



Advanced Molecular Detection
Southeast Region Bioinformatics

R ggplot2

09/09/2024

Updates

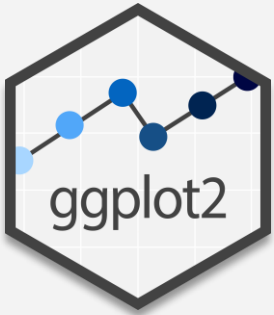
What's ahead?

Update-

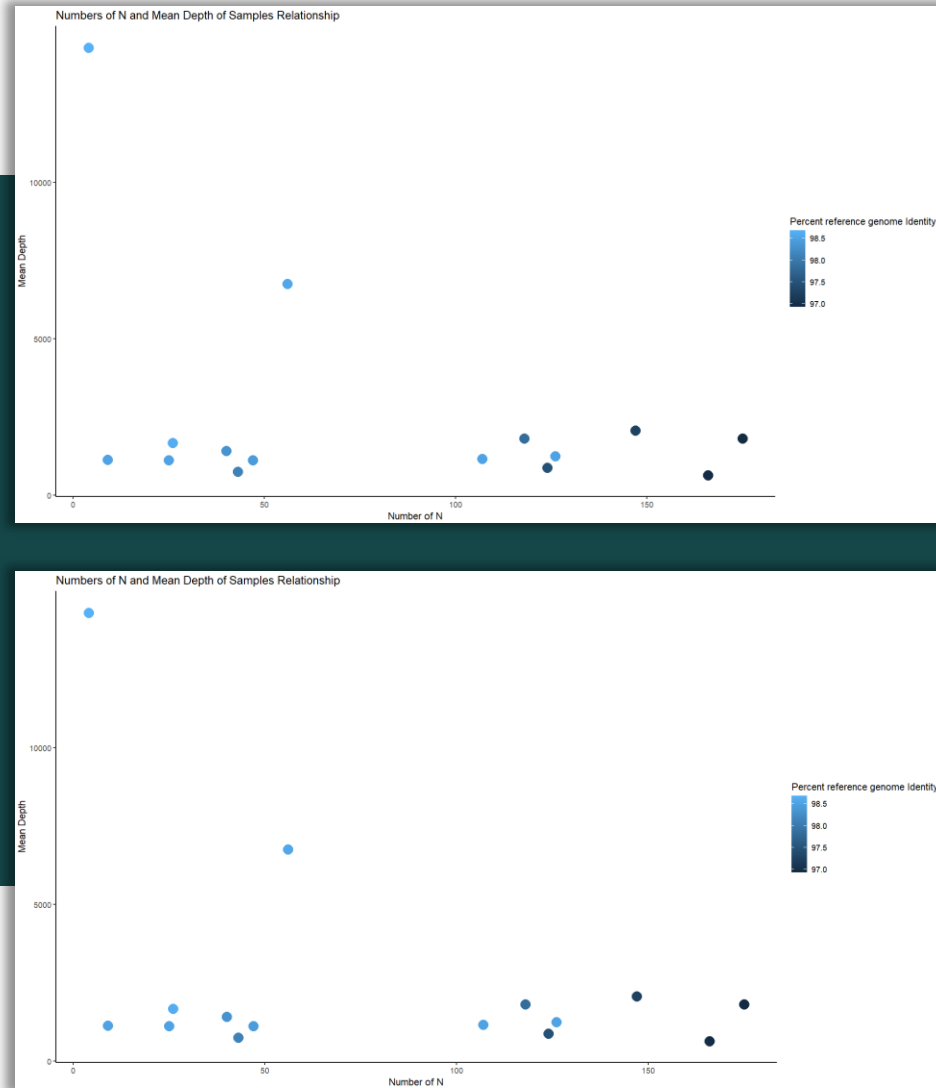
- NF Tower purchase is approved by FDOH

Office Hour-

- **September 30** – ggtree
- **October 14** – To be determined



Overview



Synopsis

- An R package which allows the user to create graphs by mapping data variables
- Strictly follows The Grammar of Graphics
- Access to wide range of plots



