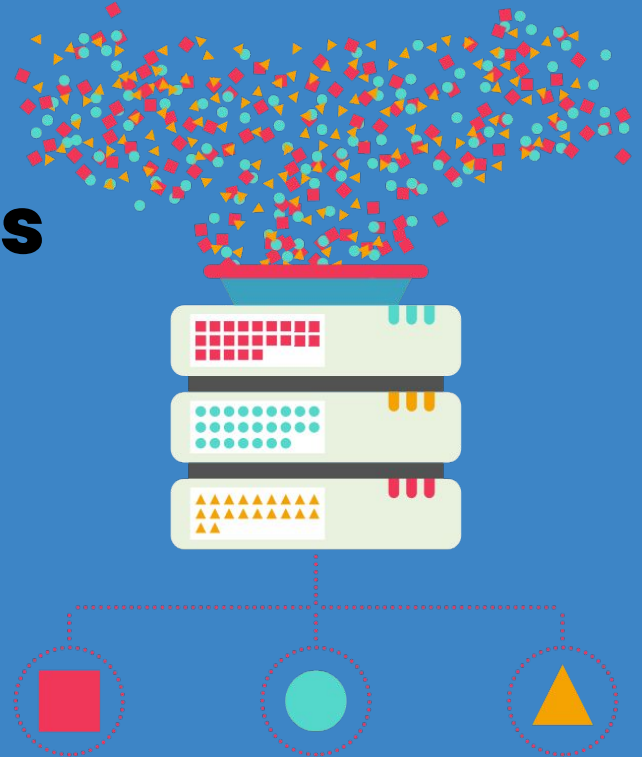


Detection of Antibiotic resistance genes in human gut microbiota

- Al-Hassan Mohamed
- Ahmed Saleh
- Abdallah Sayed
- Sarah Ahmed
- Gamila Mohamed
- Shahd Wael
- Heba Allah Mohamed



Summary

We used a DNA microarray chip covering 369 resistance types to investigate the relation of antibiotic resistance gene diversity with humans' age.

Meta-genomic DNA from fecal samples of 123 healthy volunteers of four different age groups:

- preschool Children (CH)
- School Children (SC)
- High School Students (HSS)
- Adults (AD)

The results showed that

- 80 different gene types were recovered from the 123 individuals gut microbiota, among which 25 were present in CH, 37 in SC, 58 in HSS and 72 in AD.
- Further analysis indicated that antibiotic resistance genes in groups of CH, SC and AD can be independently clustered, and those ones in group HSS are more divergent.

Data resource

- <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54070>

Dataset description

IMAGE_ID	GENE_EXPR_OPTION	SEQ_ID	PROBE_ID	POSITION	X	Y	MATCH_INDEX	SEQ_URL	PM	MM
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00101	101	359	315	725640	NA	656	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00160	160	53	375	725641	NA	606	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00162	162	240	402	725642	NA	522	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00165	165	307	97	725643	NA	1116	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00168	168	289	553	725644	NA	435	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00171	171	413	309	725645	NA	656	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00174	174	51	445	725646	NA	996	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00177	177	210	202	725647	NA	897	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00180	180	453	475	725648	NA	610	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00183	183	210	192	725649	NA	990	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00303	303	428	306	725650	NA	446	0
388797A03_100121_532_532	BLOCK1	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00429	429	127	243	725651	NA	1644	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00101	101	152	324	863244	NA	1218	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00160	160	416	188	863245	NA	588	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00162	162	141	391	863246	NA	703	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00165	165	379	389	863247	NA	462	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00168	168	49	121	863248	NA	962	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00171	171	390	304	863249	NA	627	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00174	174	304	426	863250	NA	338	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00177	177	131	139	863251	NA	994	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00180	180	46	208	863252	NA	1215	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00183	183	123	375	863253	NA	1057	0
388797A03_100121_532_532	BLOCK2	AAC2IA_L06156_264_797	AAC2IA_L06156_264_797P00303	303	15	7	863254	NA	1141	0

- **Data attributes:**

- Image ID
- Gene_exp_option => Block or subarray: a group of spots, typically made in one print round
- Seq_id
- Probe_id
- Position
- X
- Y
- Match_index
- Seq_url
- PM
- MM

- **Number of samples:** 123

- **Number of records:** 825624

T-test

- We used the mean of pre_school group as a standard value
- Then calculated t-test for each of the following groups based on it.

