

# 数据分析报告

## 基础信息

0	1	2
FileName	20240117_Sync_Y0004_03_H01_Run0002.hdf5	20240117_Sync_Y0004_04_H01_Run0002.hdf5
NP	DPN2674-E. Coli-PN-L617-N042:N101-6Plinker-十烷-60min	DPN2675-E. Coli-PN-L844-N042:N101-6Plinker-交联磷脂-60min
CYCLE	1. 0. 1. 0	1. 0. 1. 0
Voltage	-80. 0. 80. 0	-80. 0. 80. 0
Aperture	W49-C023	W49-C023
totalavarage_identity	86. 51317045	87. 21041874
total_sub_len	338272773	429476992
subs_more_than_80_pr	0. 904232606	0. 927678657
subs_more_than_83_pr	0. 869081945	0. 886902306
subs_more_than_85_pr	0. 814283304	0. 832003851
subs_more_than_88_pr	0. 528438702	0. 588008402
subs_more_than_90_pr	0. 174748164	0. 215196943
valid_reads_num	53391	73168
invalid_reads_num	2724	2792
total_matched_reads_num	56115	75960
total_reads_len	439963453	524216687
total_valid_reads_len	426102834	507649072
total_valid_reads_coverage	0. 793875905	0. 846011577
global_min_len	104	106
global_max_len	66702	69919
global_median_len	8651	7595
global_mean_len	7476. 055769	6805. 067529
average_linker_speed	2. 253074993	1. 963799447
average_basecall_lifetime	3534. 795024	3528. 997273
single_pore	0. 0439	0. 029
multi_pore	0. 061	0. 0769
suspected	0. 0033	0. 0025
appearance	0. 116	0. 1524

上孔信息

20240117\_Sync\_Y0004\_04\_H01\_Run0002 - 上孔信息

Key	Rate	Num
single_pore	0.029	15204.352
multi_pore	0.0769	40317.7472
suspected	0.0025	1310.72
appearance	0.1524	79901.4912

20240117\_Sync\_Y0004\_04\_H01\_Run0002 - 上孔信息

Key	Rate	Num
single_pore	0.029	15204.352
multi_pore	0.0769	40317.7472
suspected	0.0025	1310.72
appearance	0.1524	79901.4912

单孔信号良率

# 目录

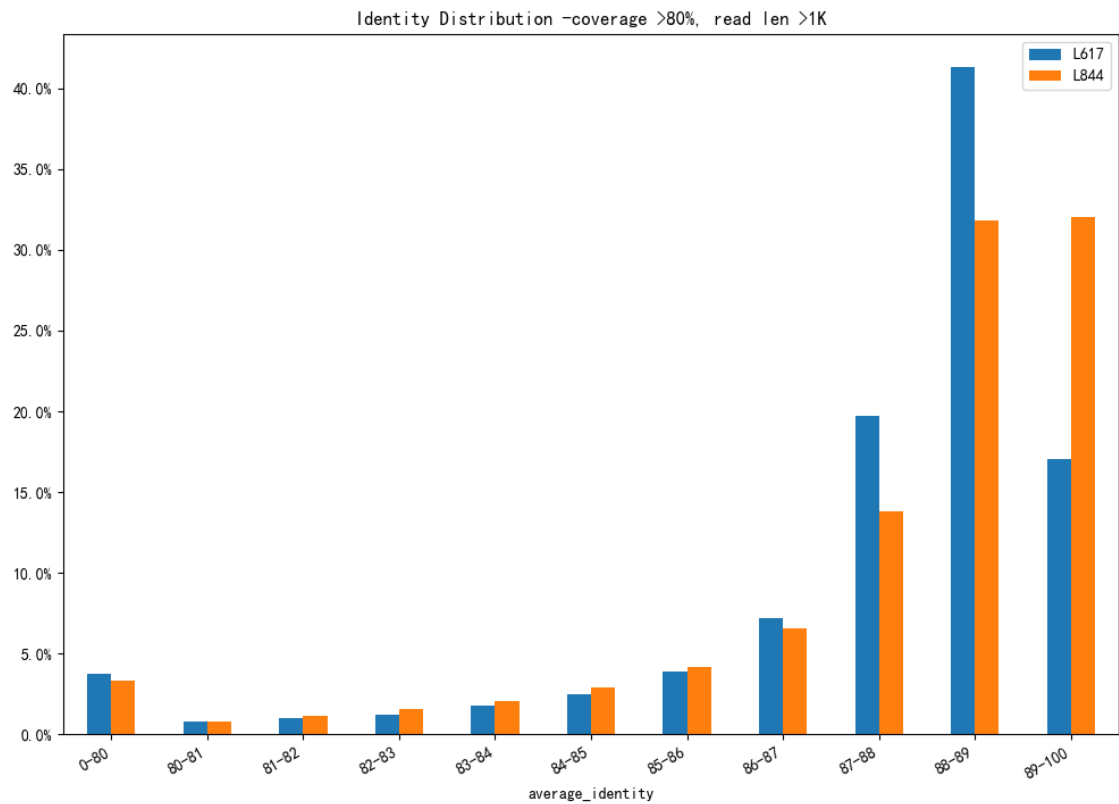
---

- 1. 准确率相关统计
  - 1.1 准确率分布
  - 1.2 与读长关系
  - 1.3 错误类型统计
  - 1.4 准确率差异分析
- 2. 读长相关统计
- 3. 速度相关统计
- 4. AR相关统计
  - 4.1 与Block Rate的关系
  - 4.2 AR分布
- 5. DW相关统计
  - 5.1 与Block Rate的关系
  - 5.2 DW分布
- 6. Capture Rate相关统计
  - 6.1 与Block Rate的关系
  - 6.2 CR分布
- 7. Pore-LT与Basecall-LT相关统计

