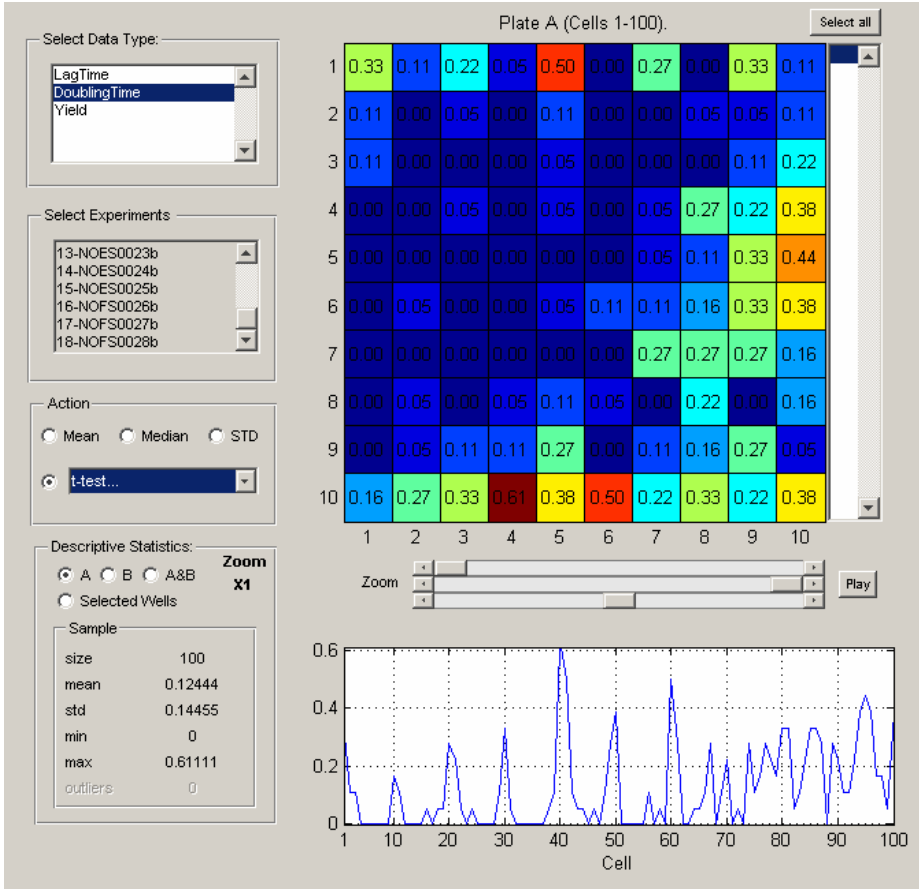
Ph.D.-project Dimitrii Zholud (adv. ON, HR)



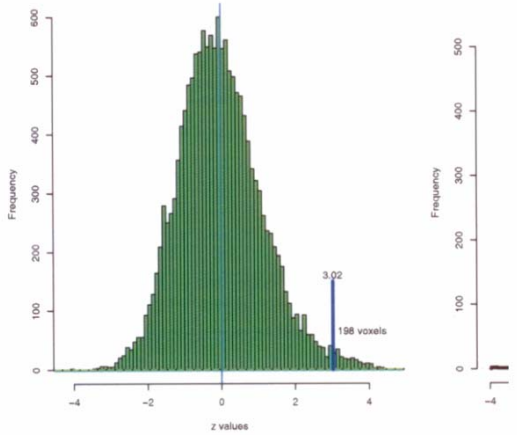
- BioScreen C Analyzer, yeast cells
- wildtype genes
- "doubling time" – rate of growth
- normalization by subtracting mean of 4 "reference wells" on plate
- 1 df t-test between plate A and B
- 5% nominal significance level
- proportion of "significant" t-values over 18 replicates
- 12.4% of t-values "significant"

How correct for spatial effects?

How handle non-normality?

How handle non-normality?

"Efron" -- transform t-values to N(0, 1) with theoretical transform, fit N (μ, σ) to these to find cut-off value, look at histogram:



"tail" problem → use methods from extreme value theory?

- tail plots for t-distribution in experiment: straight lines – but not the t-lines

- theory: suppose T(f) is t-statistic computed from arbitrary finite mean, cont. differentiable, distribution. Then $P(\mathrm{T}(\mathrm{f}) > u) \sim cu^{-f}$ with *c* depending on the underlying distribution

- estimate *c*, use EV methods to check model?
- can this be done if one doesn't have access to wildtype experiment?
- theory for F-statistics, modified t-statistics
- non-asymptotic results? use GPD? conservative – non-conservative
- dependence?