

miRNA as a factor in gene expression noise.

Jan Majta

October 2, 2014

- In single cell there is approx. 20k genes giving 50k - 1M of total mRNA molecules.
 - Gene expression falls under Zipf distribution for human (a-c), murine (d), *C.elegans* (e) and yeast (f) cells.
 - It's the same for data from HEK cells system.

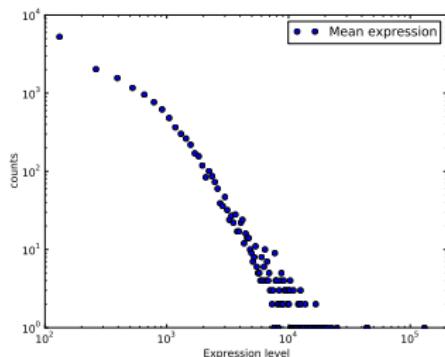


Figure: Data from HEK cells system sequencing

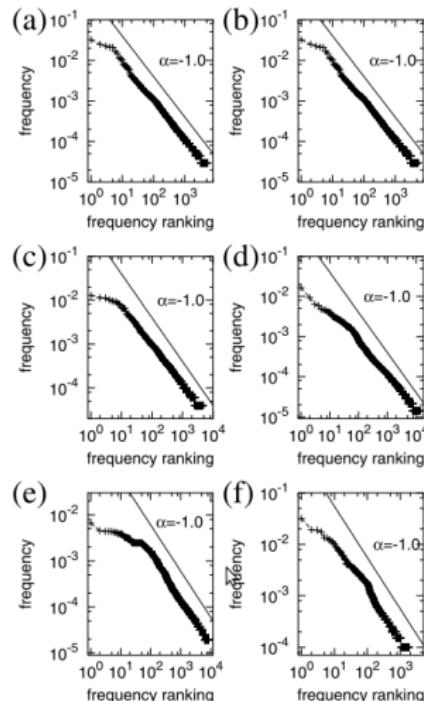
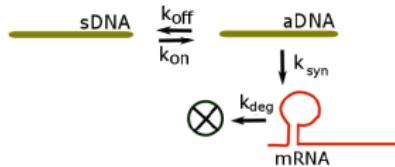


Figure: Furusawa, 2003

Model of expression:

- $E = \frac{k_{on} k_{syn}}{k_{deg}(k_{on} + k_{off})},$
- $Var = E(1 + \frac{k_{off} k_{syn}}{(k_{on} + k_{off})^2 + k_{deg}(k_{on} + k_{off})})$



- $\frac{k_{on}}{k_{off}}$ and $\frac{k_{syn}}{k_{deg}}$ are essential for expression. Absolute values changes noise.
- Gene expression distribution depends on $\frac{k_{on}}{k_{off}}$ relation
- $\rho(m) = \frac{\Gamma(\frac{\lambda}{\delta} + m)}{\Gamma(m+1)\Gamma(\frac{\lambda}{\delta} + \frac{\gamma}{\delta} + m)} \frac{\Gamma(\frac{\lambda}{\delta} + \frac{\gamma}{\delta})}{\Gamma(\frac{\lambda}{\delta})} (\frac{\mu}{\delta})^m {}_1F_1(\frac{\lambda}{\delta} + m, \frac{\lambda}{\delta} + \frac{\gamma}{\delta} + m, -\frac{\mu}{\delta})$

