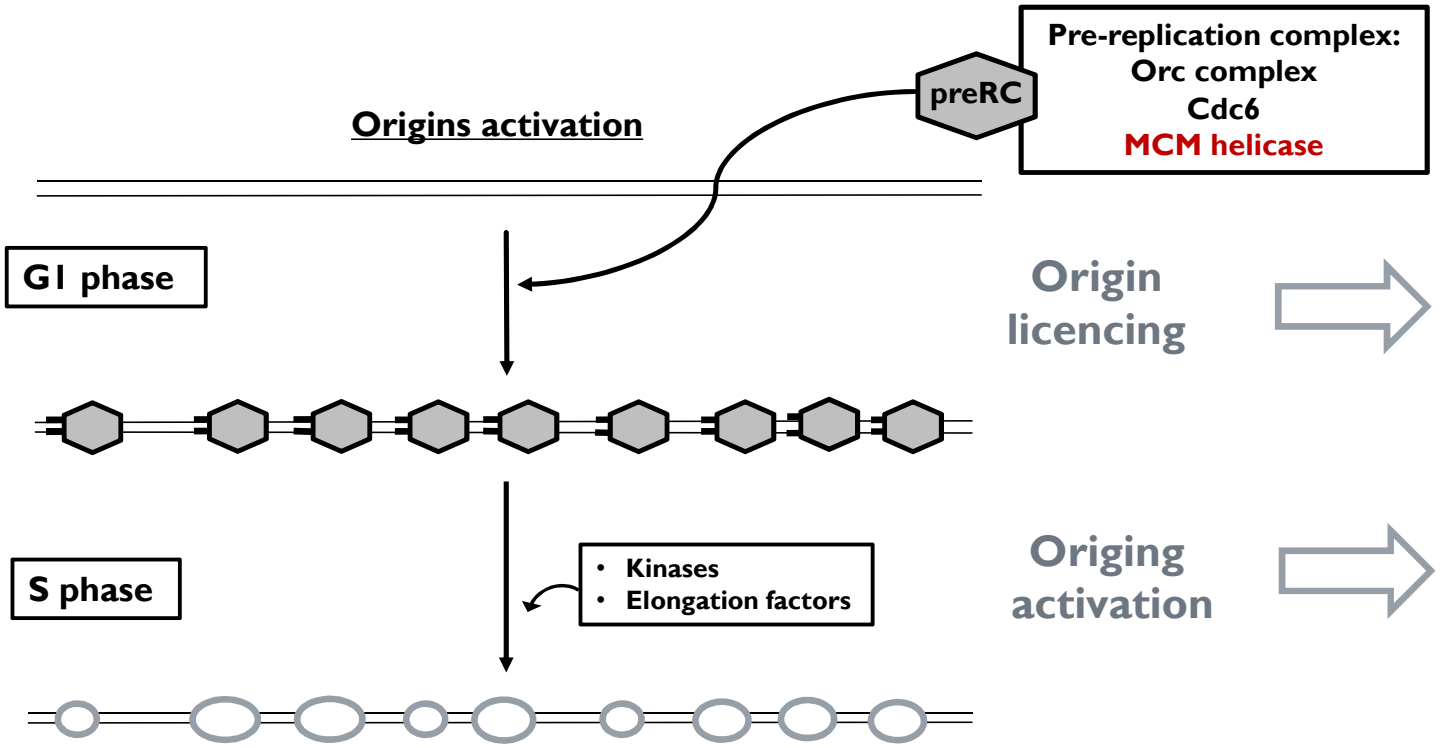
A NUMERICAL MODEL FOR THE SPATIO-TEMPORAL PROGRAM OF DNA REPLICATION IN *XENOPUS* EARLY EMBRYOS

DILETTA CIARDO

SUPERVISORS: KATHRIN MARHEINEKE , ARACH GOLDAR

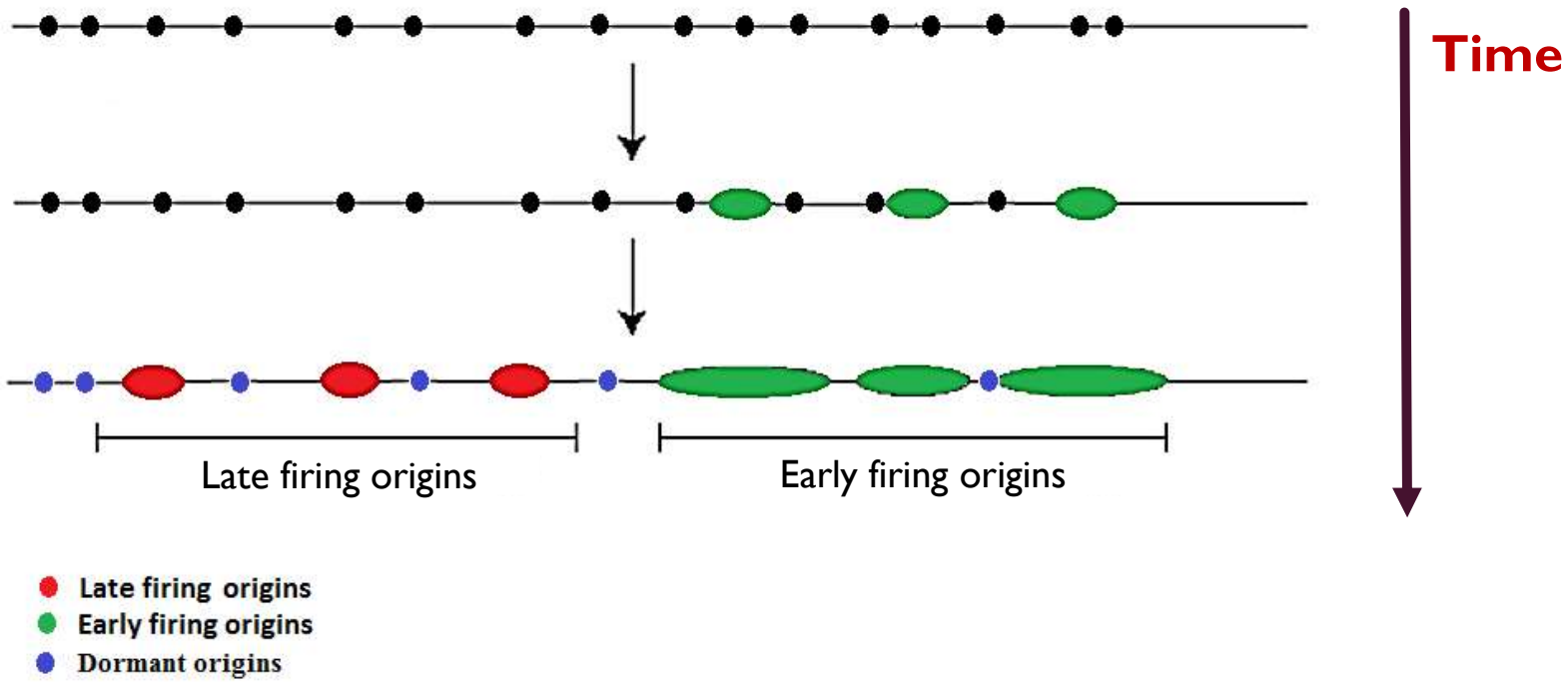
Department of Genome Biology,
Institute for Integrative Biology of the Cell (I2BC)
CNRS, CEA, Paris South University, Gif sur Yvette, France

Initiation of DNA Replication origins in eukaryotes:



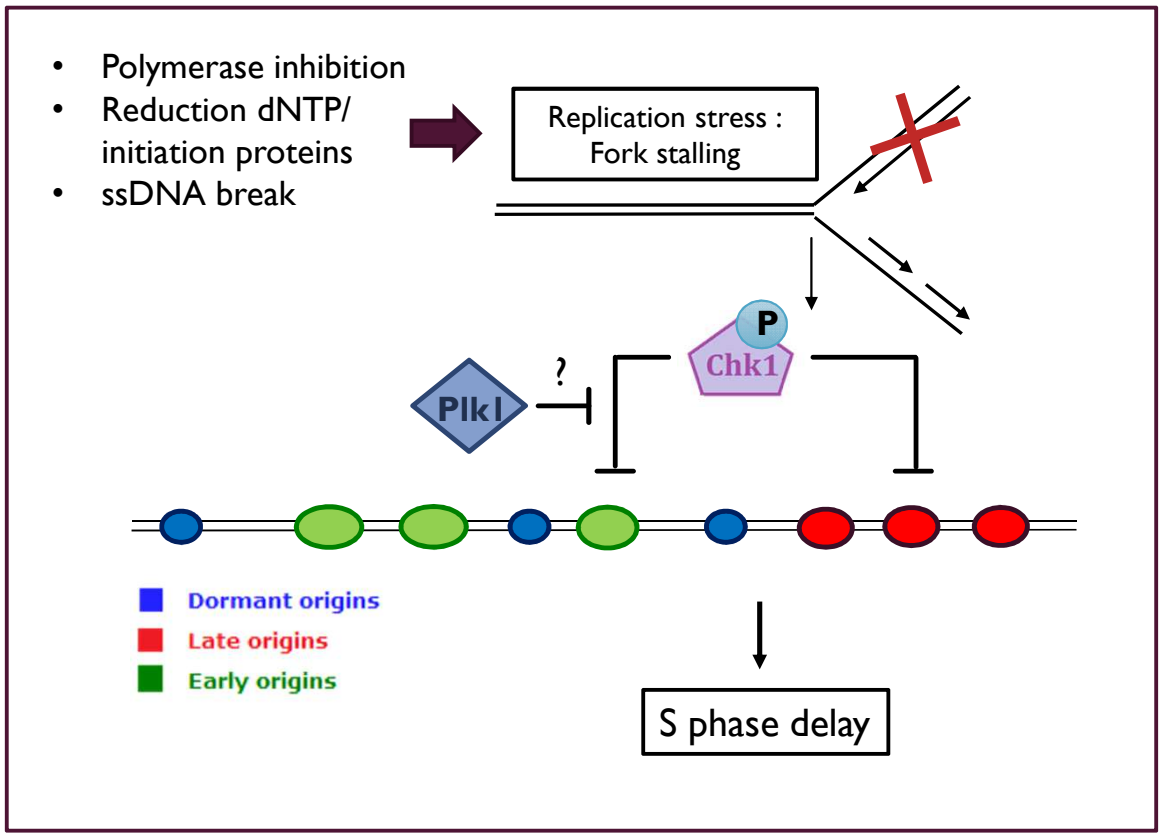
- **Thousands of origins**
- **No clear consensus sequence for pre-RC binding**

Spatio-temporal program of replication:



- Origins grouped in clusters
- Different times of firing

Intra-S phase checkpoint and the spatio-temporal program of replication:



- Activated in response to stalled forks
- Inhibits activation of late origins and delay mitosis entry.
- Hypothesis: Plk1 inhibits Chk1 action also in absence of exogenus replication stress?

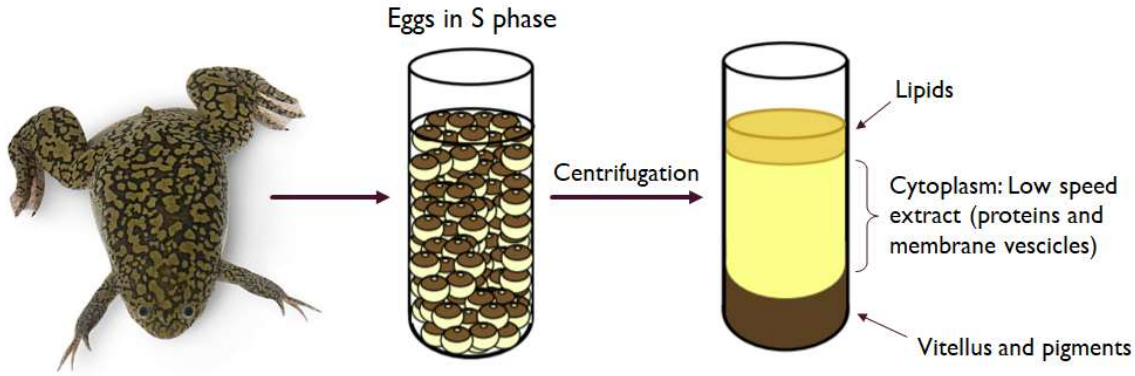
Marheineke &Hyrien (2004)

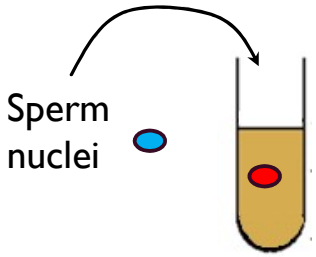
Trenz et al. (2008)

Platel et al. (2015)

Experimental system:

Xenopus laevis in vitro replication system





Stop the reaction at different times during S phase

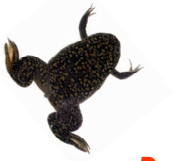
Purify nuclear or chromatin bound proteins for Gel Electrophoresis/Western Blotting

Purify DNA for DNA combing or radioactivity incorporation

- Mimics the S phase in early Xenopus development
- 100 % replication
- Synchronous S phase entry
- Initiation is sequence independent
- Random initiation
- No trascription

Visualisation of replication origins by DNA combing in the Xenopus in vitro system:

1. DNA labelling *in vitro*

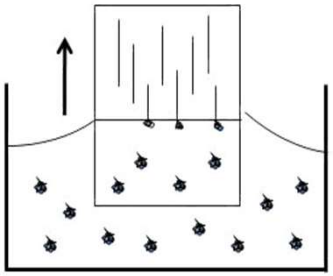


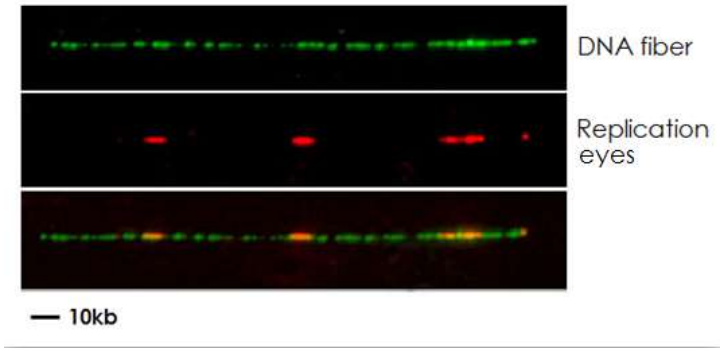
S phase Xenopus egg extract
+ sperm nuclei

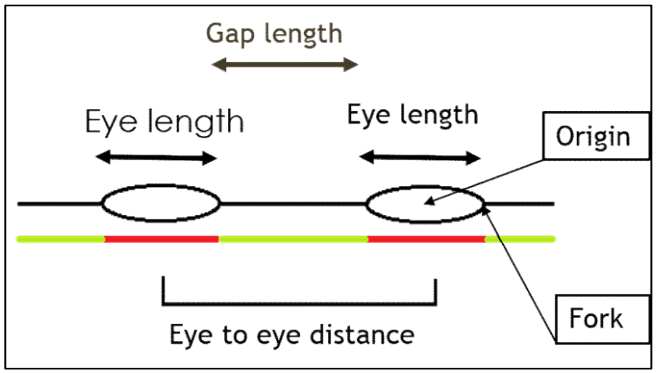
Pulse of Biotin-dUTP → Replication

Newly synthetized DNA is labelled

2. DNA fiber stretching







Replication parameters:

Replication fraction:

$$f(t) = \frac{\sum_i l_i}{l_{TOT}}$$

Fork density:

$$N_f(t) = \frac{N_{fork}}{l_{fiber}}$$

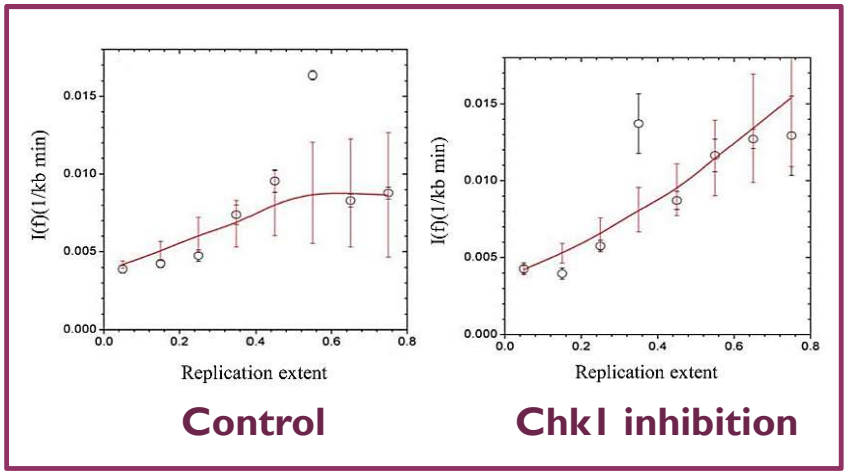
Frequency of initiation:

$$I(t) = \frac{N_{origins}}{\left(1 - f(t)\right) * l_{fiber} * \Delta t}$$

Analysis of temporal program of replication in Xenopus early embryos:

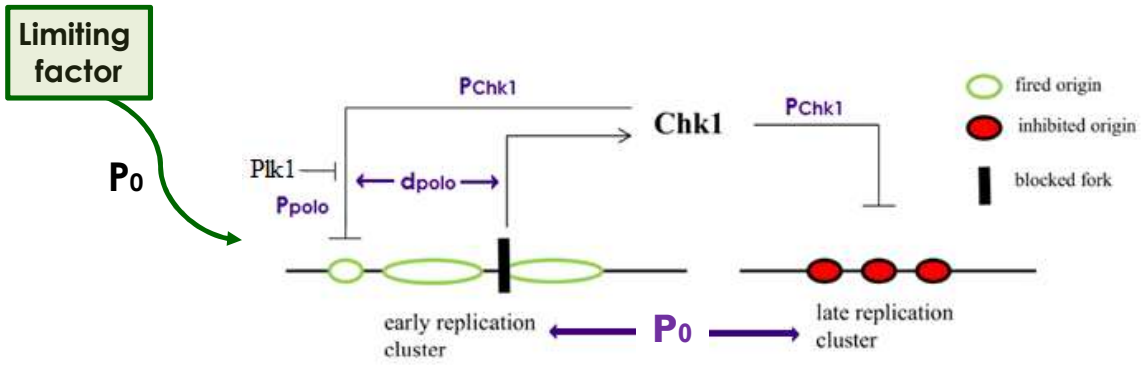
Temporal program:

(Monte Carlo numerical simulations with simplex optimization algorithm)



Platel et al. (2015)

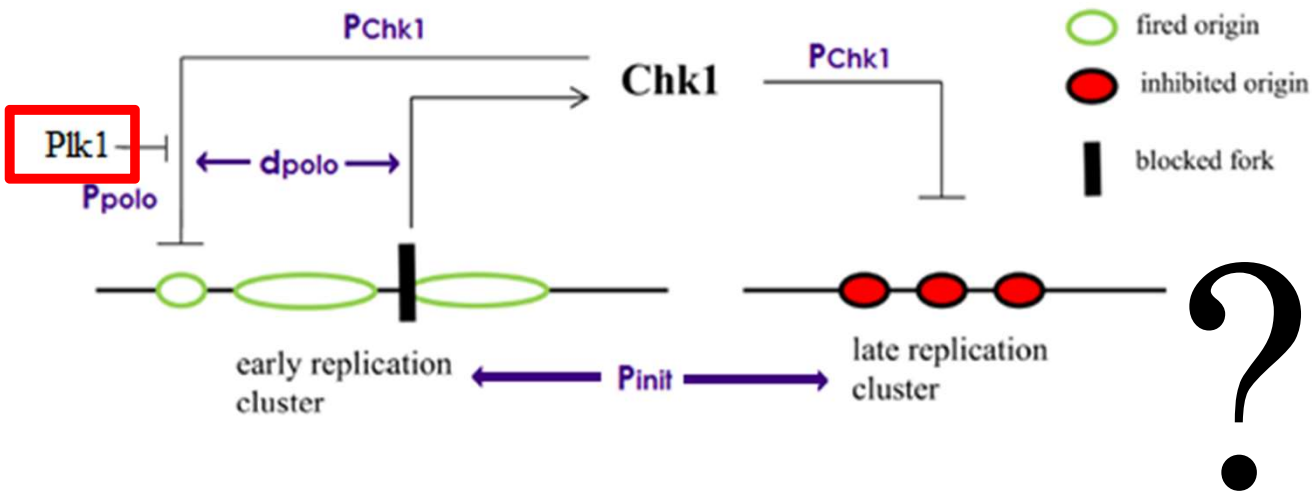
Result:
Inhibition of late firing origins in late clusters, but not in already active clusters



P_{polo} : Probability of action of Plk1
P_{chk1} : Probability of action of Chk1
P_{init} : Probability of action of limiting factor
d_{polo} : Distance of action of Plk1

Number of Chk1 and Plk1 exactly equal to the number of forks

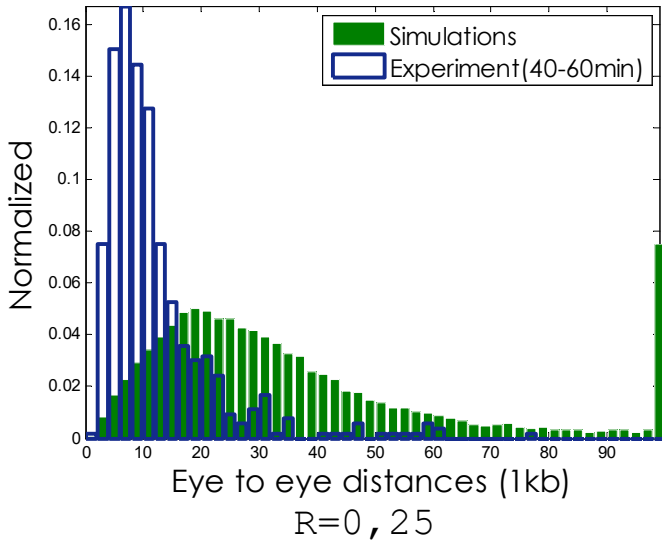
Objectives:



- Understand if the model can reproduce the spatial program of DNA replication in *Xenopus* early embryos
  - Comparison of numerical and experimental eye-to-eye distance (ETED) distributions

Previous model does not reproduce the special program of replication:

Extraction of the ETED distribution using the previous model (Monte Carlo simulations):



Parameters:

N_0: 1	d_{polo}: 45000
J: 4/120	N_{Chk1} : 1
P_0: 0.01	$\%_{Chk1}$: 1
P_{Chk1} : 0.99	P_{Polo} : 0.01

- Old parameters
- Old scenario
- New code

Analysis of the **spatial** and temporal replication program in *Xenopus early embryos*:

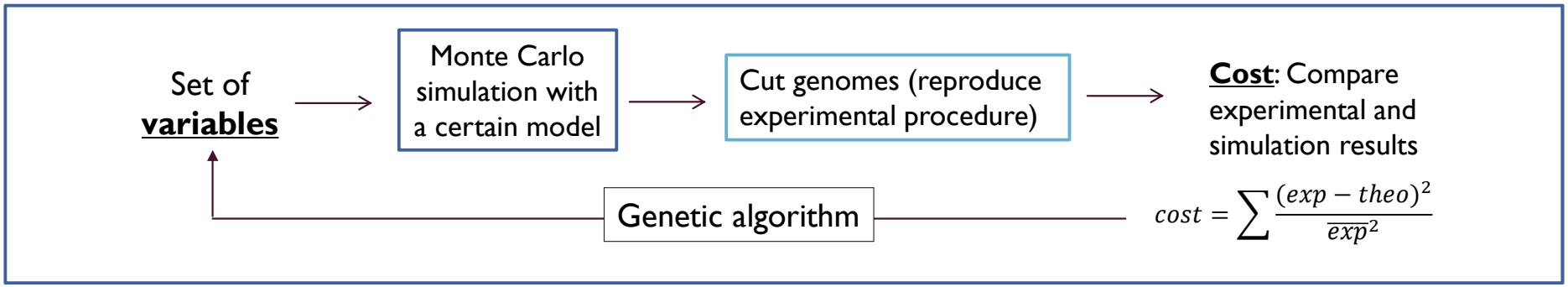
- Consider both spatial and temporal program (ETED considered)
- Length of simulated genome comparable to real length
- Experimental procedure reproduced in simulations
- Time points considered separately

Strategy:

OBJECTIVE: Find a model that describes the experimental data in presence or absence of Chk1 inhibition

HOW? Suppose a model and optimize the fit of replication parameters

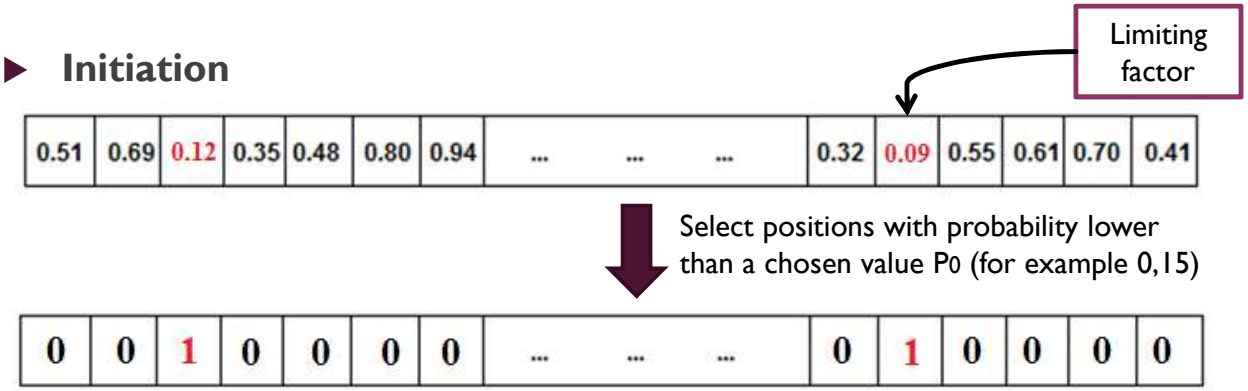
ALGORITHM:



RESULT: Family of sets of variables of a given model that better describe the experimental data

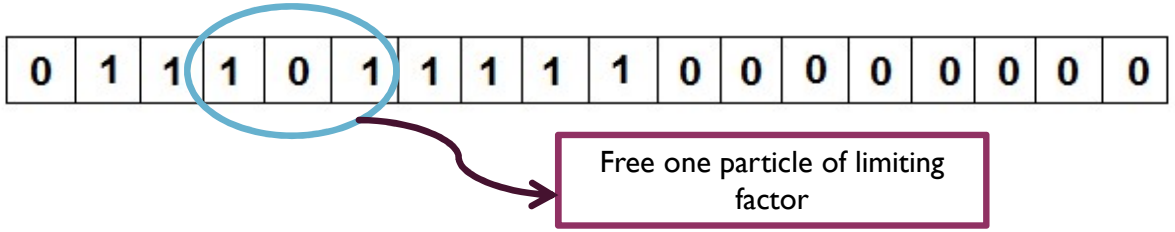
Dynamic Monte Carlo:

- **Initiation**



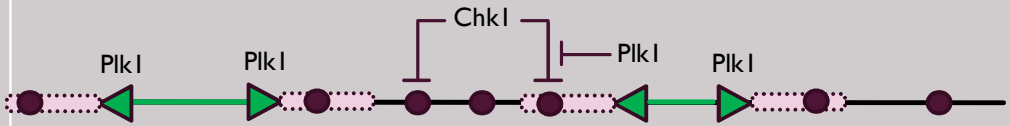
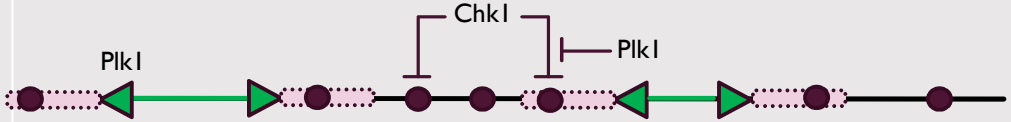
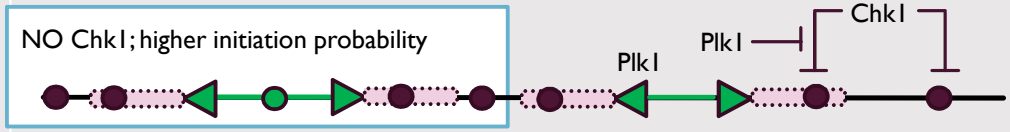
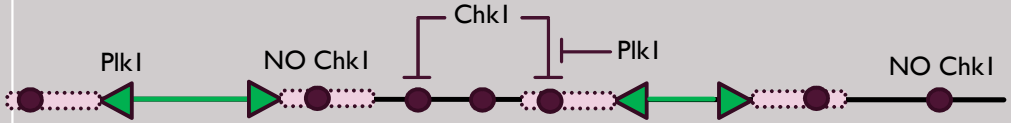
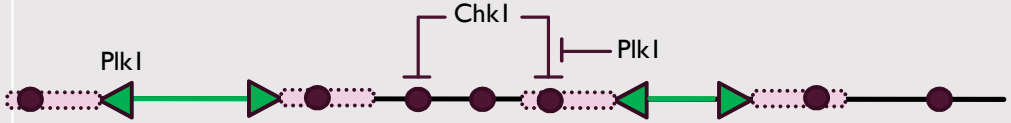
- **Elongation**

- **Termination/merger**



Goldar et al. (2008)

Optimization of different models with genetic algorithm:

Number of variables (k)	Model	Lowest score (cost/(N-k-1))
9		0,1
10	 Forks with action of Plk1 chosen randomlly at each step	0,14
10	Same as above, but action of Plk1 remain on forks until they merges	0,13
10		0,04
9		0.16
8		0,14