Overall Code: `ggplot(usedata, aes(x=numN, y=mean_depth, color= percent_ref_genome_cov)) +
geom_point(size=5)+ labs(title="Numbers of N and Mean Depth of Samples Relationship",
x= "Number of N", y="Mean Depth", color="Percent reference genome Identity")+ theme_classic()`

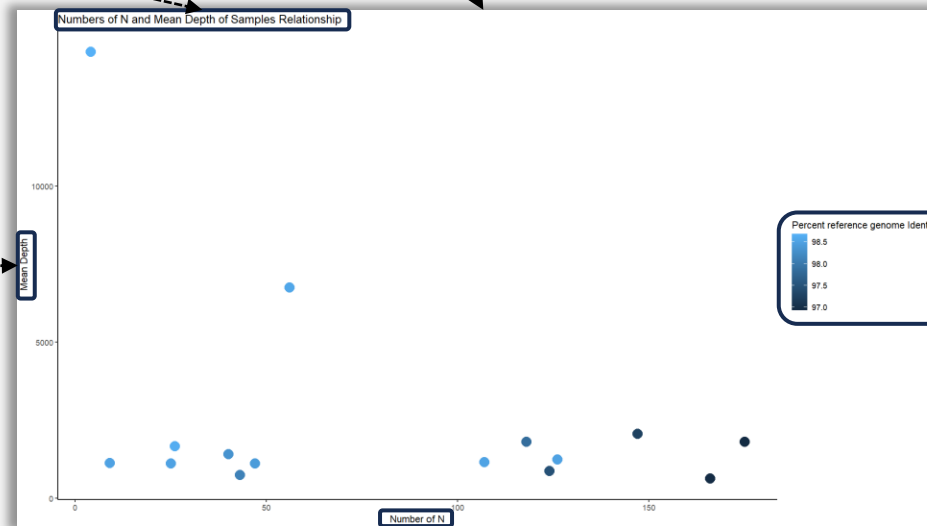
Main Title=

`labs(title="Numbers of N and Mean Depth of Samples Relationship")`

Closer Look

Title Axis=

`labs (y="Mean
Depth")`



Legend Title=

`labs(color="Perc
ent reference
genome Identity")`

Example:

Data = `usedata`

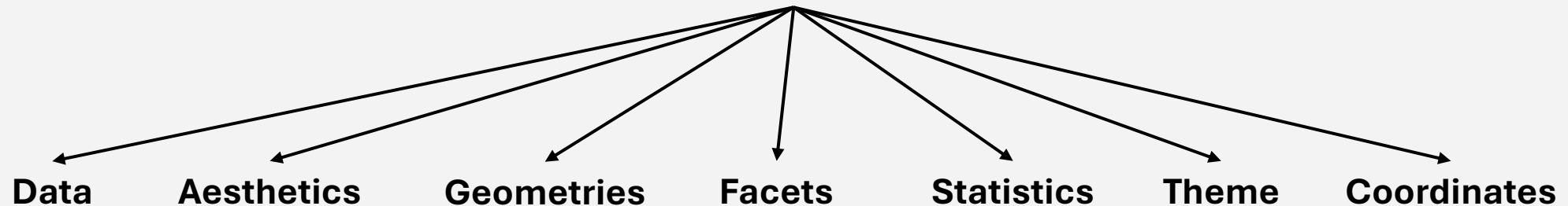
X Variable = `aes(x=numN)`

Y Variable = `aes(y=mean_depth)`

3rd Variable or Color Scheming (Reference to the Legend)= `aes(color=percent_ref_genome_cov)`

Grammar of Graphics

Overall Code: `ggplot(usedata, aes(x=numN, y=mean_depth, color= percent_ref_genome_cov)) + geom_point(size=5)+ labs(title="Numbers of N and Mean Depth of Samples Relationship", x= "Number of N", y="Mean Depth", color="Percent reference genome Identity")+ theme_classic()`



How Do We Begin?