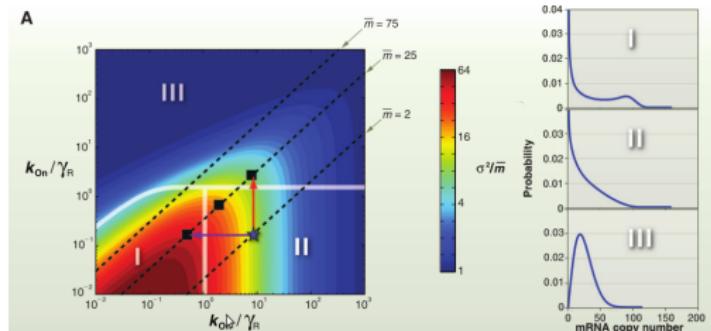


Figure: Different expression distributions. Munsky, 2012

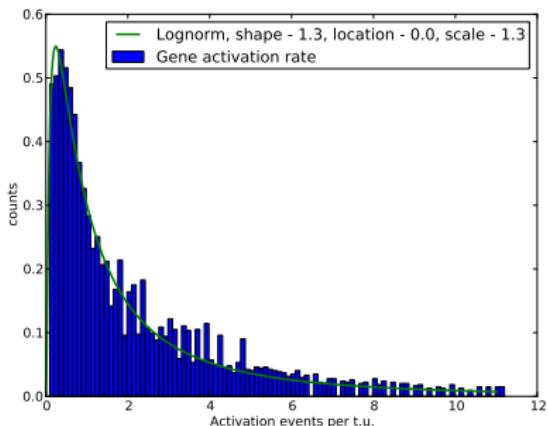


Figure: Activation rate (k_{on})

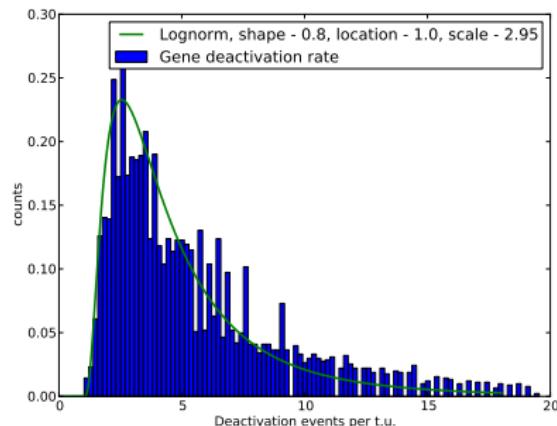


Figure: Deactivation rate (k_{off})

- Data from Kim, 2013
- For almost all genes $\frac{k_{on}}{k_{off}} < 1$
- Rates per time unit leaves space for tuning.

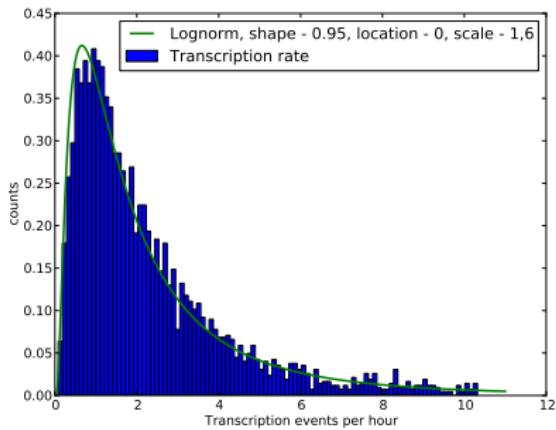


Figure: Transcription rate $S = \frac{k_{on}k_{syn}}{k_{on}+k_{off}}$

- Data from Schwanhausser, 2011.
- $k_{deg} = \ln(2)/t_{1/2}$

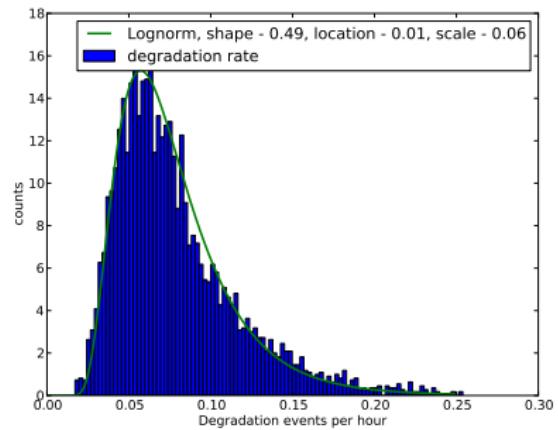


Figure: Degradation rate (k_{deg})

- $k_{syn} = \frac{transcription(k_{on}+k_{off})}{k_{on}}$
- It's only empirical solution, worth improving.
- To generate *in silico* cell, for every gene each parameter was drawn from its distribution