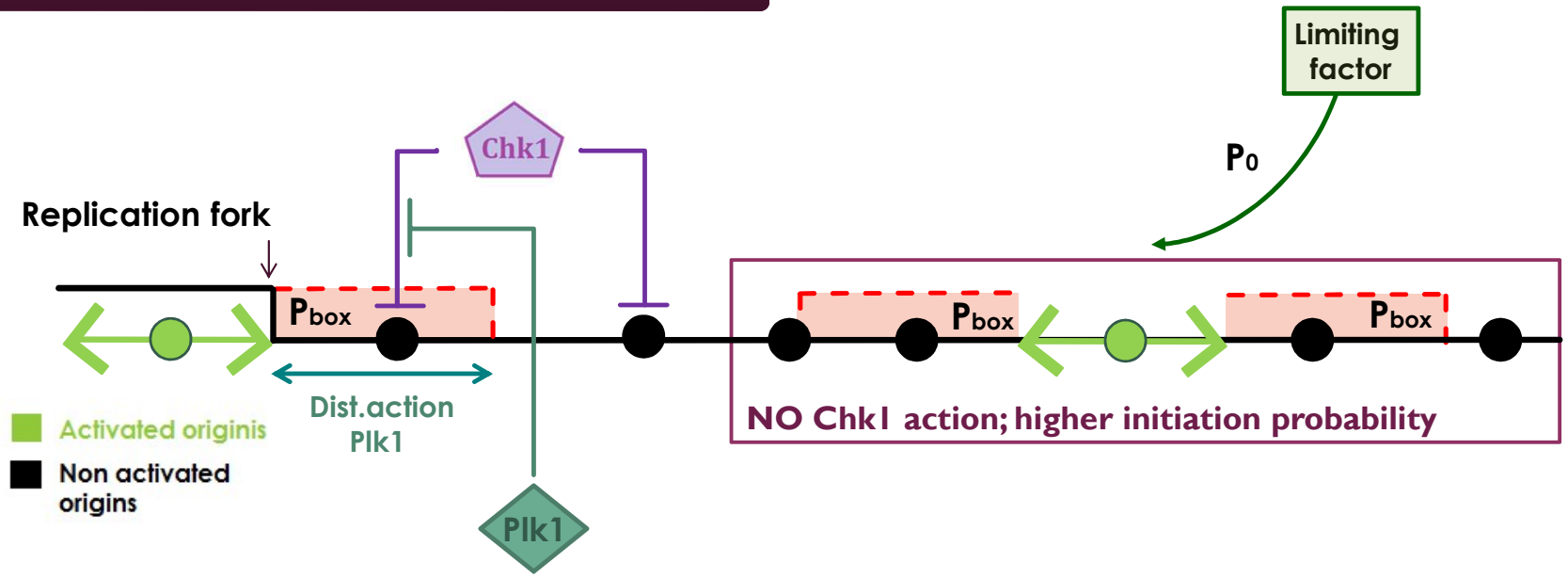
Legend:

- Origin
- Fork
- Action Plk1 and Pbox
- N=87 Num. of fitted points
- K Num. of variables

Not limited number of potential origins

Limited number of potential origins

Elements of the model:

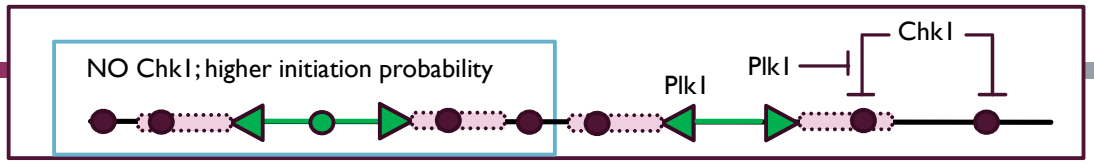


Parameters:

N0: initial number of limiting factor
J: Rate of import of limiting factor (s^{-1})
P0: Initial probability of initiation
PChk1: Probability of inhibition by Chk1
Pbox: Increased prob. close to replication forks

dbox: Distance of action of Pbox and Plk1 (kb)
N.regions: Number of regions
Dregions: Half of the length of regions (kb)
PPolo : Prob. of action of Plk1

Results from best model (Absence of Chk1 inhibition):



Analysis of early S phase:

Parameters:

N0: initial number of limiting factor

J: Rate of import of limiting factor (s^{-1})

P0: Initial probability of initiation

PChk1: Probability of inhibition by Chk1

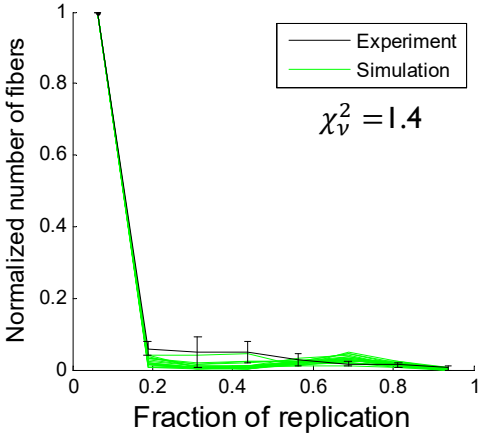
Pbox: Increased prob. close to replication forks

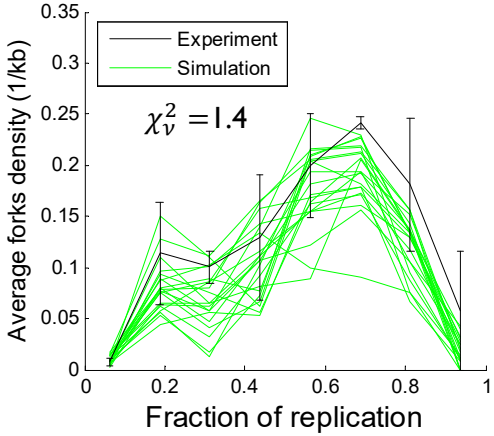
dbox: Distace of action of Pbox and Plk1 (kb)

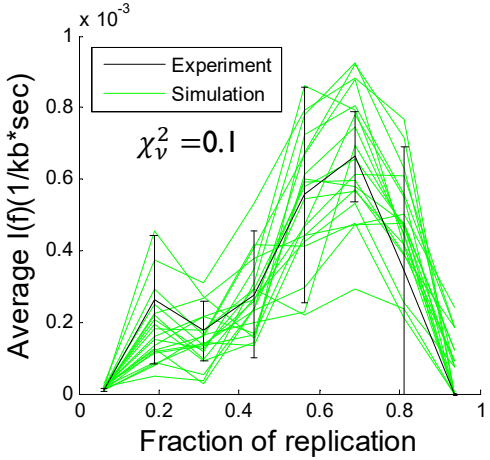
N.regions: Number of regions

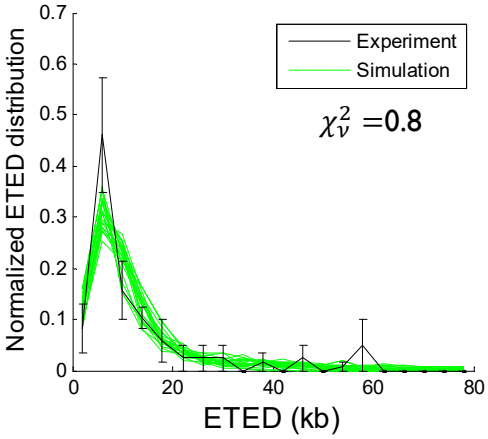
Dregions: Half of the length of regions (kb)

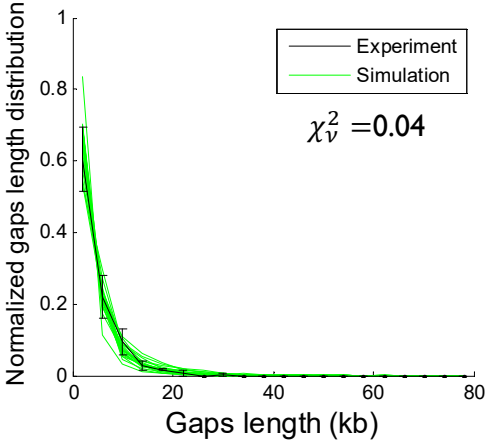
PPolo : Prob. of action of Plk1

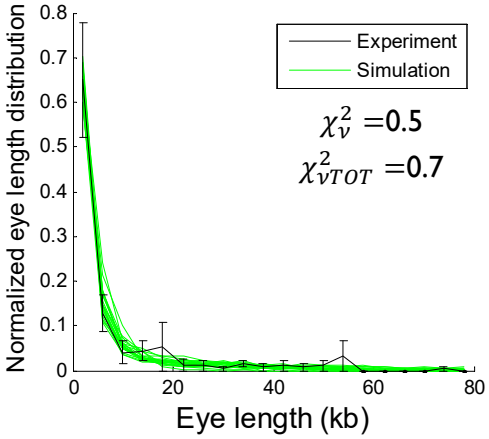




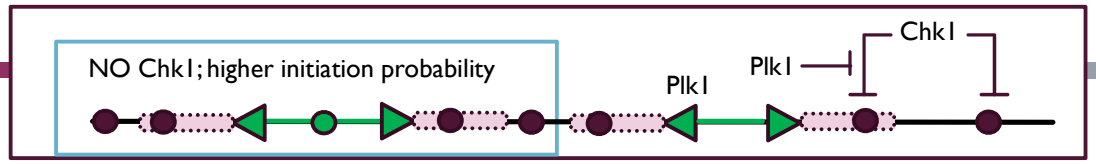








Results from best model (Chk1 inhibition):