# 1. 准确率相关统计

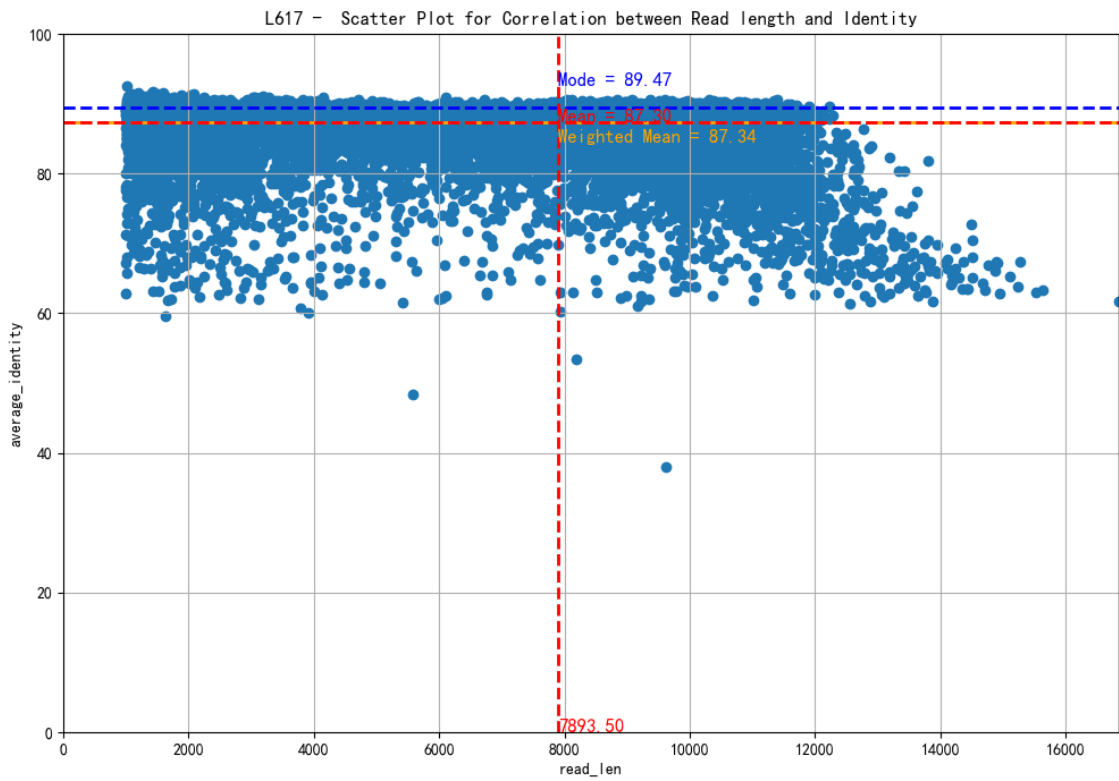
## 1.1 准确率分布

Fig 1-1-1 准确率分布直方图

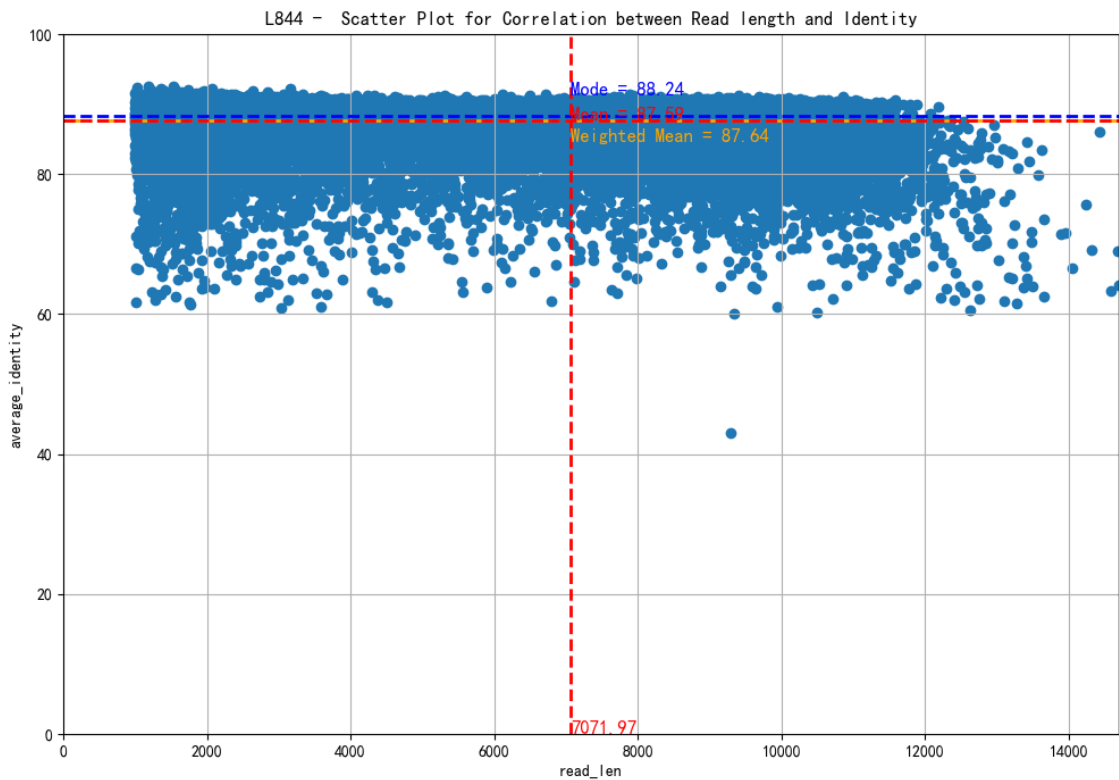


## 1.2 与读长关系

L617 - Read len与Identity相关性散点图



L844 - Read len与Identity相关性散点图



1.3 错误类型统计

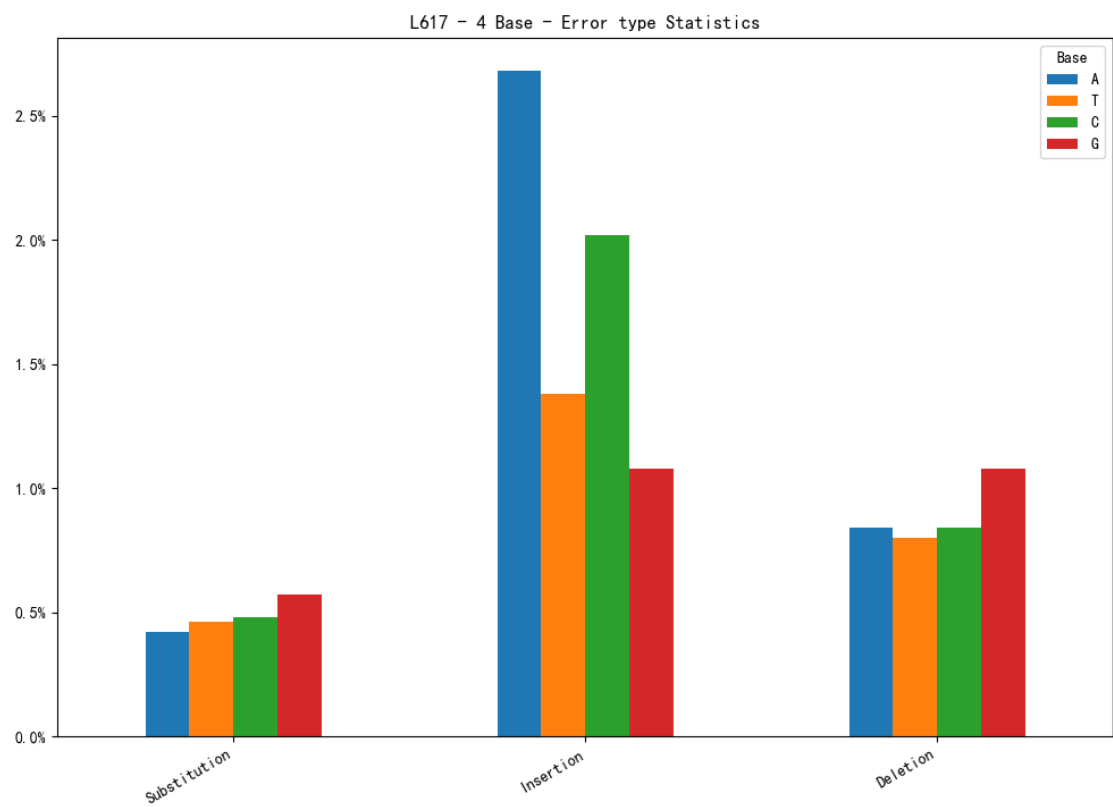
L617 - 错误类型统计汇总表

Base	Substitution	Insertion	Deletion	IDS_count
X	6240098 (1.93%)	23135472 (7.16%)	11516598 (3.56%)	40892168 (12.65%)
A	1365227 (0.42%)	8658818 (2.68%)	2706088 (0.84%)	12730133 (3.94%)
T	1496098 (0.46%)	4473405 (1.38%)	2597179 (0.8%)	8566682 (2.64%)
C	1536808 (0.48%)	6525365 (2.02%)	2715844 (0.84%)	10778017 (3.34%)
G	1841963 (0.57%)	3477888 (1.08%)	3497491 (1.08%)	8817342 (2.73%)