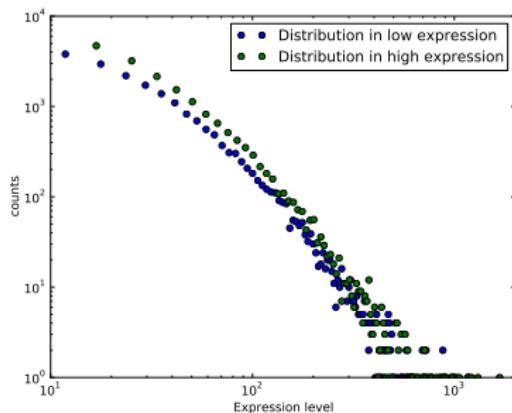


Figure: Distribution of gene expression.

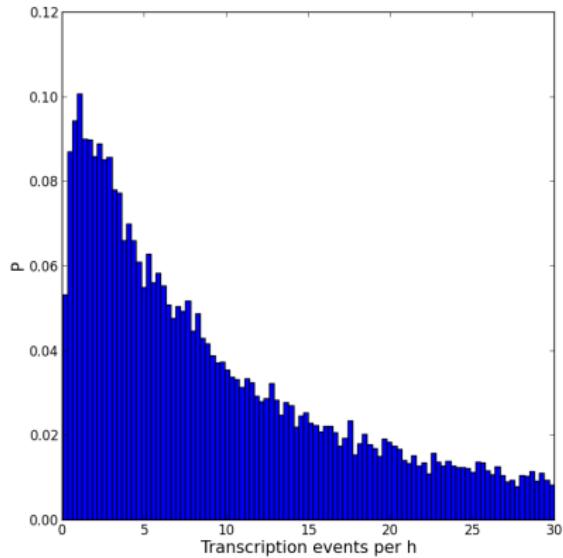


Figure: Distribution of calculated k_{syn}

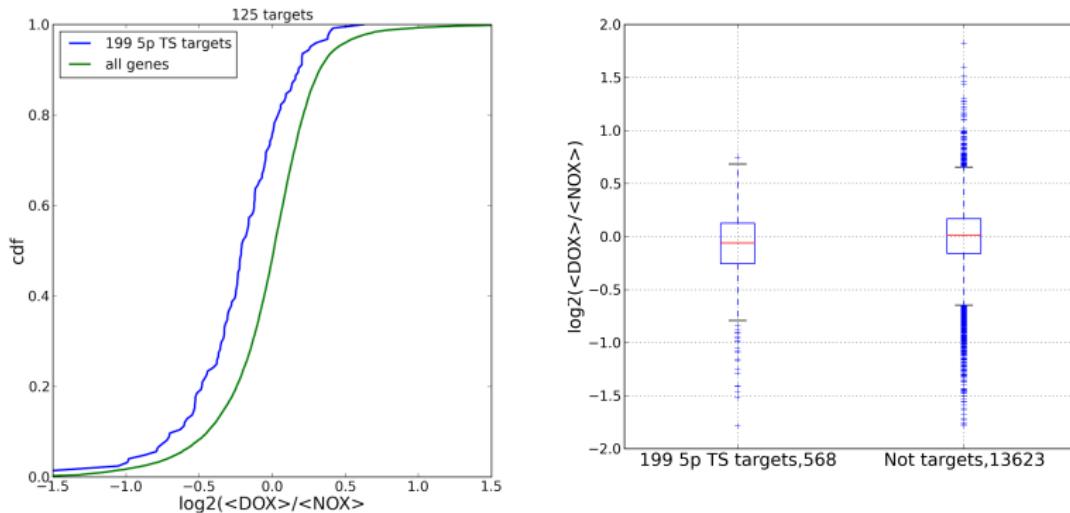


Figure: Change in gene expression after hsa-miR-199 expression induced.

- We assume, that miRNA should downregulate expression noise
- To mimic miRNA impact degradation rate was changed.
- Detecting such behaviour may be difficult due to: small change of noise and expression/sequencing noise itself

How to measure expression noise?