Analysis of early S phase:

Parameters:

N0: initial number of limiting factor

J: Rate of import of limiting factor (s^{-1})

P0: Initial probability of initiation

PChk1: Probability of inhibition by Chk1

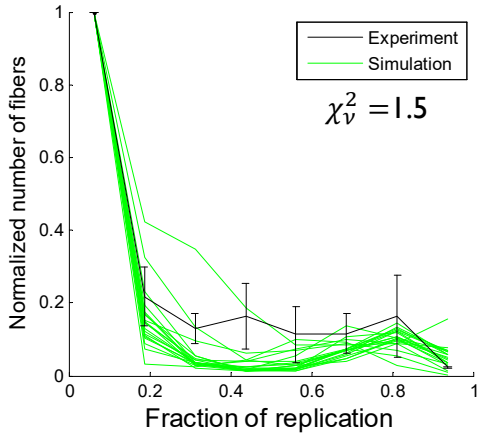
Pbox: Increased prob. close to replication forks

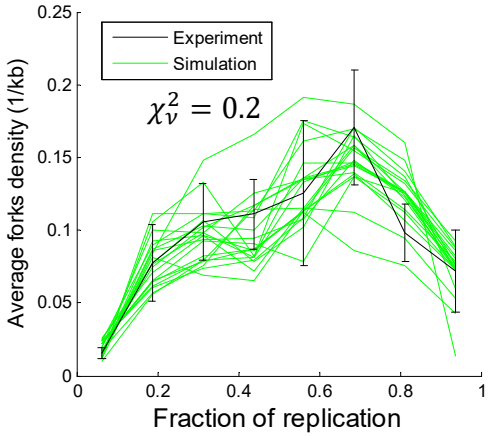
dbox: Distace of action of Pbox and Plk1 (kb)

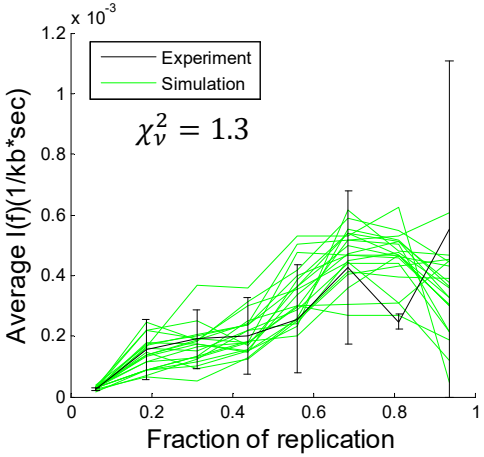
N.regions: Number of regions

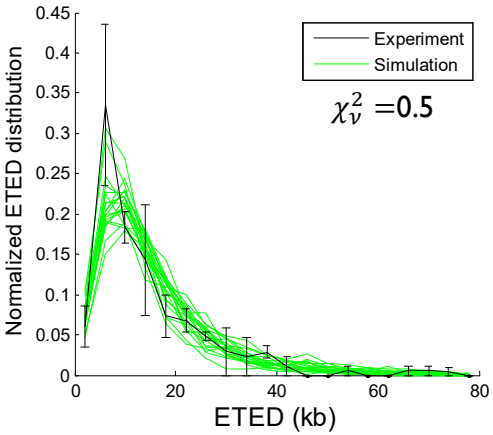
Dregions: Half of the length of regions (kb)

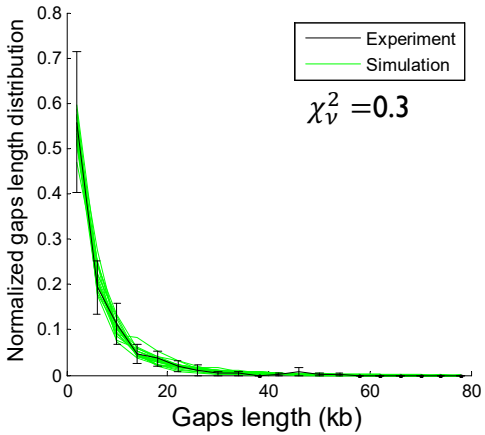
PPolo : Prob. of action of Plk1

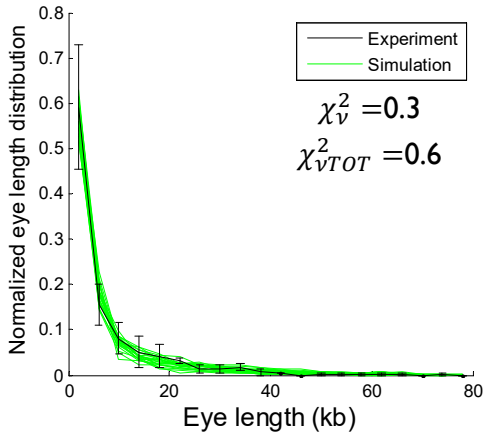






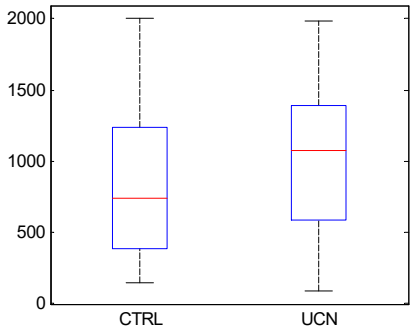




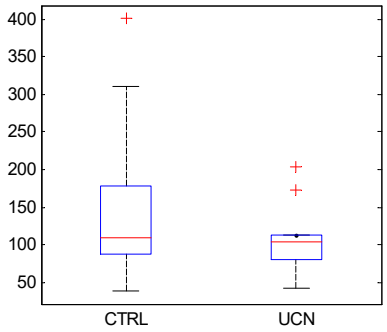


Comparison family of sets in absence and presence of Chk1 inhibition:

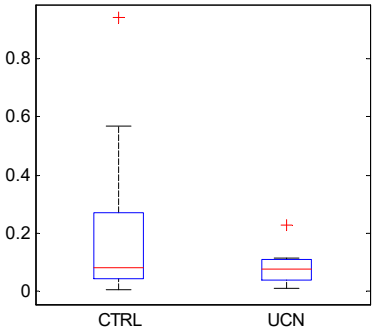
Initial number limiting factors (N)



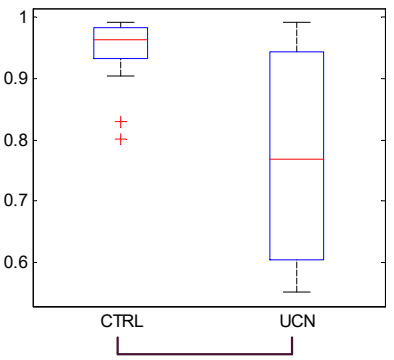
Limiting factors rate (J)



Probability of initiation (P_0)

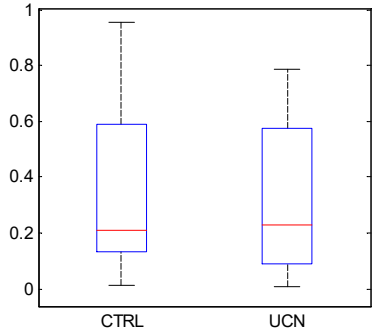


Probability of inhibition (P_{Chk1})