L844 - 错误类型统计汇总表

Base	Substitution	Insertion	Deletion	IDS_count
X	7855724 (1.83%)	32286509 (7.52%)	12896508 (3.0%)	53038741 (12.35%)
A	1735989 (0.4%)	11821533 (2.75%)	3020774 (0.7%)	16578296 (3.85%)
T	1851350 (0.43%)	5972314 (1.39%)	2797345 (0.65%)	10621009 (2.47%)
C	1938989 (0.45%)	9798747 (2.28%)	3233150 (0.75%)	14970886 (3.48%)
G	2329394 (0.54%)	4693921 (1.09%)	3845242 (0.9%)	10868557 (2.53%)

L617 - 错误类型统计汇总图



L844 - 错误类型统计汇总图

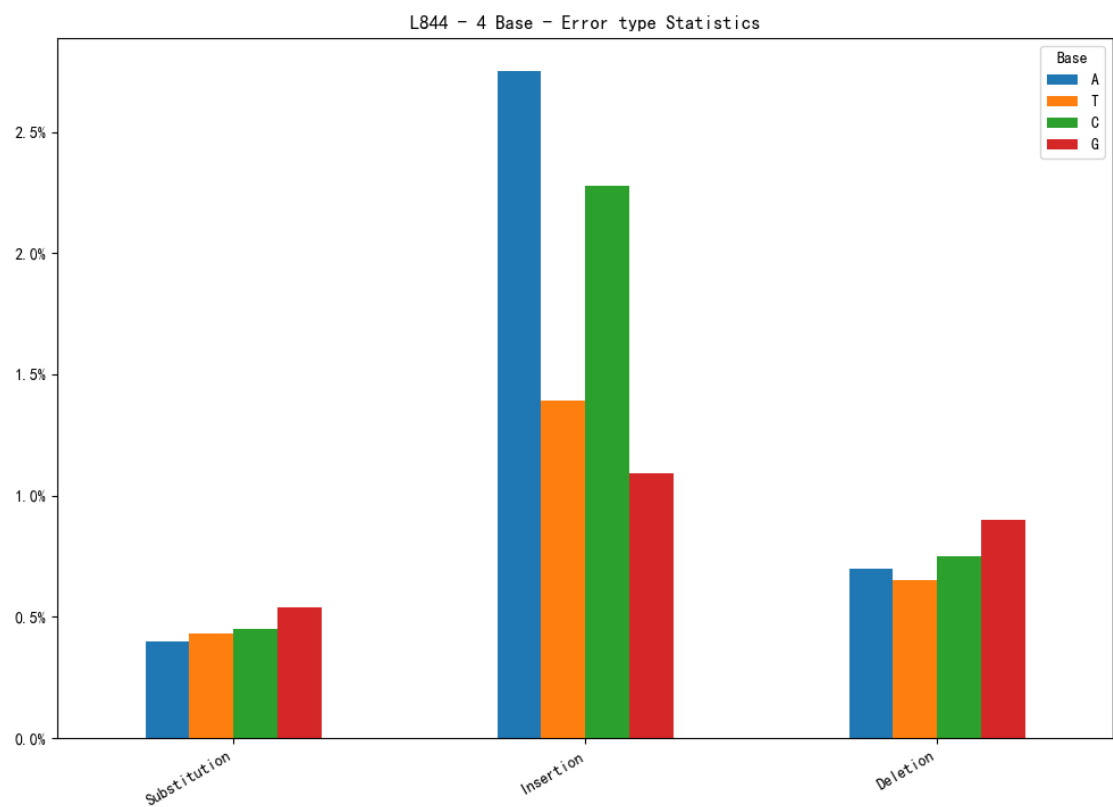


Fig 1-3-1 总-错误类型统计

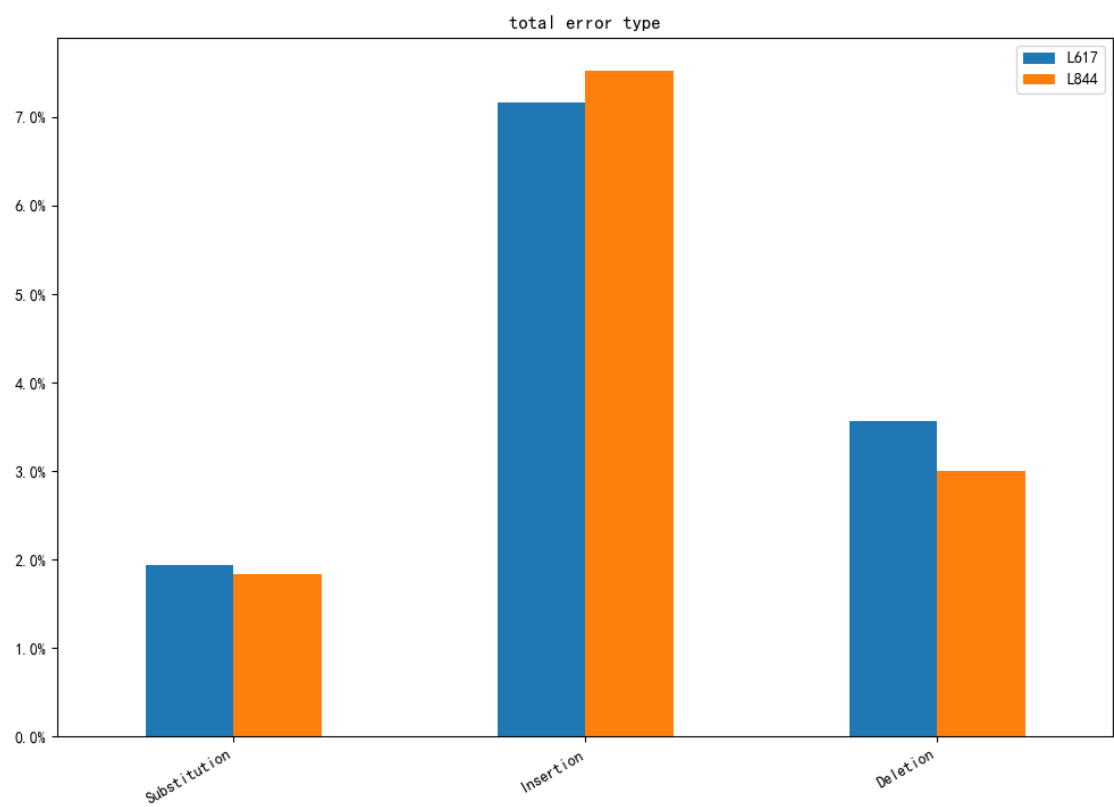
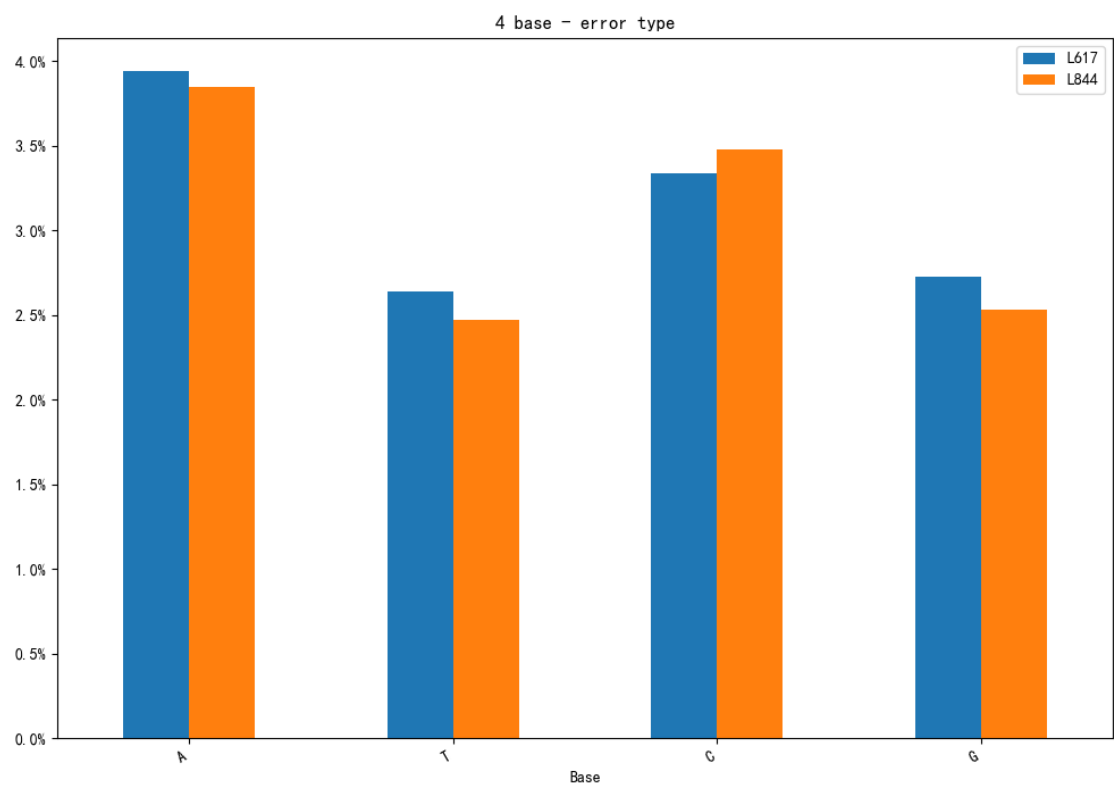


