151 | 0.999669391 | 0.924612514 | 229879 | 35-151 | 0.988798455 | 0.831628813 |
| CU | 249256 | 35-151 | 0.999739224 | 0.92342411 | 249256 | 35-151 | 0.98856196 | 0.825889848 |
| ET | 154873 | 35-151 | 0.999612586 | 0.923130565 | 154873 | 35-151 | 0.988267807 | 0.831687899 |

* **结果目录：Data\_statistics.xlsx**

### 富集效率

测序数据与产物进行比对，看一下目标片段富集的效率。

采用方法：Blat。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sample | CT | CU | ET | EU |
| 有效reads数目 | 176425 | 187837 | 108077 | 178287 |
| 总reads数目 | 229879 | 249256 | 154873 | 240531 |
| 富集效率 | 0.77 | 0.75 | 0.7 | 0.74 |

* **结果目录：Data\_statistics.xlsx**

### 片段有效reads数

统计在每个产物片段上的有效reads数目。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sample | CT | CU | ET | EU | HE | N100 |
| ADHFE1\_ | 9439 | 7788 | 5223 | 9250 | 4647 | 3322 |
| APC\_ | 10115 | 10729 | 8206 | 11415 | 5103 | 3691 |
| CH25H\_ | 3168 | 2717 | 1118 | 2859 | 2481 | 1139 |
| CHST2\_ | 2092 | 1584 | 1679 | 693 | 1123 | 1001 |
| CNR1\_ | 299 | 377 | 293 | 596 | 377 | 703 |
| DPY19L2P4\_ | 8852 | 8439 | 4598 | 9869 | 6822 | 4444 |
| ELMO1\_ | 6323 | 6039 | 2640 | 5905 | 4974 | 2893 |
| EOMES\_ | 100 | 119 | 57 | 84 | 61 | 50 |

* **结果目录：Data\_statistics.xlsx**

### 盐化效率

统计每个样本的有效reads中，盐处理后碱基C转变为T的效率。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Sample | CT | CU | ET | EU | HE | N100 | N101 |
| Transferred(C->T) | 5627946 | 6008981 | 3562479 | 5545491 | 4368268 | 3078478 | 1665907 |
| All Base(C) | 5726152 | 6114043 | 3618245 | 5639650 | 4428568 | 3121771 | 1691984 |
| Efficiency | 98.28% | 98.28% | 98.46% | 98.33% | 98.64% | 98.61% | 98.46% |

Transferred(C->T)：盐处理后碱基C转变为T的数量

All Base(C)：参照序列中碱基C总数（非CpG位点）

Efficiency：盐处理转化效率

* **结果目录：methylation.xlsx**

1. **甲基化统计**

### 甲基化比例分析

分析每个甲基化位点在样本中甲基化的比例。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Target | POS. | Distance to initiation codon | POS.In Genome | Type | CT | CU | ET |
| TRIM71\_ | 36 | 109 | 32859463 | CG | 0.871351351 | 0.949213162 | 0.918478261 |
| TRIM71\_ | 53 | 126 | 32859446 | CG | 0.721081081 | 0.756795422 | 0.745923913 |
| TRIM71\_ | 60 | 133 | 32859439 | CG | 0.711351351 | 0.928469242 | 0.714673913 |
| TRIM71\_ | 74 | 147 | 32859425 | CG | 0.885405405 | 0.956366237 | 0.903532609 |
| TRIM71\_ | 81 | 154 | 32859418 | CG | 0.943965517 | 0.966380544 | 0.964673913 |
| TRIM71\_ | 91 | 164 | 32859408 | CG | 0.884698276 | 0.962088698 | 0.918478261 |
| TRIM71\_ | 96 | 169 | 32859403 | CG | 0.945043103 | 0.961373391 | 0.938858696 |
| TRIM71\_ | 100 | 173 | 32859399 | CG | 0.954741379 | 0.972818312 | 0.944293478 |