Installation-

Option 1: `install.packages("ggplot2")`

Or

Option 2: `install.packages("tidyverse")`



Load Package From Option 1 or 2-
`library(...)`



Load and Modify (if needed) Dataset-
For example: `usedata<-read_csv("Tampa_Dengue_TL_report.csv")`



Generate the Plot-

Overall Code: `ggplot(usedata, aes(x=numN, y=mean_depth, color= percent_ref_genome_cov)) +
geom_point(size=5)+ labs(title="Numbers of N and Mean Depth of Samples Relationship",
x= "Number of N", y="Mean Depth", color="Percent reference genome Identity")+ theme_classic()`

Applications

Overall Code: `ggplot(usedata, aes(x=numN, y=mean_depth, color= percent_ref_genome_cov)) +
geom_point(size=5)+ labs(title="Numbers of N and Mean Depth of Samples Relationship",
x= "Number of N", y="Mean Depth", color="Percent reference genome Identity")+ theme_classic()`



Example 1: `ggplot(usedata, aes(x=numN, y=mean_depth, size= percent_ref_genome_cov)) +
geom_point(color="Red")+ labs(title="Numbers of N and Mean Depth of Samples Relationship", x= "Number of N",
y="Mean Depth", size="Percent Reference Genome Identity") + theme_classic()`



Example 2: `ggplot(usedata, aes(x=numN, y=mean_depth, color=percent_ref_genome_cov)) + geom_point()+
geom_smooth(method = "lm", se=FALSE, color="Orange")+ labs(title="Numbers of N and Mean Depth of Samples
Relationship",x= "Number of N", y="Mean Depth") + scale_color_gradient(low="blue",high = "red")+ theme_minimal()`



Example 3: `ggplot(usedata, aes(x=numN, y=mean_depth)) + geom_point()+ facet_wrap(~percent_ref_genome_cov)+
labs(title="Numbers of N and Mean Depth of Samples Relationship",x= "Number of N", y="Mean Depth") +
theme_minimal()`



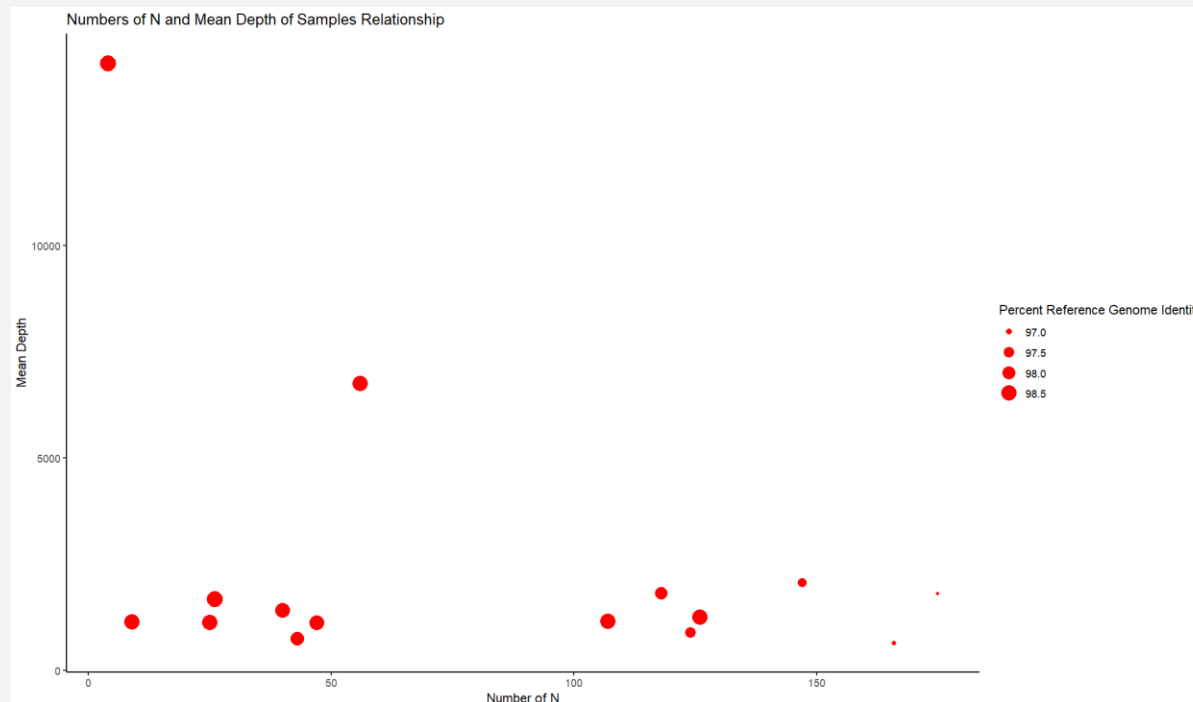
Example 4: `ggplot(usedata, aes(x=numN, y=mean_depth, color=percent_ref_genome_cov)) + geom_point(size=5)+
facet_wrap(~percent_ref_genome_cov) + labs(title="Numbers of N and Mean Depth of Samples Relationship",x=
"Number of N", y="Mean Depth", color="Percent Reference Genome Identity")+ theme_minimal()+
scale_color_gradient(low="blue",high = "red")`