Fig 1-3-2 四种碱基-总错误统计






1.4 准确率差异分析


1.4.1 单因素方差分析

Table1-4-1 单因素方差分析统计表

Fig 1-4-1

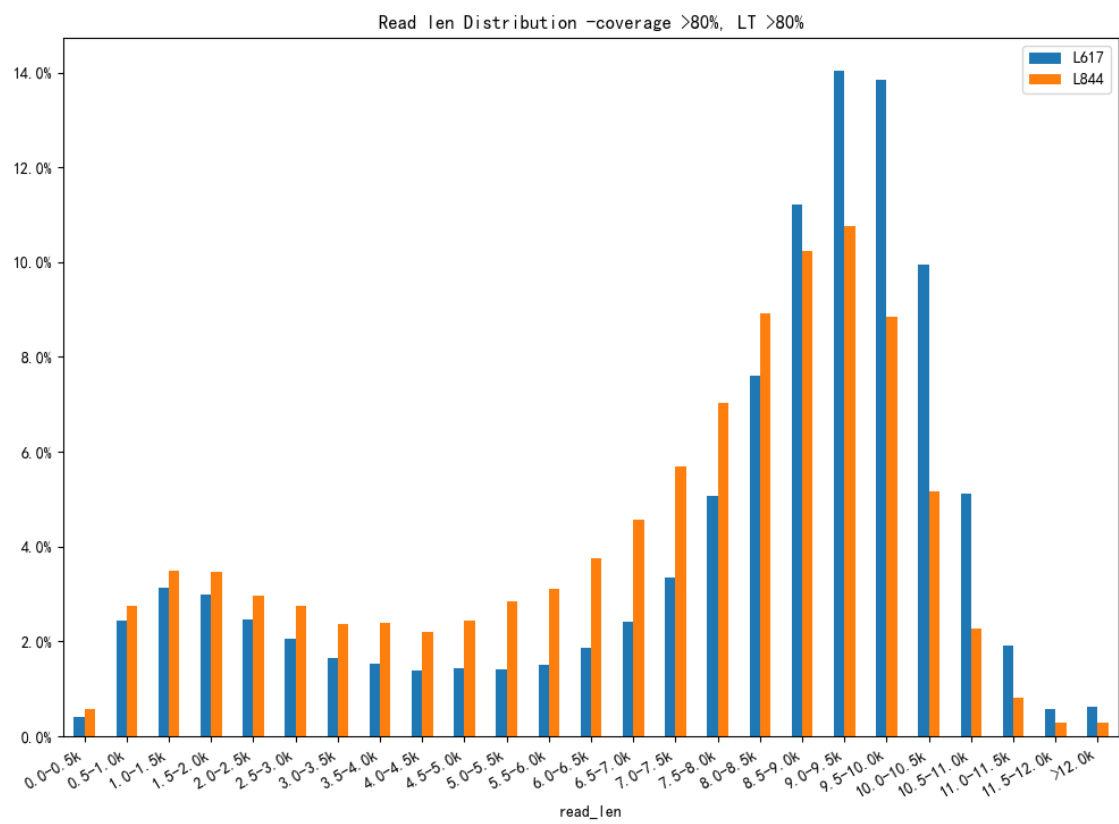
1.4.2 秩和检验

Table1-4-2秩和检验统计表

Fig 1-4-2

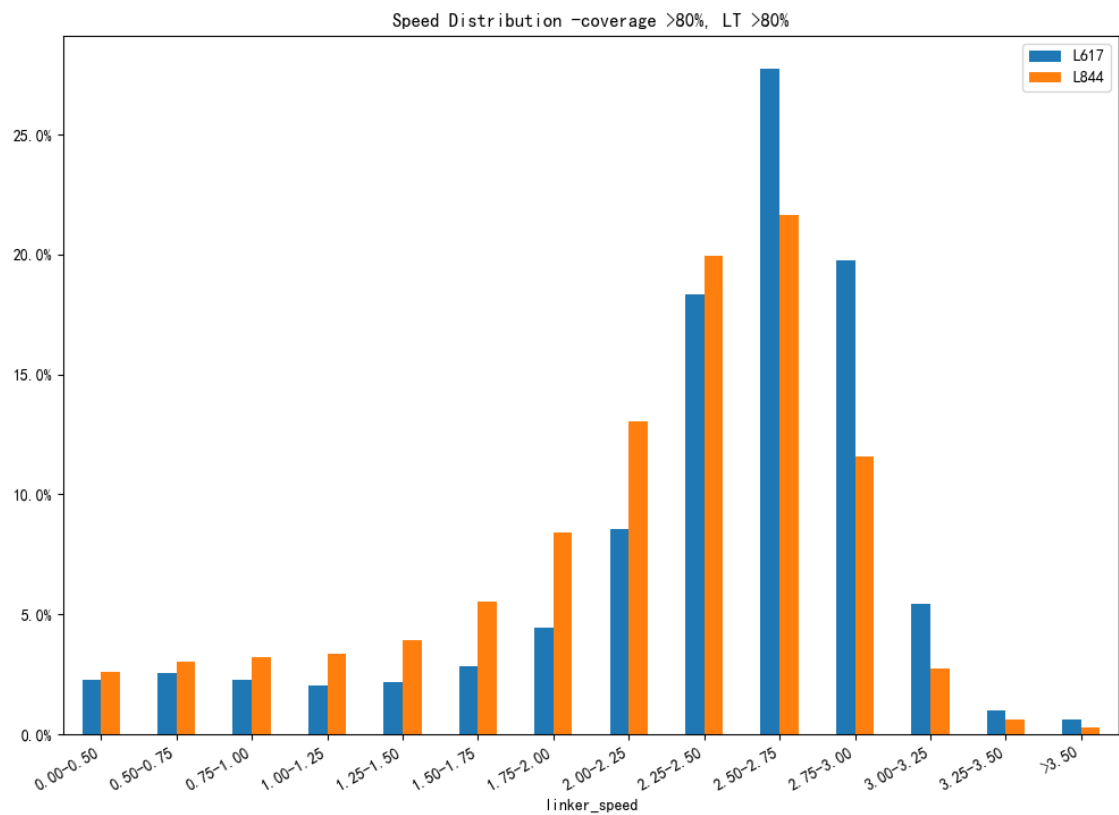
2. 读长相关统计

Fig 2-1 信号LT>60min 读长分布占比图



3. 速度相关统计