- **School Children (SC) =>**

```
t = -900.55, df = 275207, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 11.3309
95 percent confidence interval:
 10.07840 10.08384
sample estimates:
mean of x
 10.08112
```

- **High School Students (HSS) =>**

```
t = -637.79, df = 275207, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 11.3309
95 percent confidence interval:
 10.73427 10.73793
sample estimates:
mean of x
 10.7361
```

- **Adults (AD) =>**

```
t = -739.78, df = 275207, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 11.3309
95 percent confidence interval:
 10.04998 10.05675
sample estimates:
mean of x
 10.05337
```

Visualizations – Sample

Box-plot

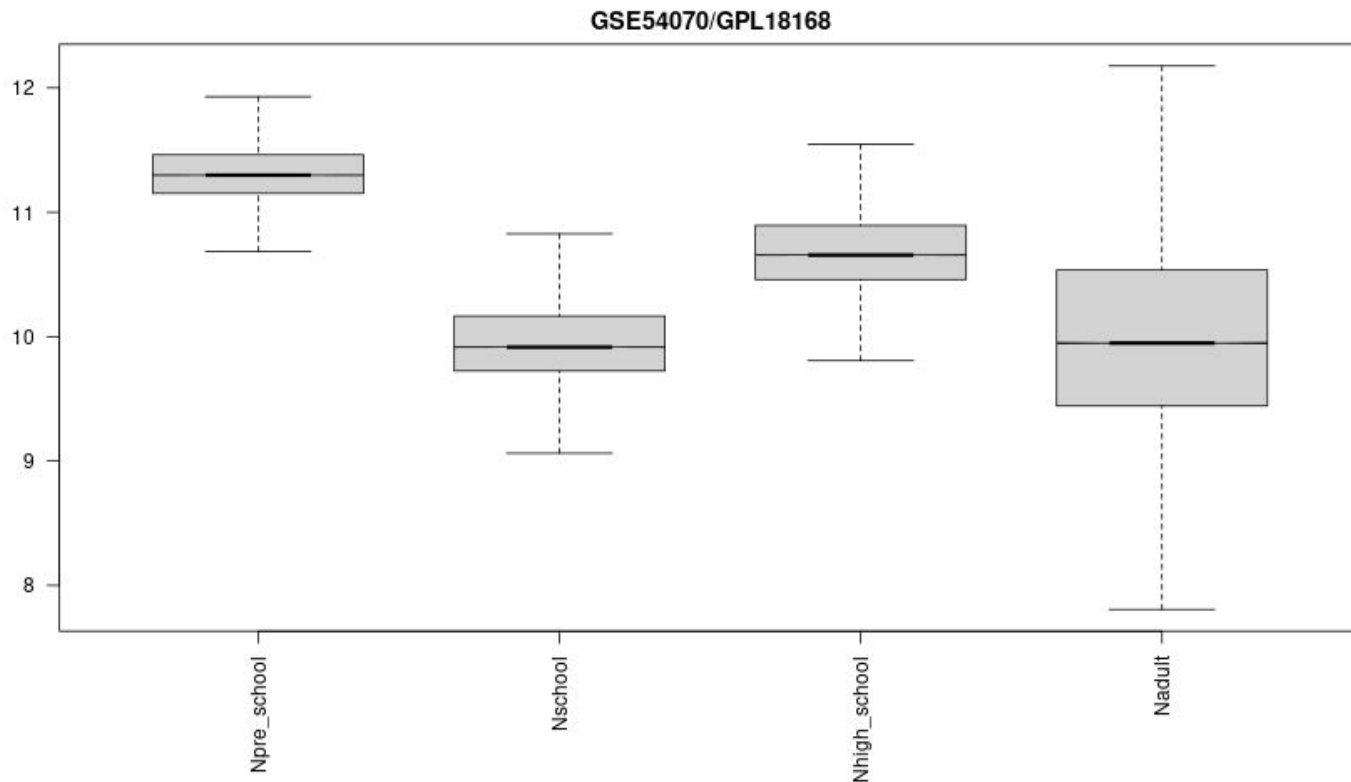
to show the distribution of data.

```
> summary(Npre_school)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 8.164  11.152 11.298 11.331 11.462 16.000

> summary(Nschool)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 8.475  9.724  9.916 10.081 10.165 15.925

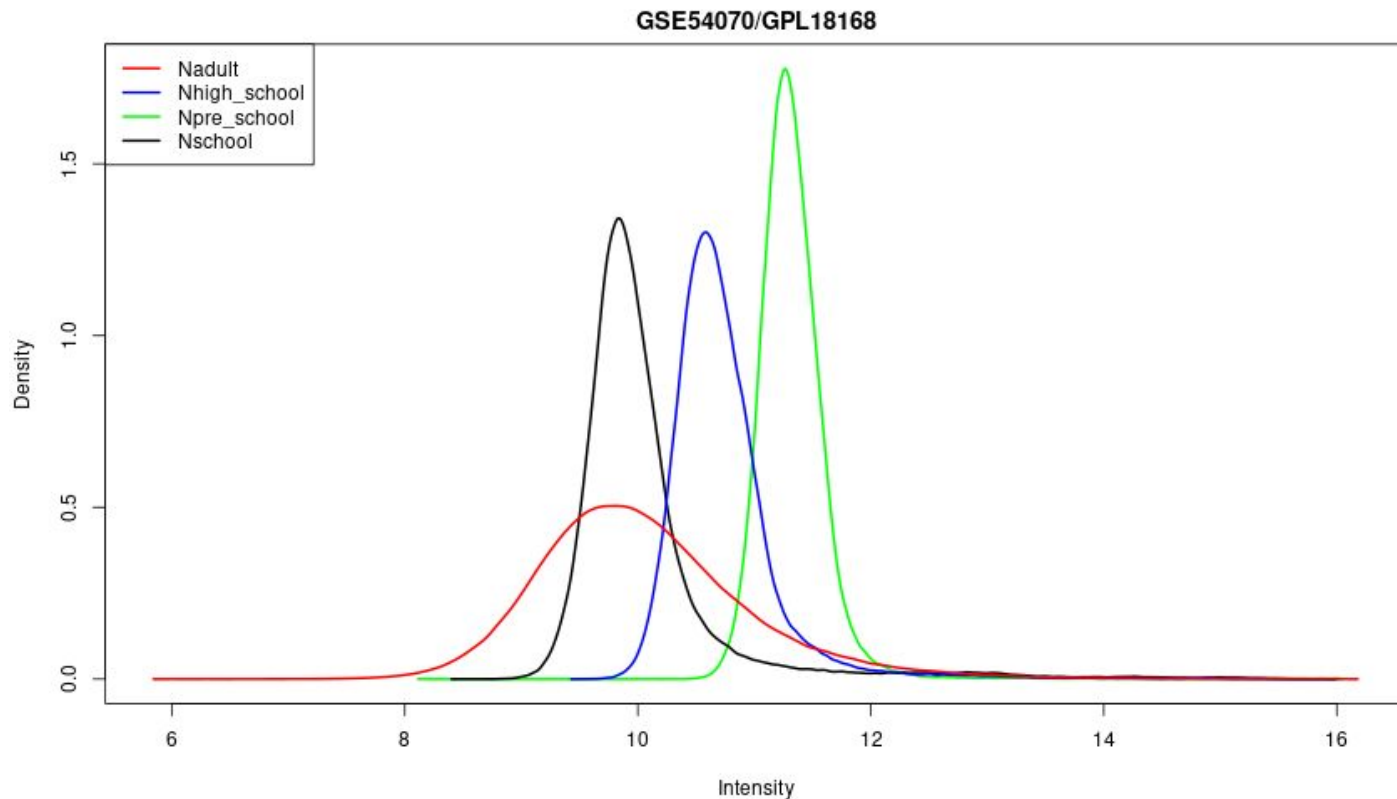
> summary(Nhigh_school)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 9.50  10.46  10.66 10.74 10.89 15.90

> summary(Nadult)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 6.022  9.443  9.948 10.053 10.536 16.000
```