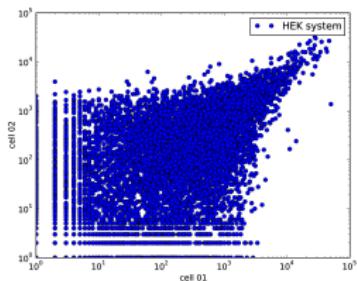


Figure: Expression noise between two HEK cells.

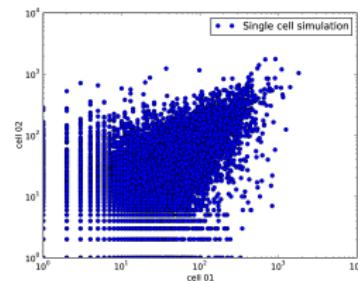


Figure: Expression noise between two *in silico* cells.

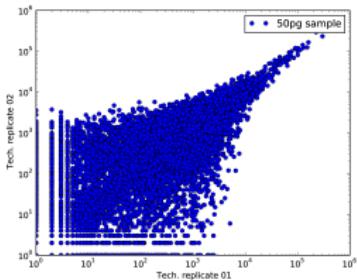
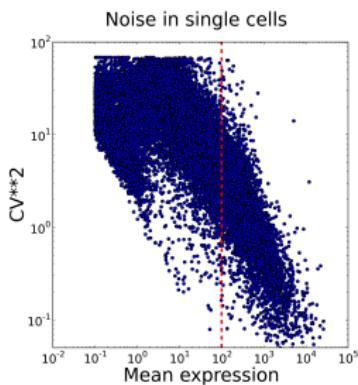


Figure: Noise of sequencing 50pg sample – technical replicates.



Simulations with one parameter set as variable. To observe noise change as reaction to increased degradation rate some conditions must be fulfilled:

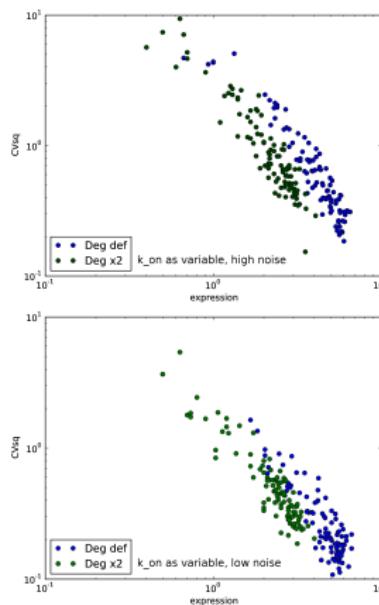


Figure: k_{on} , high and low noise level.

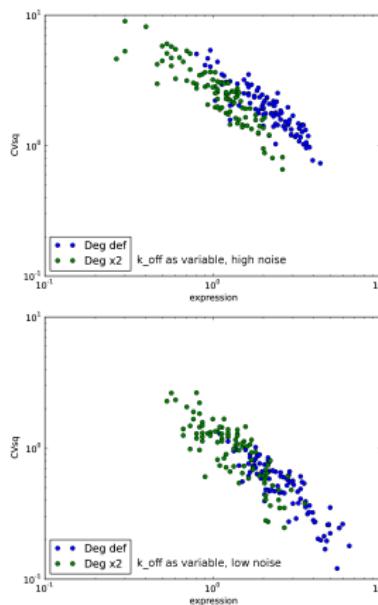


Figure: k_{off} , high and low noise level.