Fig 3-1 信号LT>60min 速度分布占比图



4. AR相关统计

4.1 与Block Rate的关系

L617 - AR与BR相关性散点图

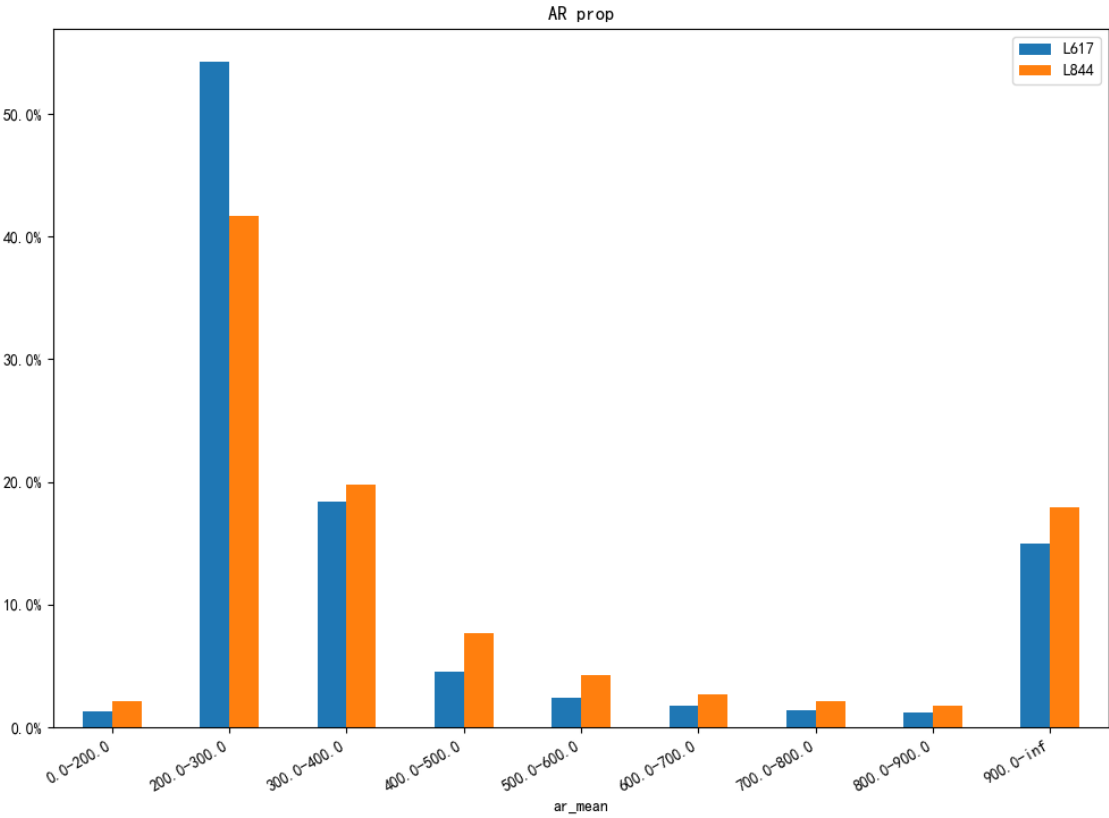


L844 - AR与BR相关性散点图

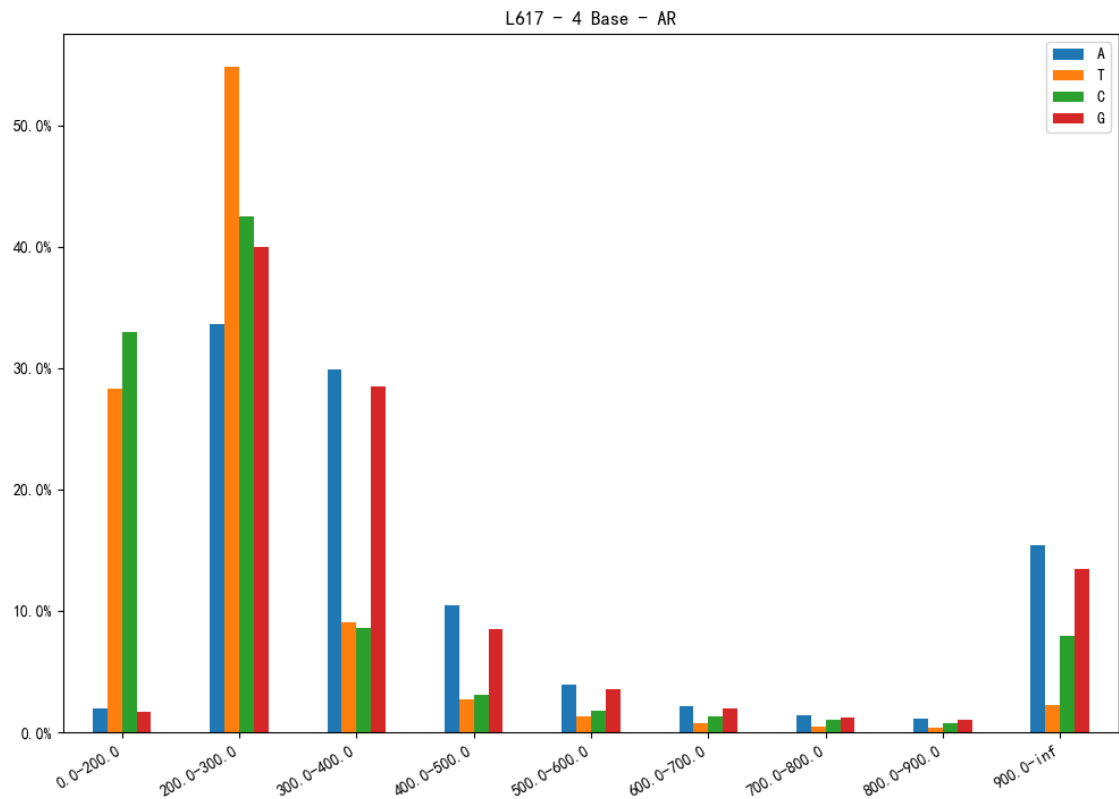


4.2 AR分布

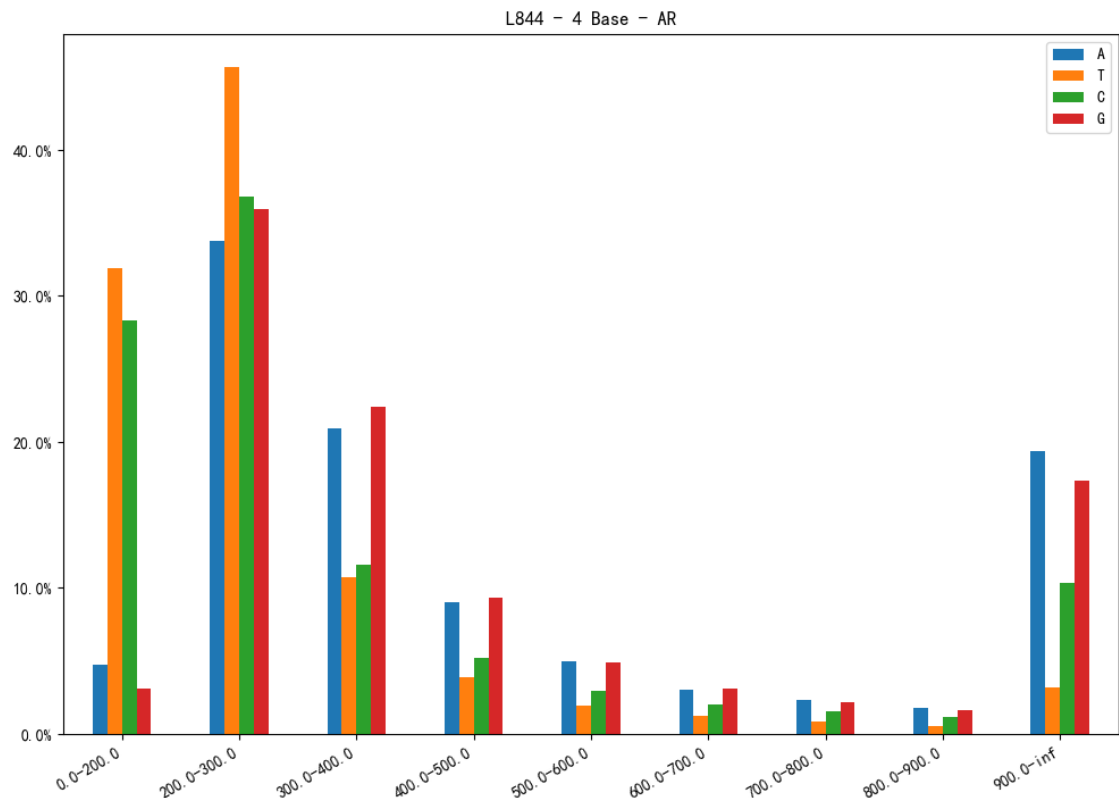
Fig 4-2-1 AR分布条形图



L617 四种碱基AR占比图



