

RNAseq - Internal Standard

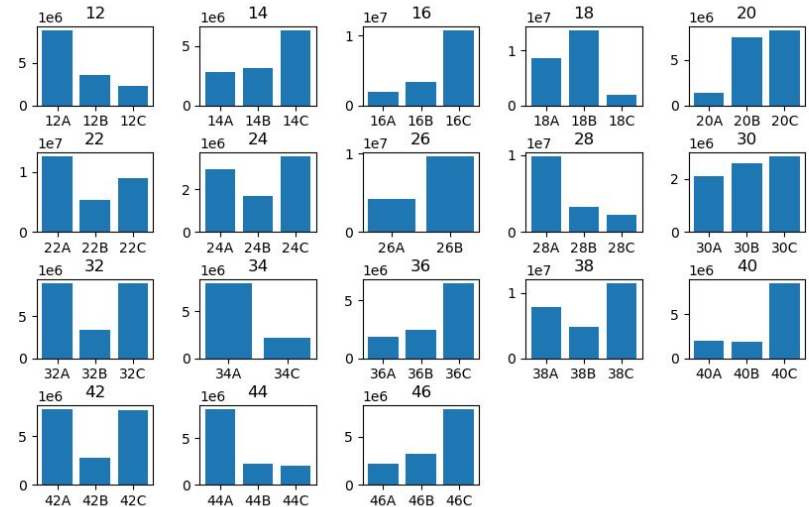
Why use Internal Standards in RNAseq?

RNAseq data is collected in a **relative framework**:

- Gene abundance is calculate as percent of sequence library

Limitation: cannot provide information on extent/direction of changes in any particular gene

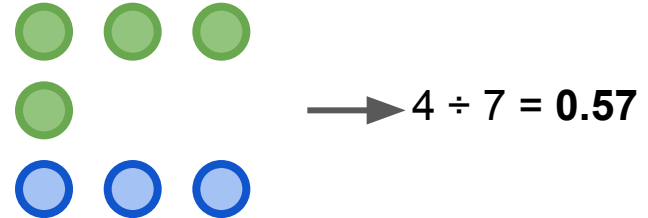
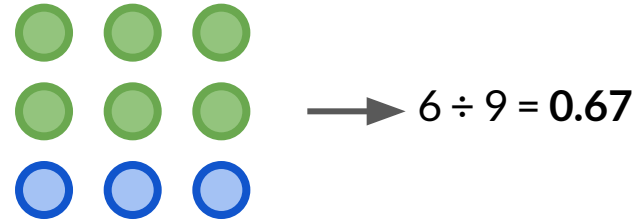
Counts of total reads mapped



Why use Internal Standards in RNAseq?

Limitation: cannot provide information on extent/direction of changes in any particular gene

Example: a decrease in expression of a transcript can be due to either:

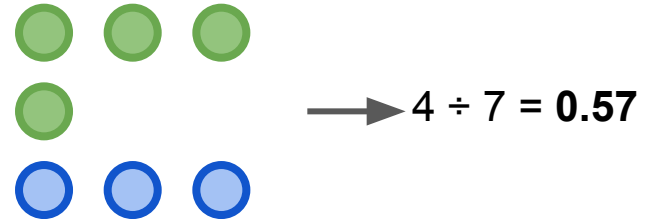
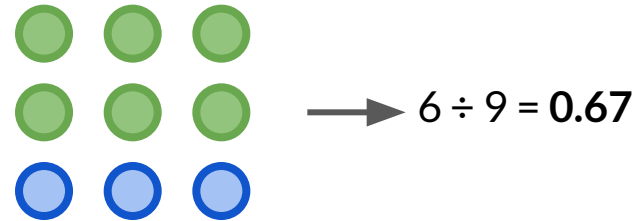


Why use Internal Standards in RNAseq?

Limitation: cannot provide information on extent/direction of changes in any particular gene

Example: a decrease in expression of a transcript can be due to either:

1. A decrease in abundance of that transcript

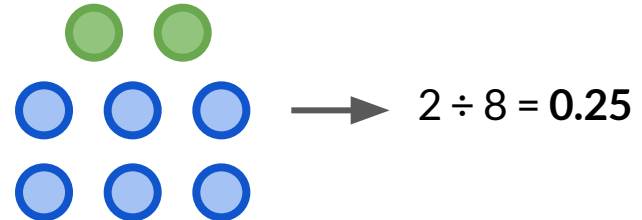
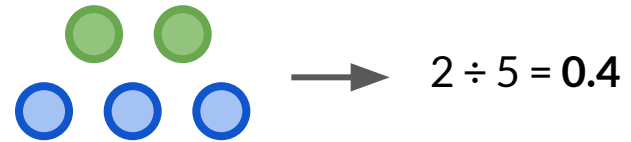


Why use Internal Standards in RNAseq?

