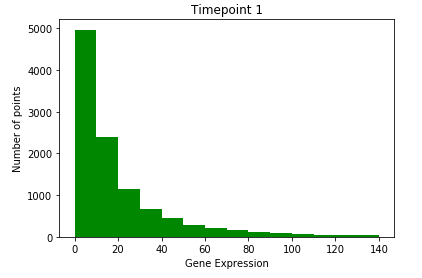
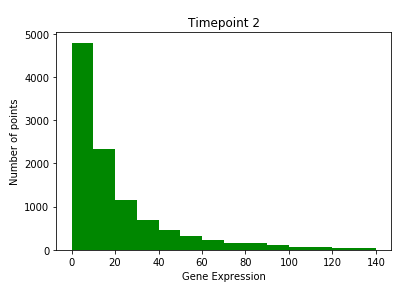
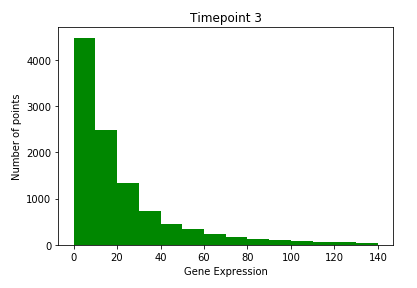
**Exploratory Analysis**

Number of points: 11087 (eliminated missing values)

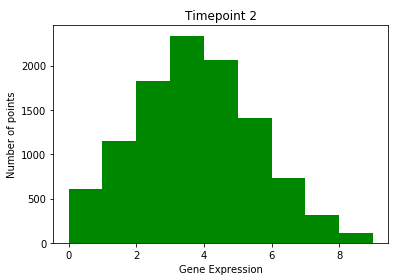
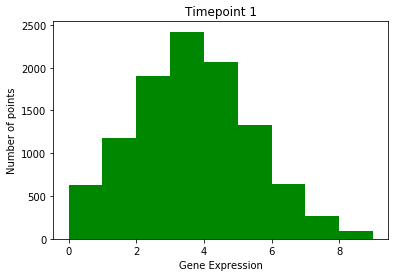
1. Gene Expression Value plot for each time point
2. Histogram of raw values

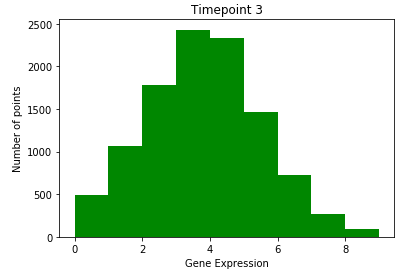




Seems like the timepoints are following an exponential distribution. However, in biology, we normalize gene expression values to log2 values.

1. Histogram of log2 values





This seems to be a Gaussian distribution.