Example 5: `ggplot(usedata, aes(x=numN, y=mean_depth))+geom_jitter(aes(color=percent_ref_genome_cov,
width=0.2,size=3))+ geom_area(position = "stack",fill="black",size=1)+ scale_color_gradient(low="blue",high =
"red")+theme_minimal()`

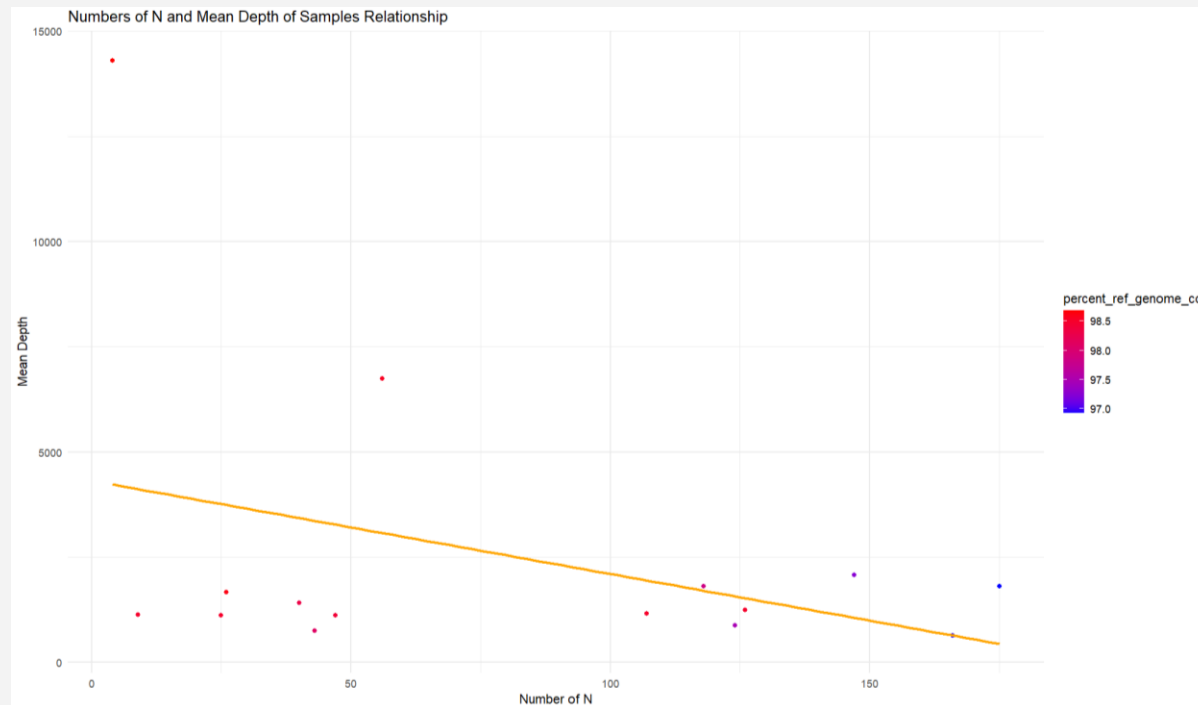
Applications Cont.