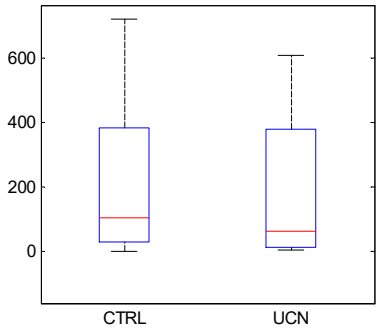


Significant difference:
t-test, α=0.05
Kolmogorov-Smirnov test, α=0.05

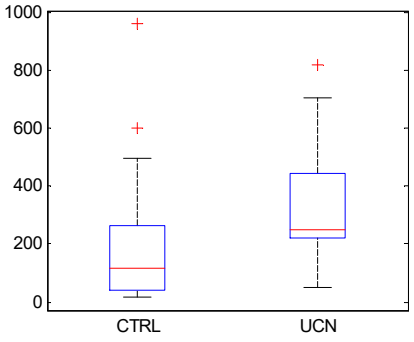
Probability close to forks (P_{box})



Action length of P_{box} and Plk1 (d_{box})

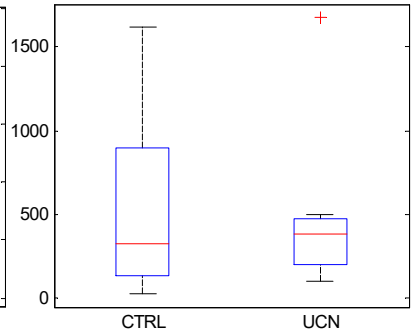


Number of domains (N_{dom})

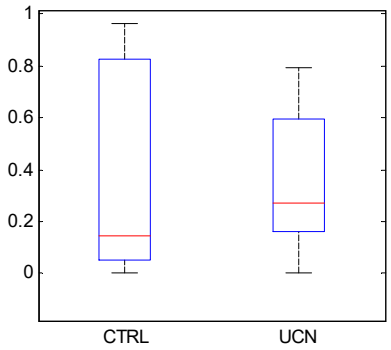


Significant difference:
Kolmogorov-Smirnov test, α=0.05

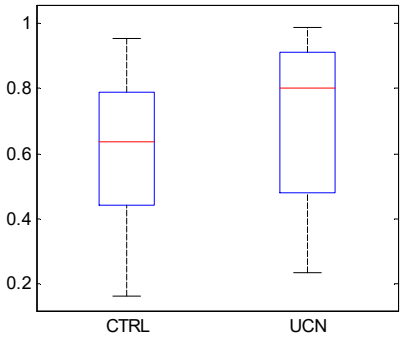
Length of domains (l_{dom})



Probability of Plk1 action (P_{polo})

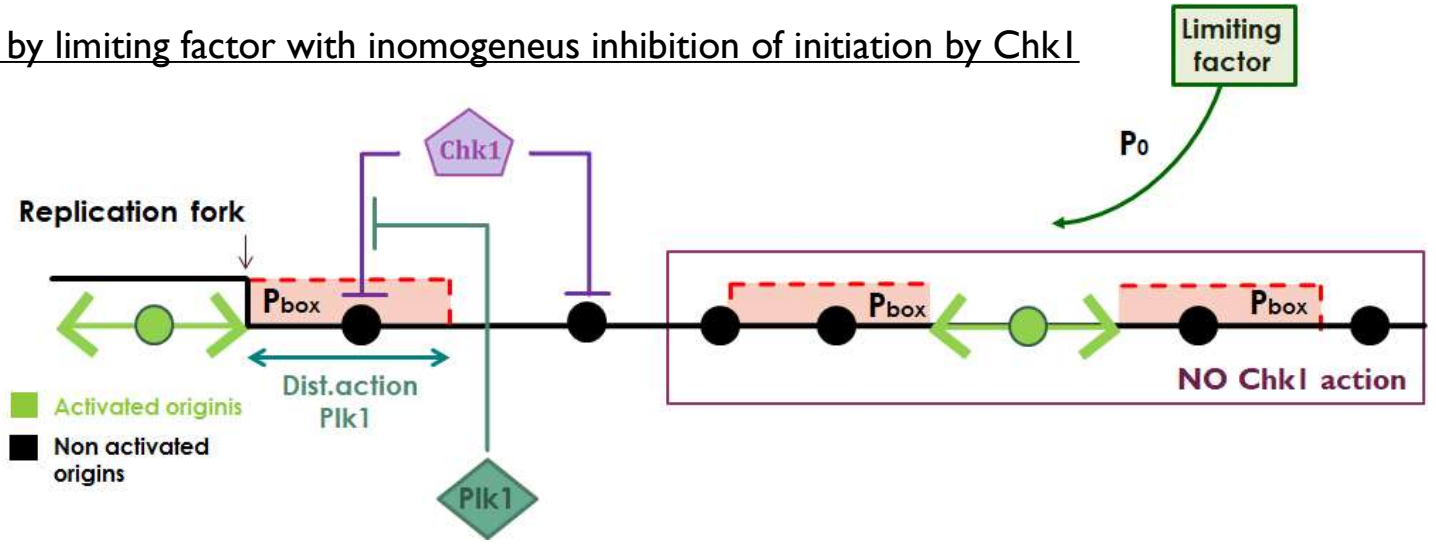


Increased probability inside domains (P_{dom})



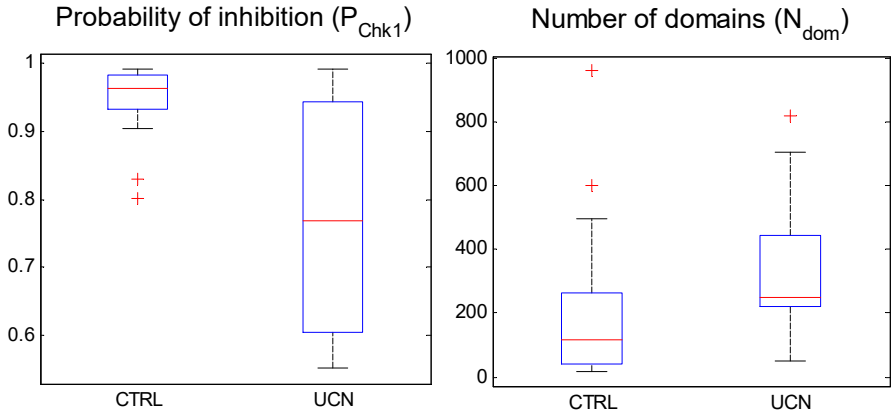
Best model:

Random initiation by limiting factor with inomogeneus inhibition of initiation by Chk1



Conclusions:

- Limiting factor interacting with probability P0 everywhere
- Probability increase close to replication forks (Pbox) (Guilbaud et al. (2011) , Lob et al. (2016))
- Strong inhibition of origin initiation by Chk1
- Inhibition of Chk1 action by Plk1 close to active forks
- Regions without action of Chk1, with high probability of origins initiation (Initiated early in S phase)
- Length of domains:
  $\bar{l}_{CTRL} = 1030 \pm 220kb, \;\; \bar{l}_{UCN} = 911 \pm 201kb$



Acknowledgements:

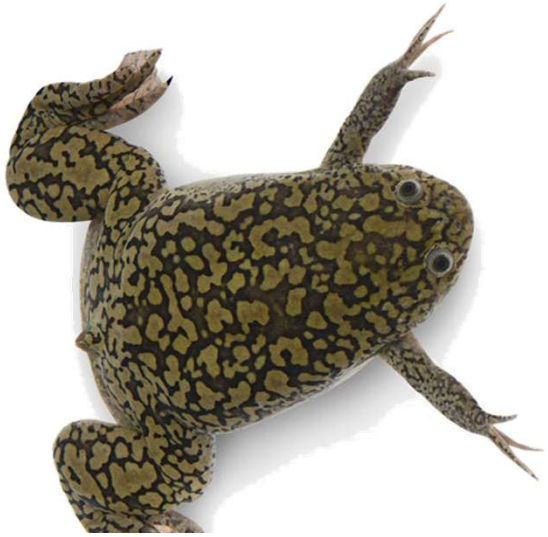
Team Kathrin Marheineke
Group « Dynamique de la Réplication de l'ADN chez les eucaryotes supérieurs »

Équipe:
MARHEINEKE Kathrin
HACCARD Olivier
NARASSIMPRAKASH Hemalatha
BAZIN Melanie (L3 student)

Team Julie Soutourina
Group « Régulation transcriptionnelle des génomes »

Équipe:

SOUTOURINA Julie
GOLDAR Arach
DENBY WILKES Cyril
WERNER Michel
GIORDANENGO-AIACH Nathalie
GOPAUL Diyavarshini



Thank you for Your attention!