1. Descriptive Statistics

**cdRPKM0 cdRPKM1 cdRPKM2**

**count** 11087.000000 11087.000000 11087.000000

**mean** 32.207382 35.000680 31.737374

**std** 100.702754 107.251175 79.427800

**min** 0.074259 0.073034 0.074316

**25%** 4.912363 5.002981 5.886184

**50%** 11.789318 12.465262 13.627347

**75%** 27.960392 30.408265 30.054516

**max** 2656.834519 2849.584643 2268.660153

1. Correlation Matrix

**cdRPKM0 cdRPKM1 cdRPKM2**

**cdRPKM0** 1.000000 0.969042 0.907932

**cdRPKM1** 0.969042 1.000000 0.950209

**cdRPKM2** 0.907932 0.950209 1.000000

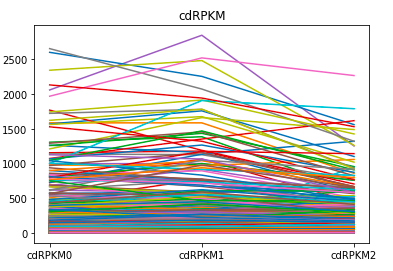
1. Skew

**cdRPKM0** 12.843991

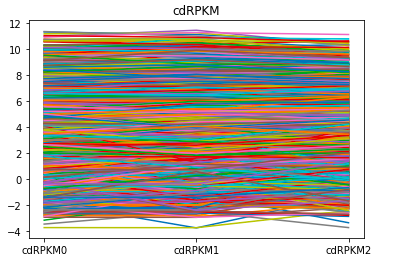
**cdRPKM1** 12.489221

**cdRPKM2** 11.373368

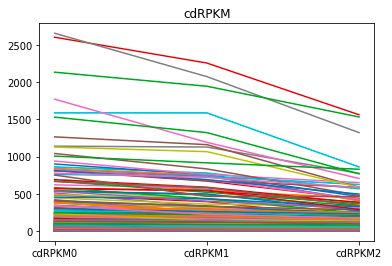
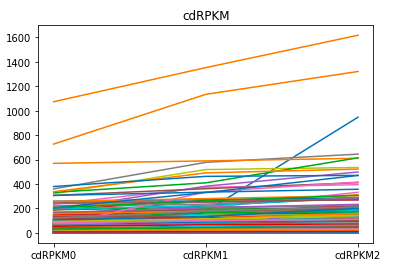
1. Plots of all gene expression values across time
2. **Raw values**

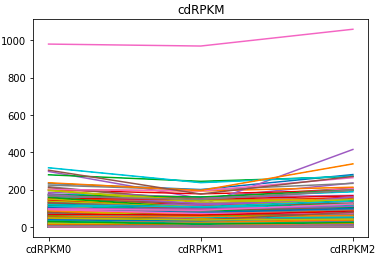
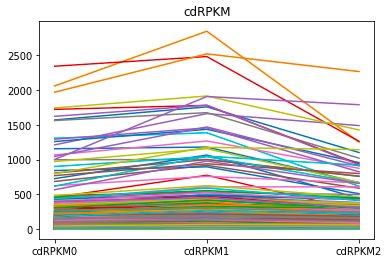
****

1. **Log plot**

****

1. **Plots according to expression features**

****

****

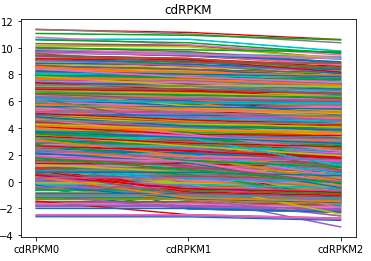
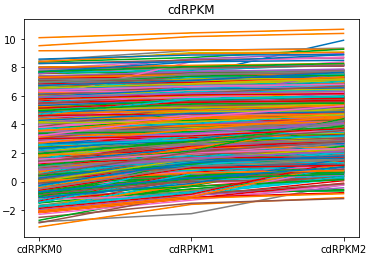
**Fig.** (Clockwise from top left)

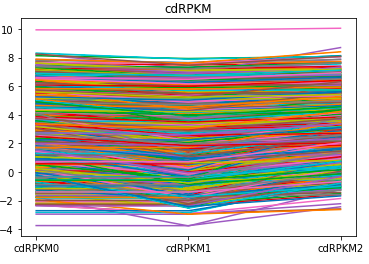
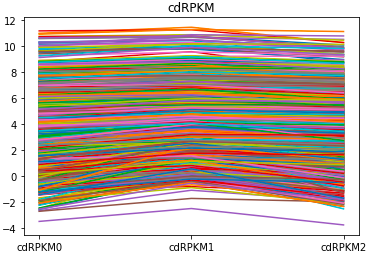
Monotonic increasing, monotonic decreasing

Increasing then decreasing, decreasing then increasing

Number of sequences in each: [2514, 1725, 3430, 3418]

1. **Log plots according to expression features**





**Fig.** (Clockwise from top left)

Monotonic increasing, monotonic decreasing

Increasing then decreasing, decreasing then increasing

Number of sequences in each: [2514, 1725, 3430, 3418]