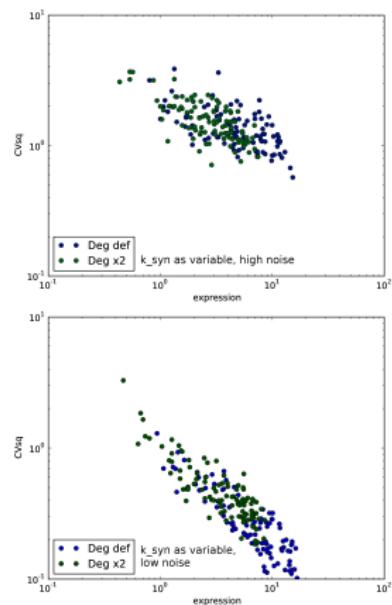
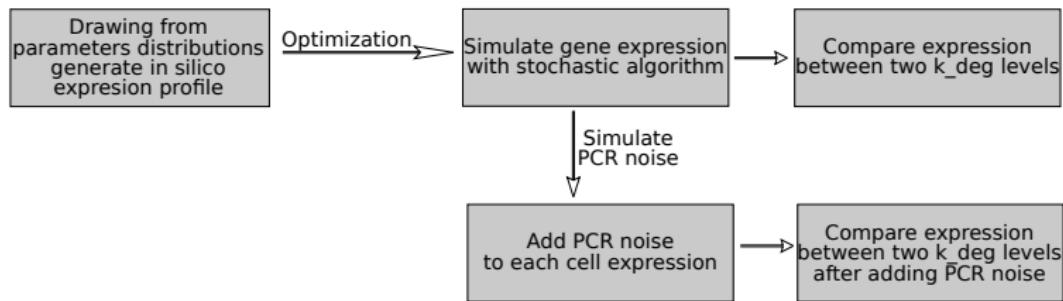


Figure: k_{syn} , high and low noise level.



- Adjusting $\frac{k_{on}}{k_{off}}$ level gives opportunity to tune noise level.

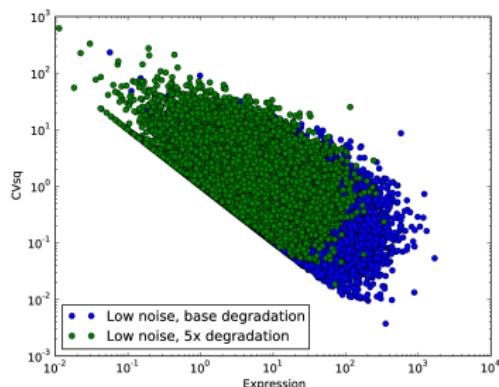


Figure: Calculated gene expression noise in function of expression level, low noise.

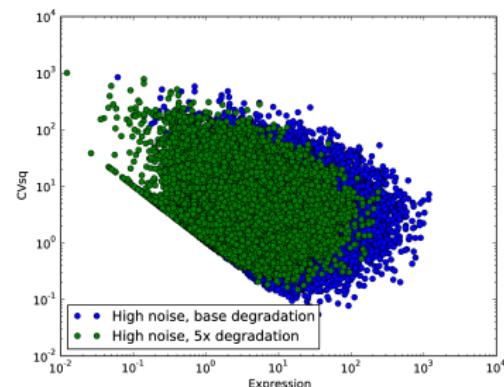


Figure: Calculated gene expression noise in function of expression level, high noise.

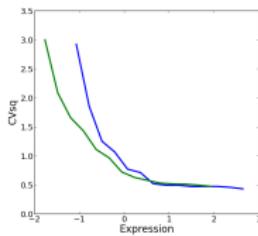
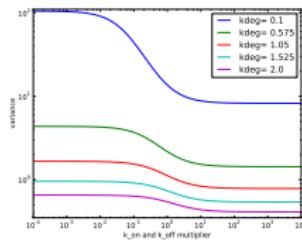
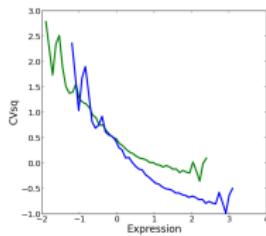


Figure: Response of variance

- Gene expression was simulated with Gillespie algorithm, using StochPy library
- For certain time period each event may happen in random time, corresponding with it's rate.

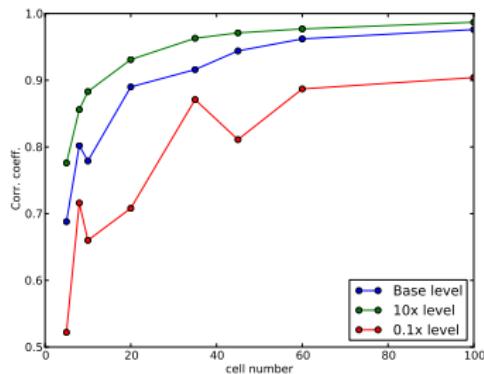


Figure: Cell number optimization.