Target：产物名称

POS：该甲基化位点在该产物上的位置

Distance to initiation codon：该甲基化位点距离编码起始位置的距离

POS.In Genome：该甲基化位点在参照基因组上的位置

Type：甲基化类型（CG表示CG位点的甲基化）

* **结果目录：methylation.xlsx**

### 单倍型分析

对每个片段计算每种单倍型。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Target | Haplotype | Depth | CT | CU | ET | EU |
| USP44\_ | ttttttttttttttt | 156924 | 0.008332639 | 0.006387736 | 0.005984043 | 0.008485363 |
| USP44\_ | ccccccccccccccc | 60728 | 0.175652029 | 0.277467263 | 0.170046543 | 0.171064913 |
| USP44\_ | tttttttttcttttt | 23913 | 0.001666528 | 0.000798467 | 0.000664894 | 0.001103097 |
| USP44\_ | tttttcttttttttt | 15172 | 0.00091659 | 0.00047908 | 0.000166223 | 0.00093339 |
| USP44\_ | cccccccccctcccc | 12047 | 0.095492042 | 0.146997764 | 0.078956117 | 0.108697497 |

Target：产物名称

Haplotype：单倍型类型

Depth：支持此单倍型的测序数量

* **结果目录：methylation.xlsx**

**参考文献：**

[1]. Feng H, Conneely KN, Wu H.A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. Nucleic 2014 Apr;42(8): e69. doi: 10.1093/nar/gku154. Epub 2014 Feb 22.

[2]. [Paul DS](http://www.ncbi.nlm.nih.gov/pubmed?term=Paul%20DS%5BAuthor%5D&cauthor=true&cauthor_uid=24518816), Guilhamon P, Karpathakis A, Butcher LM, Thirlwell C, Feber A, Beck S. Assessment of RainDrop BS-seq as a method for large-scale, targeted bisulfite sequencing. Epigenetics. 2014 May 1;9(5):678-84. doi: 10.4161/epi.28041. Epub 2014 Feb 11.

[3]. Holmes EE, Jung M, Meller S, Leisse A, Sailer V, Zech J, Mengdehl M, Garbe LA, Uhl B, Kristiansen G, Dietrich D. Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine. PLoS One. 2014 Apr 3;9(4): e93933. doi: 10.1371/journal.pone.0093933. eCollection 2014.

1. **权责声明**

**第1条**

本报告及其附件中披露的由上海天昊遗传分析中心负责设计优化的检测体系（合同中对权利归属已做说明的除外）的知识产权归上海天昊生物科技有限公司（以下简称公司）所有，任何其它单位及个人未经公司许可不得对其进行专利申请或用于商业用途。上述检测体系包括设计合成的各种核酸序列、检测反应组成、反应条件及检测流程等。

**第2条**

本报告及其附件中披露的项目内容、实验数据及结果，公司必须严格保密，未经客户单位负责人允许，严禁公司任何人对外宣传。

**第3条**

为了保证项目成果的安全，本报告及其附件中披露的实验数据及结果已通过邮件发送，而本报告及其附件中披露的实验及数据分析相关信息、原始实验数据则需客户单位的项目负责人或其联系人通过公司授权的账号在公司项目管理系统中下载。

**第4条**

本报告及其附件内容仅限于项目团队成员、客户单位的项目负责人及其联系人查阅或拷贝，公司内任何其他人员未经项目总监允许不得查阅或拷贝，公司外任何其他人员未经客户单位的项目负责人许可不可查阅或拷贝。

# 附录I：上海天昊生物科技有限公司联系方式

**业务联系**

|  |  |  |  |
| --- | --- | --- | --- |
| 总监 | 王颖 | Tel: 021-50802060\*125 | Email: |

**实验技术和流程以及实验报告咨询**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 刘燕 | Tel: 021-50802060\*133 | Email: |

**数据分析咨询**

|  |  |  |
| --- | --- | --- |
|  | 李才华 (生物信息)  Tel: 021-50802060\*117  Email: | 马瑞晓 (遗传统计)  Tel:  Email: |