Example 1: `ggplot(usedata, aes(x=numN, y=mean_depth, size= percent_ref_genome_cov)) +
geom_point(color="Red")+ labs(title="Numbers of N and Mean Depth of Samples Relationship", x= "Number of N",
y="Mean Depth", size="Percent Reference Genome Identity") + theme_classic()`



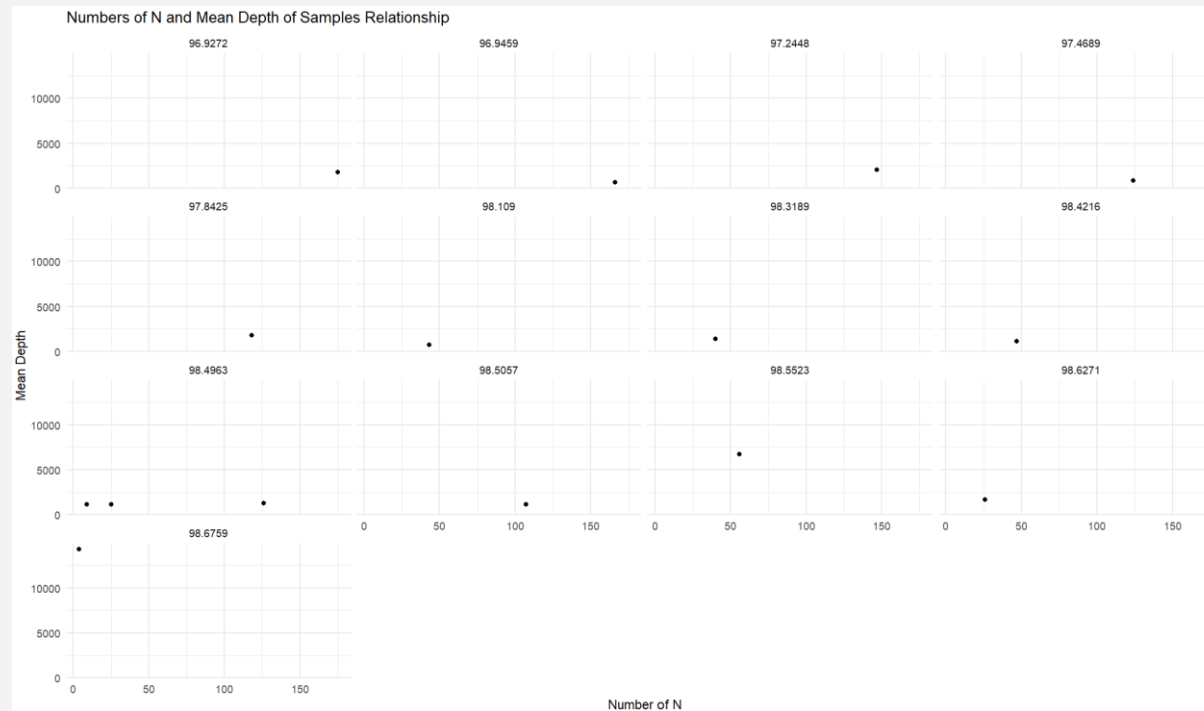
Applications Cont.