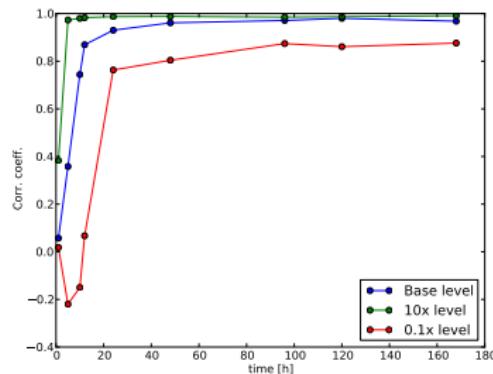


Figure: Simulation time optimization.

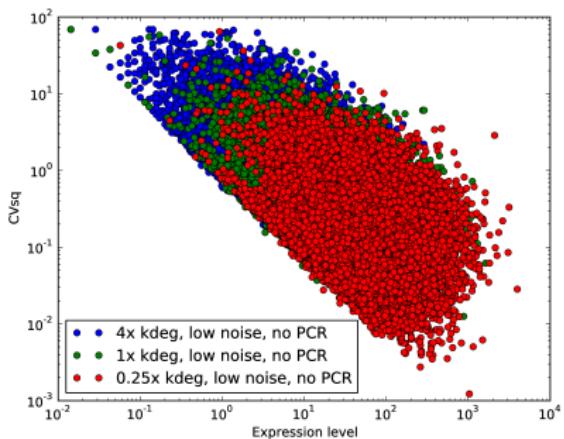


Figure: Expression for three levels (fold 4) of degradation rate for low noise level.

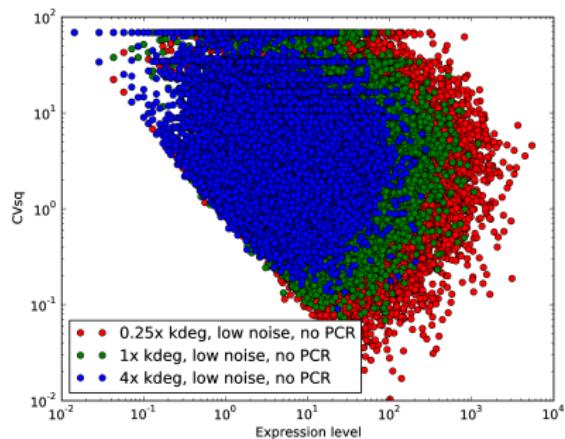


Figure: Expression for three levels (fold 4) of degradation rate for high noise level.

- For low noise level no change in noise visible
- At high noise changes visible
- Rather change in noise range, than noise level - down limit at $\frac{1}{\text{expression}}$

- Fragmentation, filtering etc. not simulated
- We simulated PCR derived noise with three parameters:
 - p - PCR efficiency
 - n - PCR cycles number
 - P - probability of amplifying certain molecule
- Our noise was compared to noise of sequencing 50pg sample
- Lower noise parameters were chosen to prevent loss of less abundant mRNAs

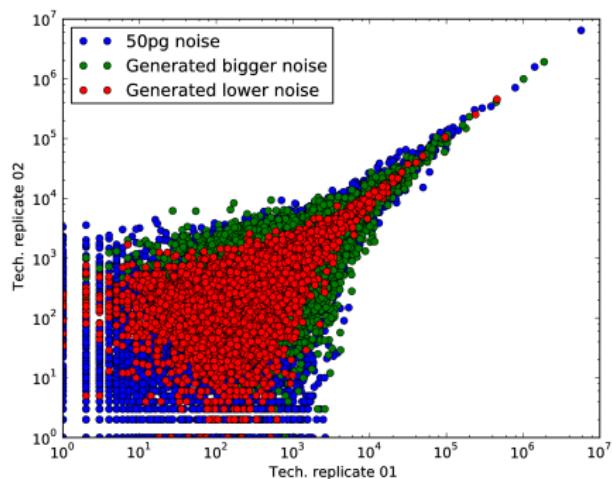


Figure: Two levels of generated PCR noise compared to 50pg noise.