Limitation: cannot provide information on extent/direction of changes in any particular gene

Example: a decrease in expression of a transcript can be due to either:

1. A decrease in abundance of that transcript
2. An increase in the abundance of unrelated transcript



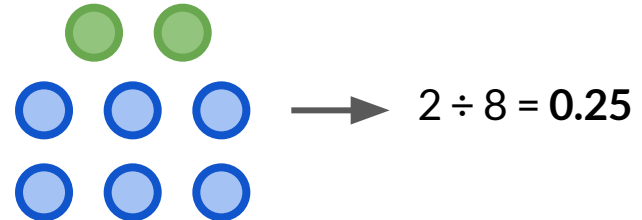
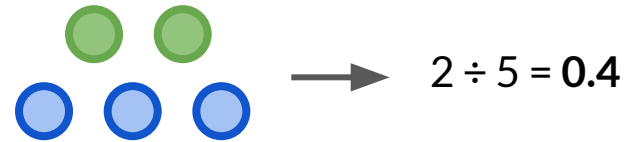
Why use Internal Standards in RNAseq?

Limitation: cannot provide information on extent/direction of changes in any particular gene

Example: a decrease in expression of a transcript can be due to either:

1. A decrease in abundance of that transcript
2. An increase in the abundance of unrelated transcript

→ Hard to differentiate between the 2



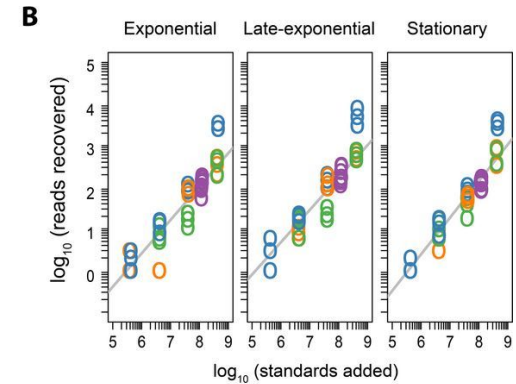
How to use Internal Standards in RNAseq?

Methods Summary:

- Add many different standard molecules at different known concentrations
- Obtain sequencing efficiency from: standards added vs. standard recovered

A

Group	ID	Length (nt)	Standards added	Symbol
A	Std. 5	979	441,334,667	○
	Std. 6	964	40,179,884	○
	Std. 1	1068	4,205,482	○
	Std. 12	991	428,059	○
B	Std. 11	996	410,023,533	○
	Std. 4	1028	39,362,526	○
	Std. 10	969	4,178,282	○
C	Std. 3	1000	408,554,669	○
	Std. 7	991	41,496,261	○
	Std. 8	1000	4,106,141	○
	Std. 2	1077	378,467	○
D	Std. 14	301	123,756,534	○
	Std. 13	613	119,426,056	○
	Std. 15	1504	115,793,038	○



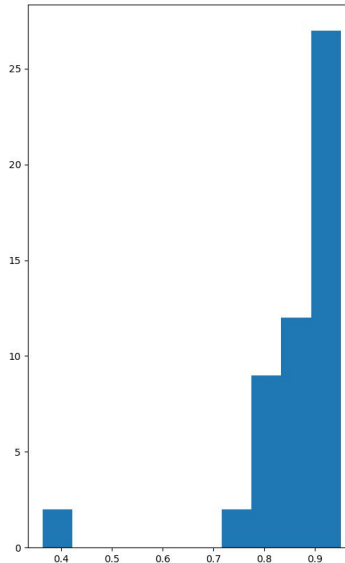
Gifford et al., 2016

How to use Internal Standards in RNAseq?

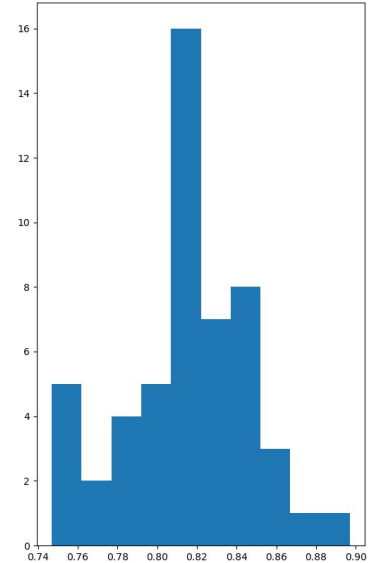
Methods Summary:

- Obtain median across all standards
- Fit linear line to median (global) standard data
- Filter standards to remove outliers
- Fit linear line to local standard data to obtain slope and intercept
- Use $y=mx+b$ to correct counts of gene transcript
- Divide transcript count by cell count to obtain transcript/cell

Slope Distribution before Filtering



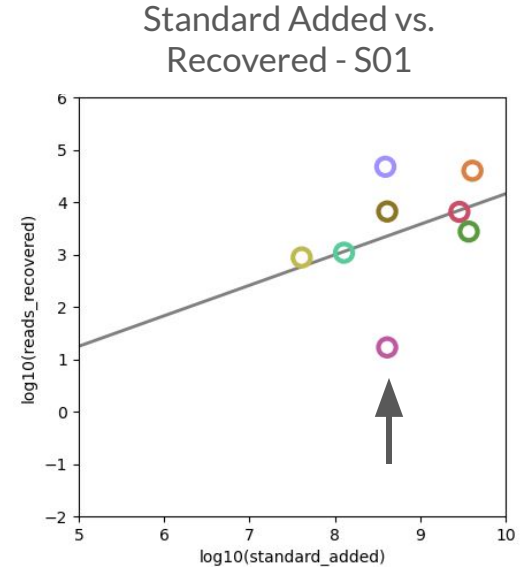
Slope Distribution after Filtering



How to use Internal Standards in RNAseq?