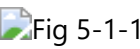
L844 四种碱基AR占比图



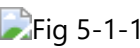
5. DW相关统计

5.1 与Block Rate的关系

L617 - DW与BR相关性散点图

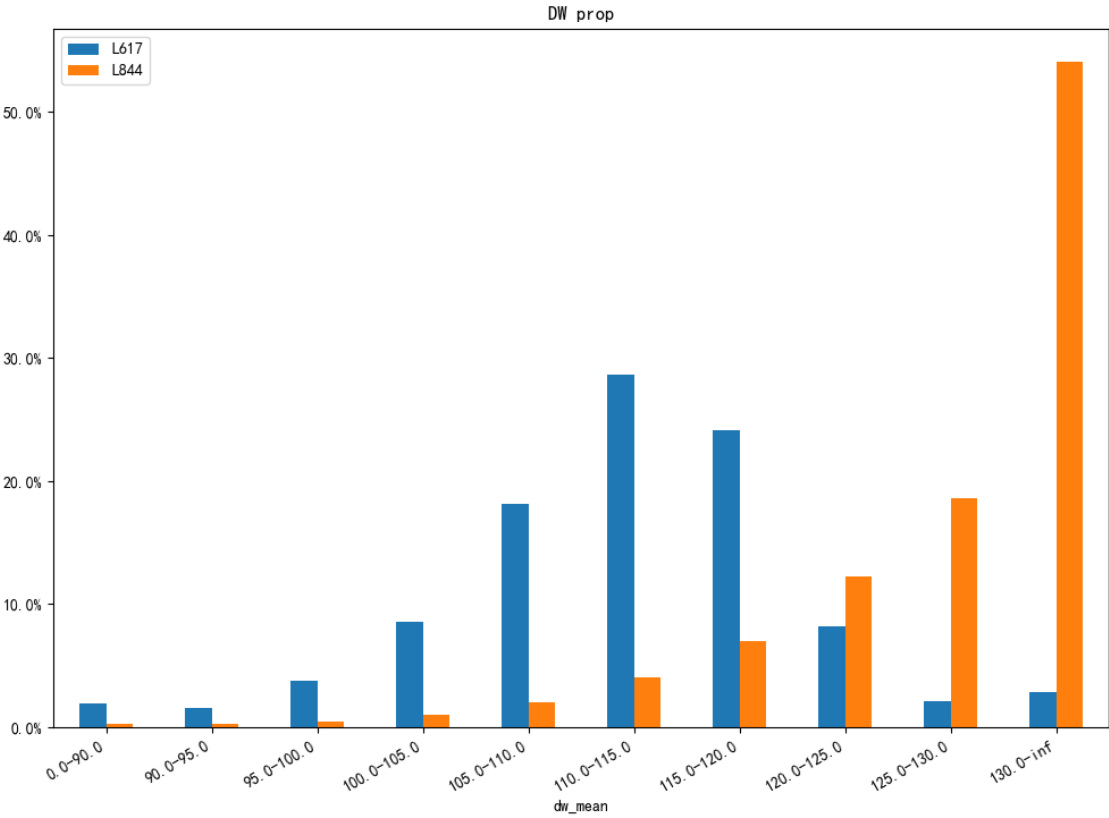


L844 - DW与BR相关性散点图

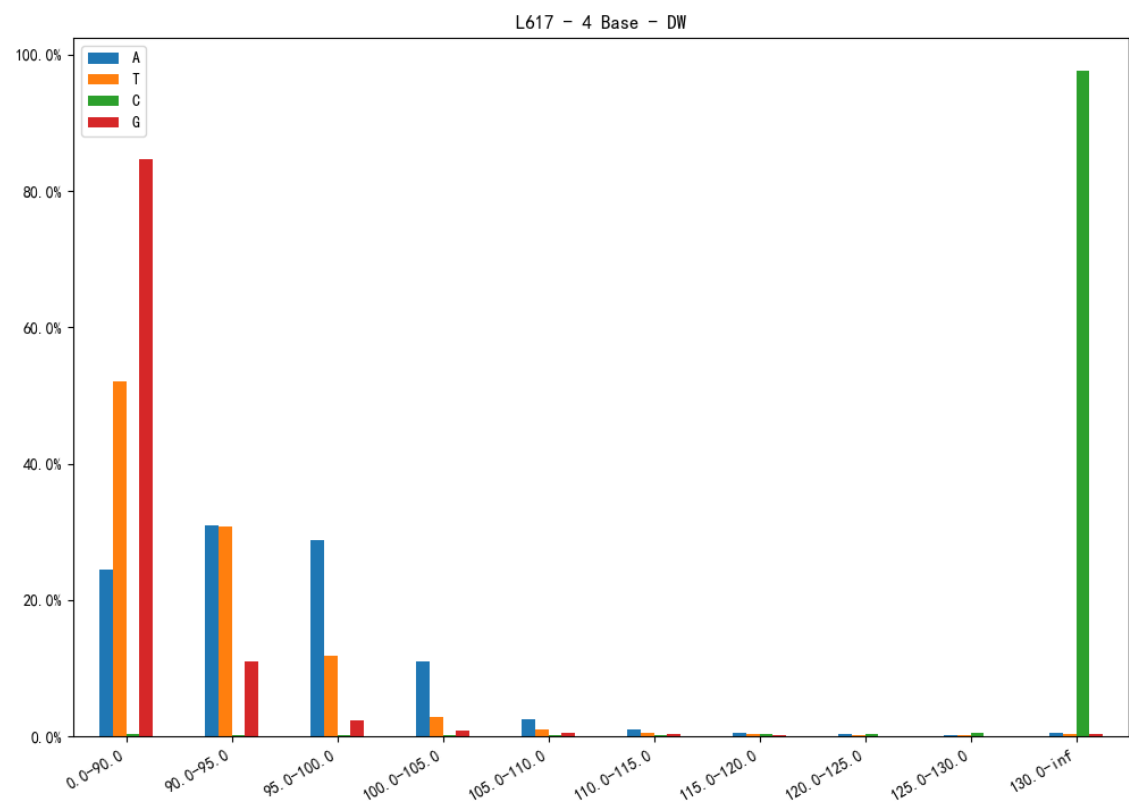


5.2 DW分布