Example 2: `ggplot(usedata, aes(x=numN, y=mean_depth, color=percent_ref_genome_cov)) + geom_point()+
geom_smooth(method = "lm", se=FALSE, color="Orange")+ labs(title="Numbers of N and Mean Depth of Samples
Relationship",x= "Number of N", y="Mean Depth") + scale_color_gradient(low="blue",high = "red")+ theme_minimal()`



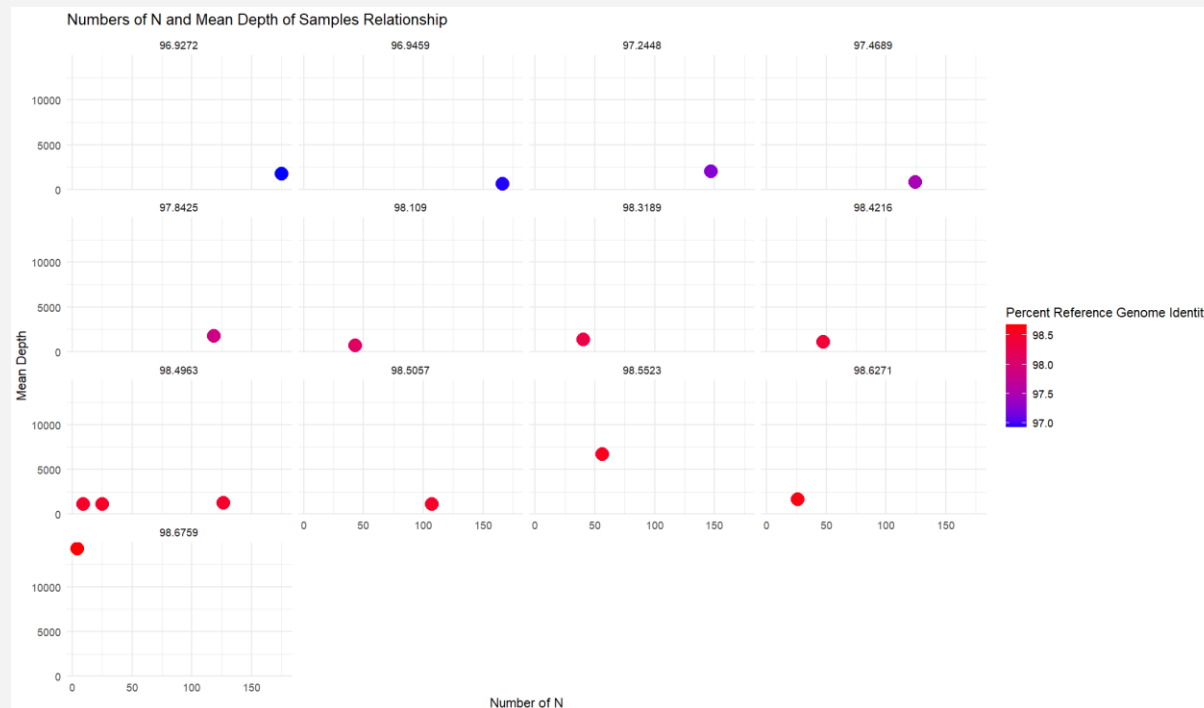
Applications Cont.

Example 3: `ggplot(usedata, aes(x=numN, y=mean_depth)) + geom_point()+ facet_wrap(~percent_ref_genome_cov)+
labs(title="Numbers of N and Mean Depth of Samples Relationship",x= "Number of N", y="Mean Depth") +
theme_minimal()`