Standard Filtering:

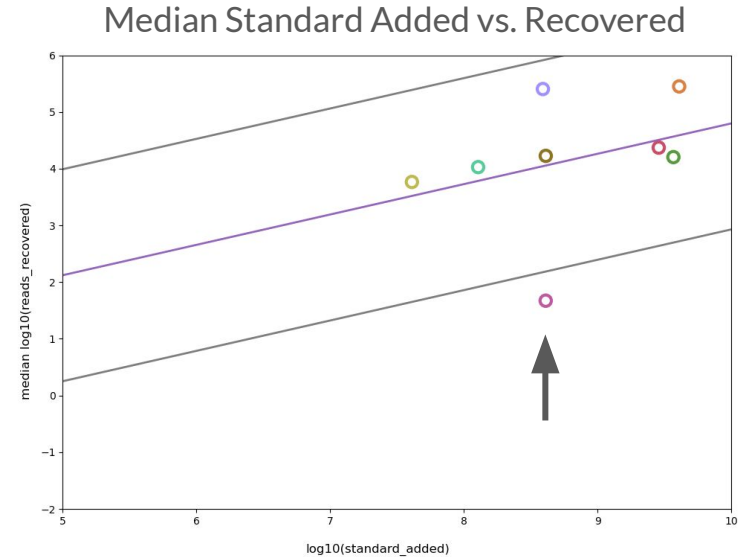
- Some standards are outliers → need to be removed



How to use Internal Standards in RNAseq?

Standard Filtering:

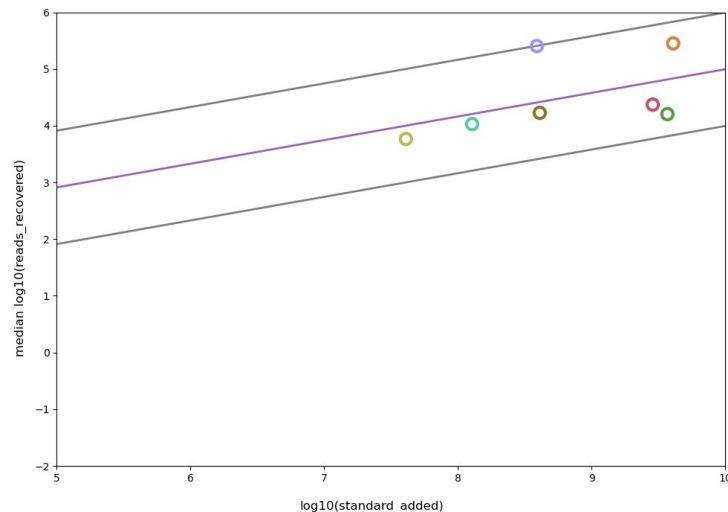
- Some standards are outliers → need to be removed
- Obtain median across all standards and filter standards globally
- Outliers are determined based on median of standards across samples



How to use Internal Standards in RNAseq?

Standard Filtering:

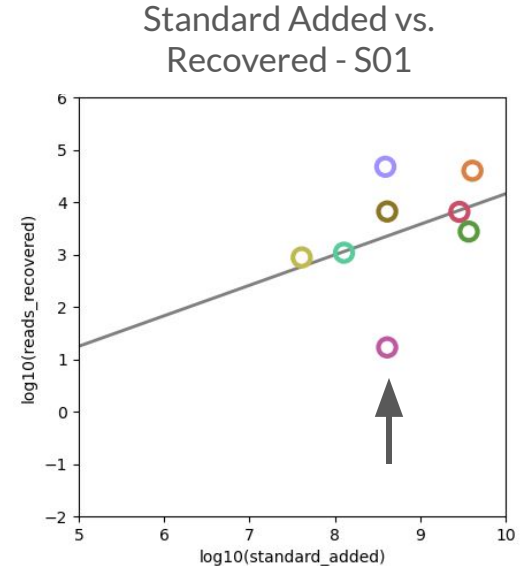
- Some standards are outliers → need to be removed
- Obtain median across all standards and filter standards globally
- Outliers are determined based on median of standards across samples



How to use Internal Standards in RNAseq?

Standard Filtering:

- Some standards are outliers → need to be removed
- Obtain median across all standards and filter standards globally
- Outliers are determined based on median of standards across samples
- Standards are then removed per sample



How to use Internal Standards in RNAseq?

Standard Filtering:

- Some standards are outliers → need to be removed
- Obtain median across all standards and filter standards globally
- Outliers are determined based on median of standards across samples
- Standards are then removed per sample
- Fit linear line to obtain slope and intercept for back-calculating transcript counts