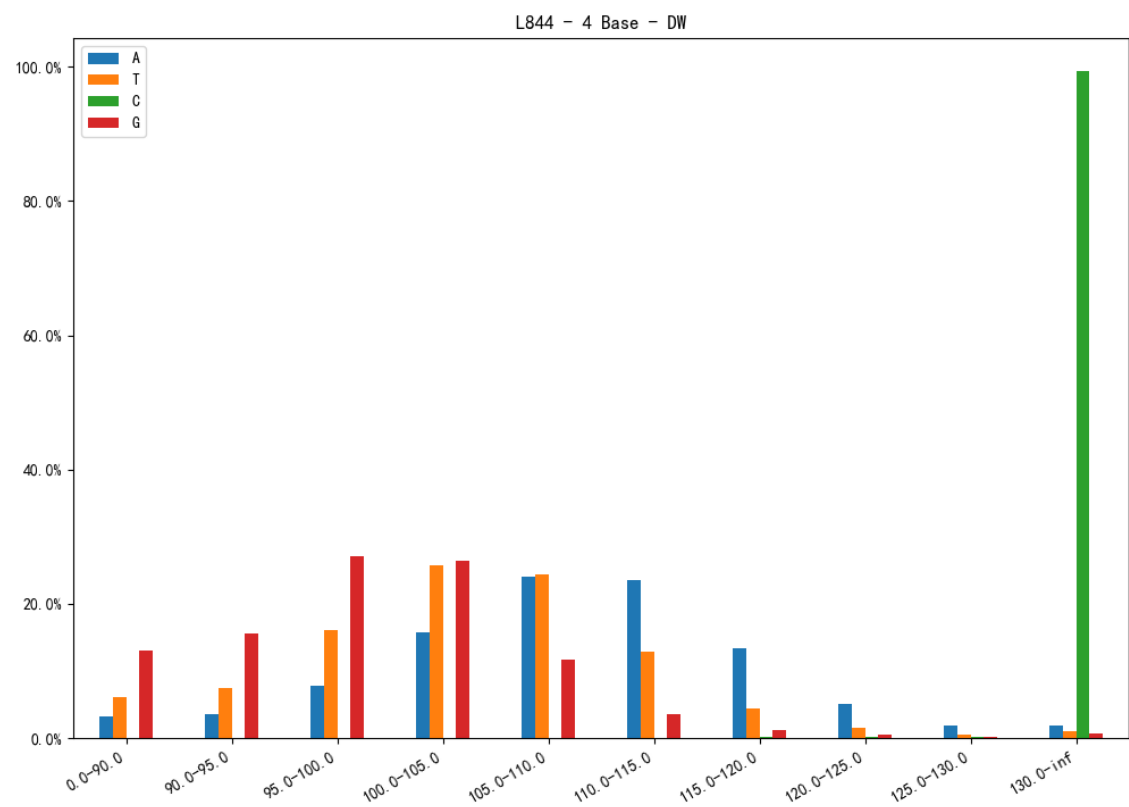
Fig 5-2-1 DW分布条形图



L617 - 四种碱基DW占比图



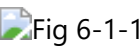
L844 - 四种碱基DW占比图



## 6. Capture Rate相关统计

### 6.1 与Block Rate的关系

L617 - CR与BR相关性散点图

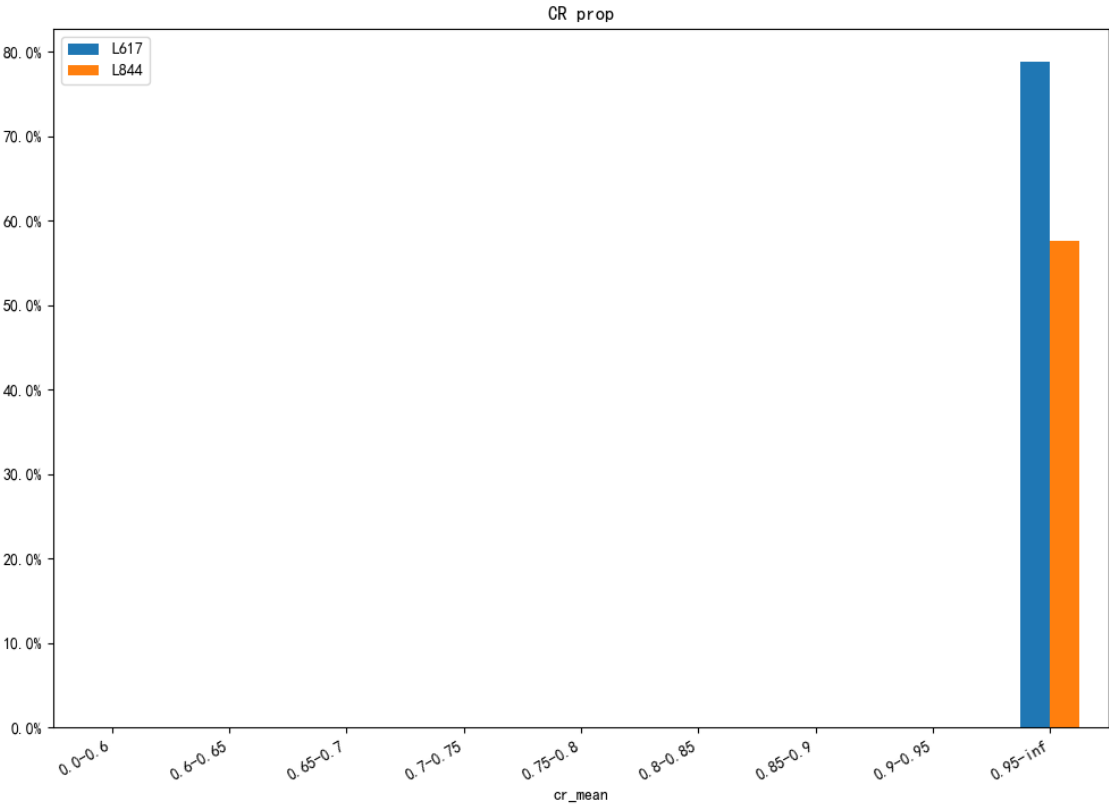


L844 - CR与BR相关性散点图