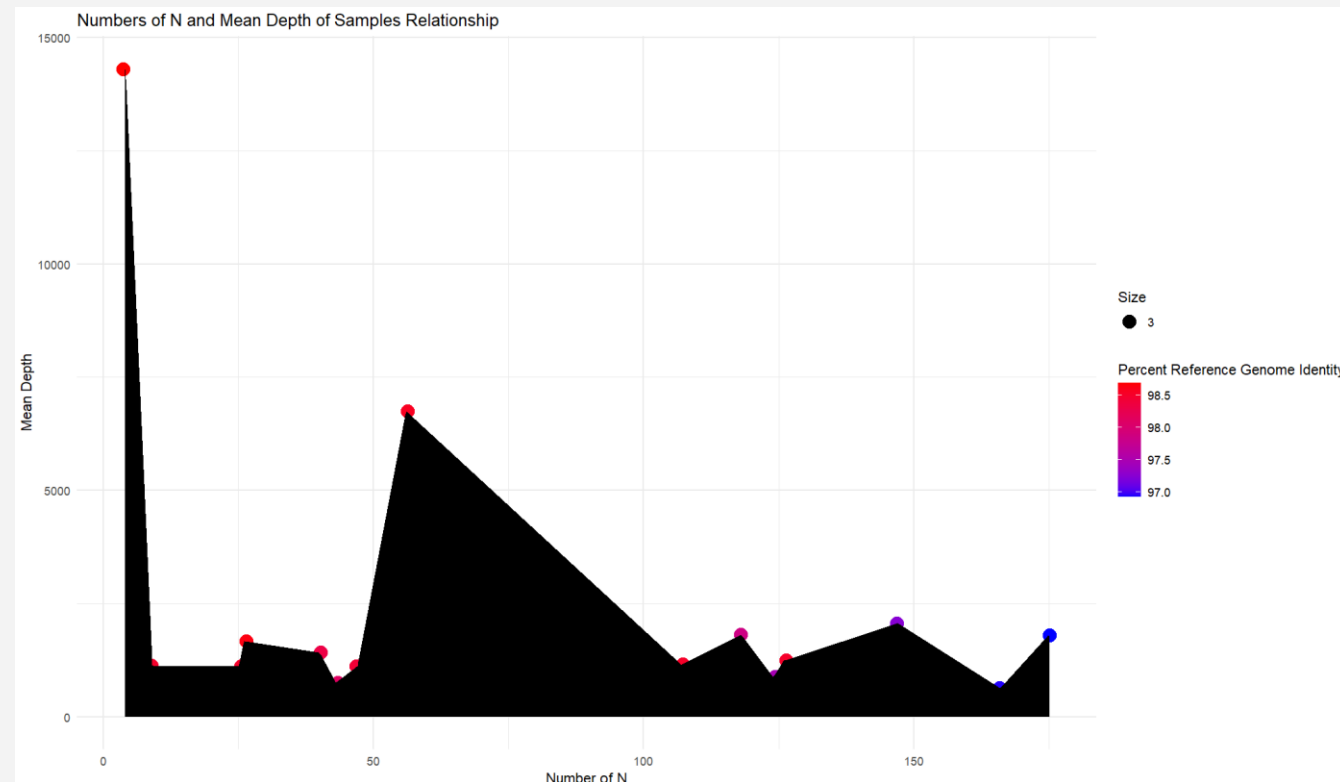
Applications Cont.

Example 4: `ggplot(usedata, aes(x=numN, y=mean_depth, color=percent_ref_genome_cov)) + geom_point(size=5)+
facet_wrap(~percent_ref_genome_cov) + labs(title="Numbers of N and Mean Depth of Samples Relationship",x=
"Number of N", y="Mean Depth", color="Percent Reference Genome Identity")+ theme_minimal()+
scale_color_gradient(low="blue",high = "red")`



Applications Cont.

Example 5: `ggplot(usedata, aes(x=numN, y=mean_depth))+geom_jitter(aes(color=percent_ref_genome_cov, width=0.2,size=3))+ geom_area(position = "stack",fill="black",size=1)+ scale_color_gradient(low="blue",high = "red")+theme_minimal()`



Conclusion

- ✓ Able to understand the conceptual aspect of ggplot2 and the purpose of The Grammar of Graphics
- ✓ Brief overview on how to simply plot a graph on ggplot2
- ✓ Different applications uses from the modification of the original R code.



Advanced Molecular Detection

Southeast Region Bioinformatics

Questions?

bphl-sebioinformatics@flhealth.gov

Molly Mitchell, PhD

Bioinformatics Supervisor

Molly.Mitchell@flhealth.gov

Nikhil Reddy, MS

Bioinformatician

Nikhil.Yengala@flhealth.gov

Sam Marcellus, MPH

Bioinformatician

Samantha.marcellus@flhealth.gov