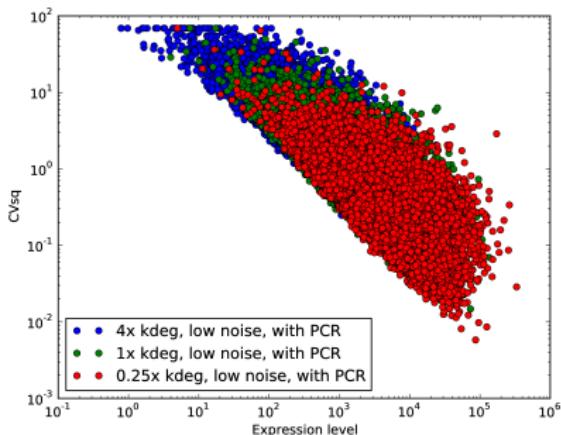


Figure: PCR noise for three levels (fold 4) of degradation rate for low noise level.

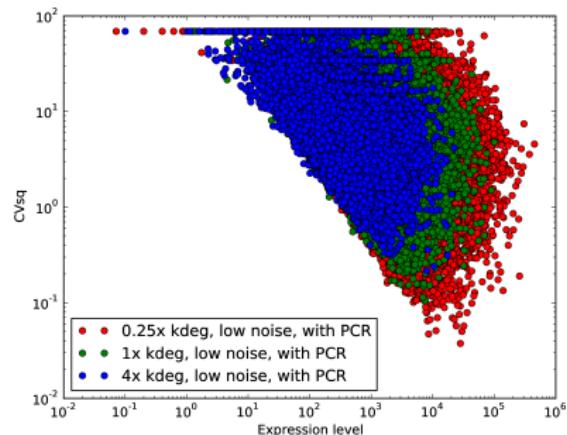
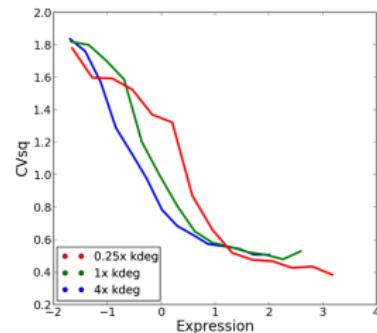
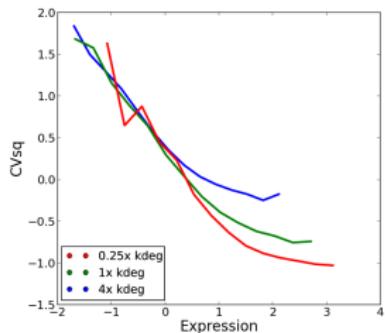
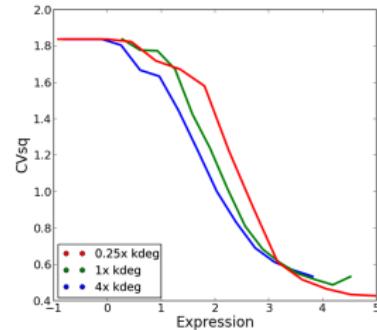
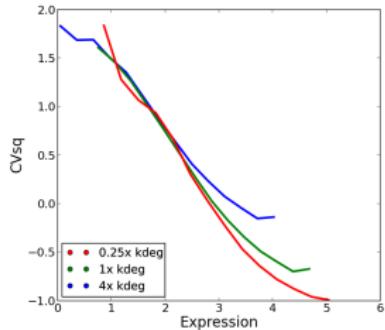


Figure: PCR noise for three levels (fold 4) of degradation rate for high noise level.

- Adding PCR step makes noise level higher, also reducing noise range
- No visible change in noise level for “real” noise level.



- Median CV_{sq} for low (left) and high (right) noise simulations expression.



- Median CV_{sq} for low (left) and high (right) noise simulations sequencing.

- Generated *in silico* cells expression profiles
- Simulated PCR noise
- Improve parameters. Look for appropriate kon/koff level when wet-lab data will be ready.
- Check if there is correlation between parameters.
- Improve PCR noise to represent more low expressed genes.
- Match *in silico* parameters and expression levels with real genes.