**公司网址**

<http://biotech.geneskies.com/>

**论文引用或致谢**

如果您认可我们的工作，在发表文章时愿意将我们放到致谢名单中，我们将非常感谢，文章中天昊的中英文名称如下：

**上海天昊生物科技有限公司 (中文)**

**Genesky Biotechnologies Inc., Shanghai, 201315 (English)**

# 附录II：上海天昊生物科技有限公司简介

上海天昊生物科技有限公司，2008年4月创建于上海浦东张江高科技园区，随着公司的不断发展和壮大，目前上海天昊已经拥有员工近100名，2500平方的办公与实验场地，建立了整套完备的实验室管理体系和标准流程，主要目标是构建一个具备国际一流水准的标准化多层次分子生物学研究分析平台，为国内外基因生物领域科研机构、医学院校以及生物制药企业提供精确、高效的基因检测分析服务。

2013年我们在北京和南京分别成立的两个办事处，力求为全国客户提供更全面、更优质的服务。

2014年被评为“国家高新技术企业”，公司总经理姜正文博士获得：国家“千人计划”，江苏省“双创人才”、苏州工业园区“领军人才”等荣誉。

自上海天昊成立以来，我们已经建立并完善了包括[ABI 3130xl](http://biotech.geneskies.com/ch/technology/partform/173.html" \t "_blank)、ABI 3730xl测序仪、实时定量PCR仪、[Illumina GAIIx](http://biotech.geneskies.com/ch/technology/partform/175.html" \t "_blank)、Illumina Miseq二代测序平台和细胞生物学分析平台在内的多个分子生物学分析研究平台。不仅先后开展了疾病基因定位、基因突变及多态性测序分析、各种通量SNP基因分型分析、基因转录水平分析、甲基化水平分析等服务项目，同时还利用目前世界先进的二代测序技术开展了全基因组测序、全外显子组测序、RNA测序、micro-RNA测序、甲基化测序等成熟的科研分析服务。此外，上海天昊在进行高质量技术服务的同时还积极集中优势力量自主创新，独立研发了[AccuCopy](http://biotech.geneskies.com/ch/technology/service/146.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/146.html" \t "_blank)多重拷贝数检测技术、[CNVplex](http://biotech.geneskies.com/ch/technology/service/168.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/168.html" \t "_blank)高通量拷贝数检测技术、[SNPscan](http://biotech.geneskies.com/ch/technology/service/137.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/137.html" \t "_blank)高通量SNP分型方法等具有国际水平的专利技术，另外利用二代测序技术不断开发出大量新的基因分析技术，例如多种策略方法的目的区域富集大样品测序技术EasyTargetTM、超高通量SNP分型技术SNPseq®、超高通量目的区域CNV检测技术CNVseq®等。截至2014年 10月底，公司申请了3项PCT专利，18项国内发明专利，5项进入实审，2项专利已经授权，获得4项软件著作权，申请了38商标著作权，已获14商标著作权。目前公司可以提供的分子生物学或基因组学相关科研服务项目超过40类，迄今已经为国内外近500多家科研院校、医疗单位和生物公司提供了超过2000项科研技术服务。

在为广大新老客户提供科研技术实验服务的同时，上海天昊集服务与产品于一身，利用自主研发的[AccuCopy](http://biotech.geneskies.com/ch/technology/service/146.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/146.html" \t "_blank)、[CNVplex](http://biotech.geneskies.com/ch/technology/service/168.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/168.html" \t "_blank)、[SNPscan](http://biotech.geneskies.com/ch/technology/service/137.html" \t "_blank)[®](http://biotech.geneskies.com/ch/technology/service/137.html" \t "_blank)、EasyTargetTM、SNPseq®和CNVseq®等全新技术为客户提供灵活定制科研项目检测试剂盒服务，为具备检测实力的大中型科研单位提供更为先进高效的检测方法和试剂。

[上海天昊遗传分析中心](http://biotech.geneskies.com/ch/about/165.html)借助多名长期从事基因及遗传分析的领域专家构成的专家咨询团队，结合准确、高效、经济的科研技术服务体系，致力于长期为分子生物学及医学遗传学领域的研究者提供高质量的科研策略咨询、实验技术服务和遗传数据分析，帮助广大科研人员获得更为优质的科研成果。

上海天昊生物科技有限公司  
地址：上海浦东新区康桥路787号9号楼  
服务热线：400-065-6886  
传真：021-50802059  
网址：[www.geneskies.com](http://www.geneskies.com/)