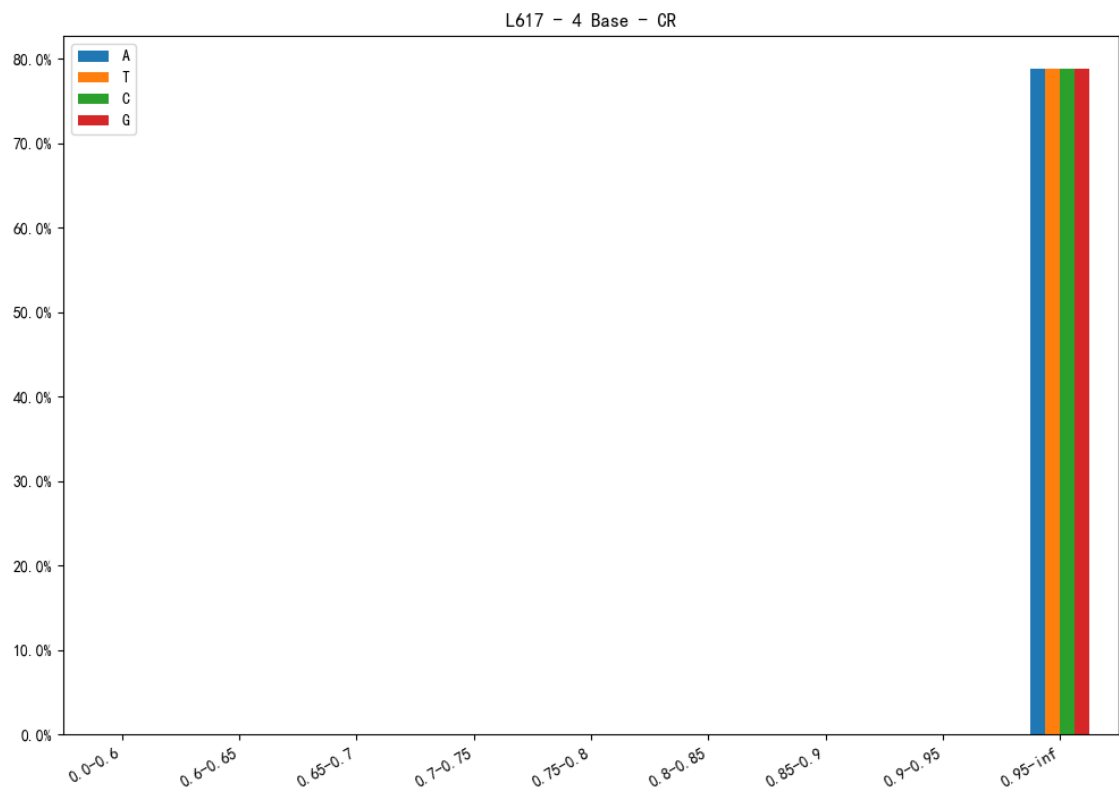
### 6.2 CR分布

Fig 6-2-1 CR分布条形图

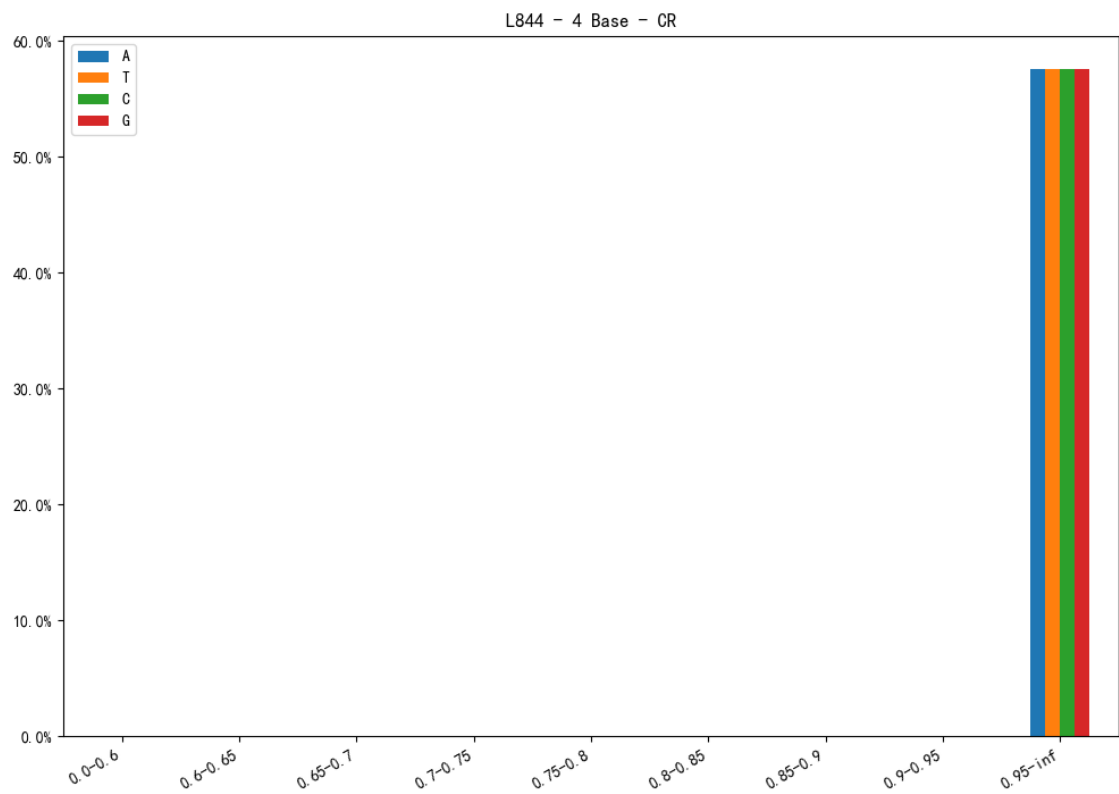




L617 - 四种碱基CR占比图

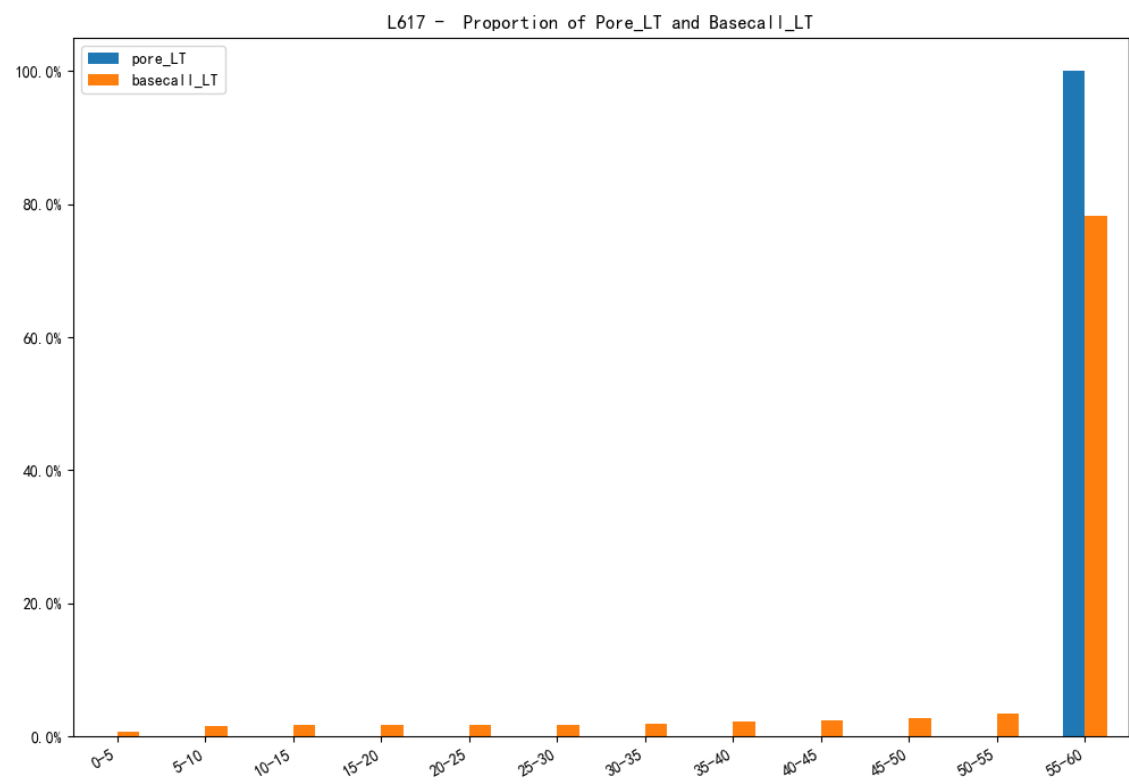


L844 - 四种碱基CR占比图



7. Pore-LT与Basecall-LT相关统计

L617 Pore-LT与Basecall-LT占比图



L844 Pore-LT与Basecall-LT占比图