Visualizations - Sample

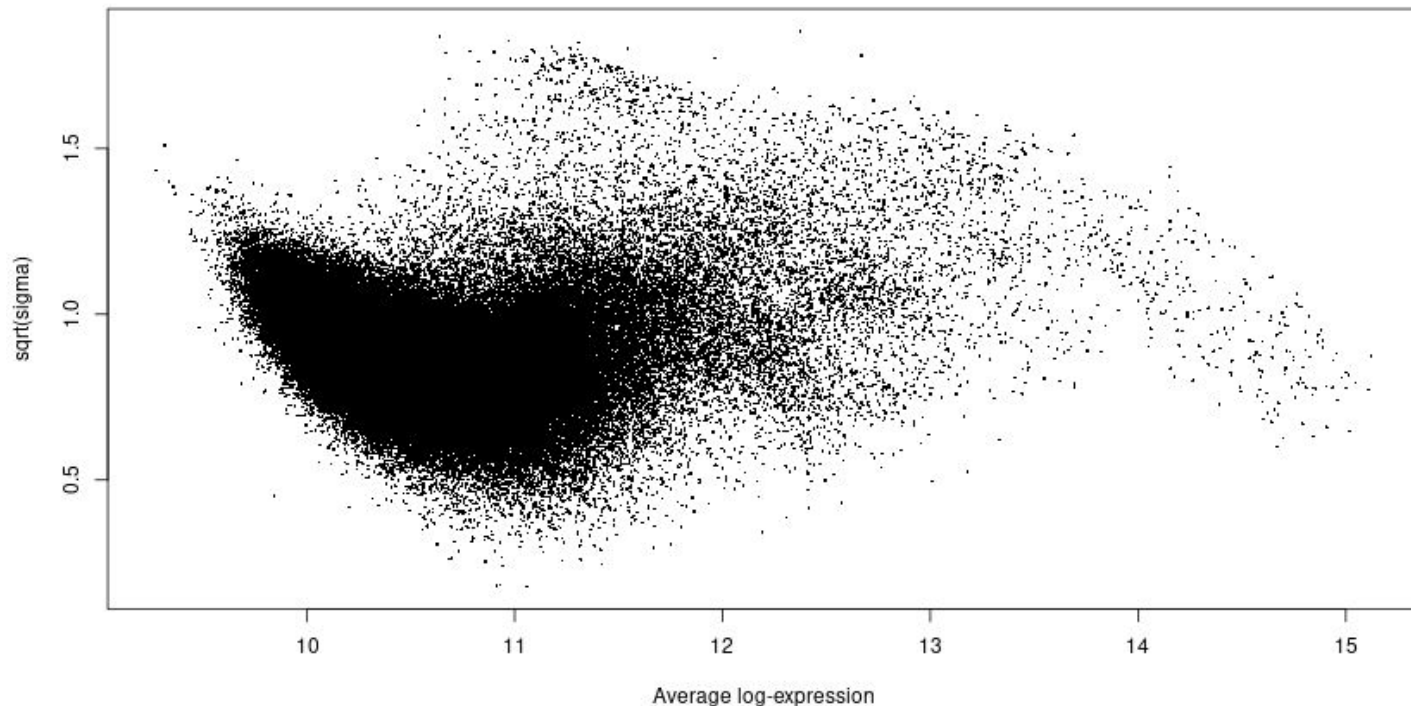
Expression value
distribution to



Visualizations – Sample

Variance of the mean

calculates how much the data points spread around the regression line.





Thanks!

Do you have any questions?