Formation and activation of DNA Replication origins in eucaryotes:

A two step process:

1. Origin licencing

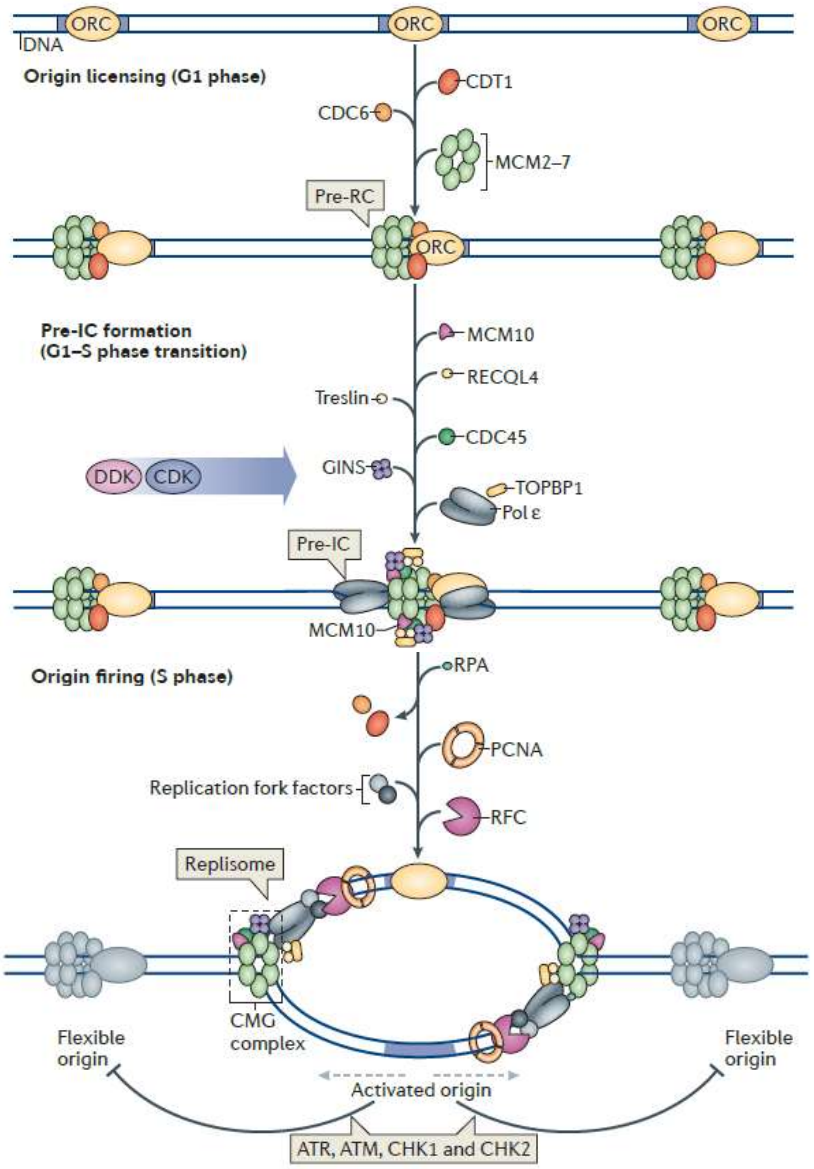
⇩

Loading pre-RC

2. Origing activation

⇙ Formation pre-IC

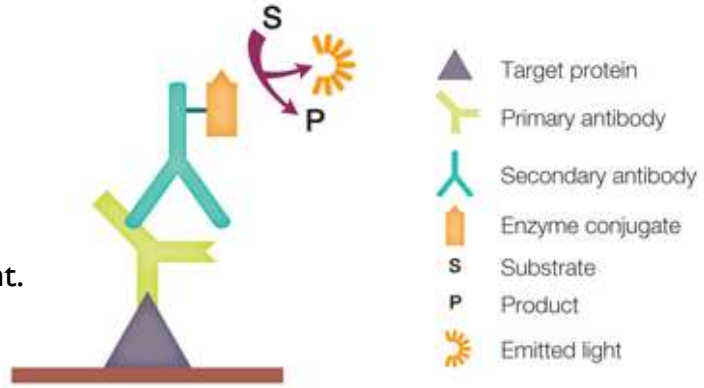
⇘ Activation of helicase

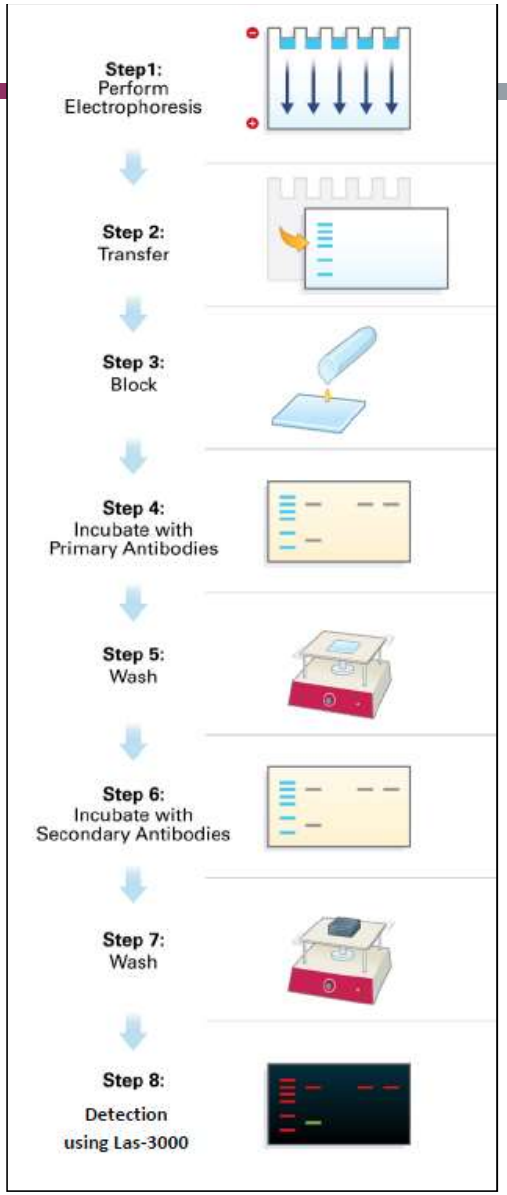


SDS-PAGE (Sodium Dodecyl Sulphate-Polyacrylamide gel electrophoresis) and Western Blot

- The SDS-PAGE gel electrophoresis separate the proteins by size;
- SDS is applied to protein samples to linearize proteins and to impart a negative charge (uniform distribution of charge per unit mass);
- The protein of interest interacts with a specific antibody (primary antibody);
- A second antibody (linked to the horseradish peroxidase) binds to the primary and allows the detection by the use of a Chemiluminescence kit.

The enzyme horseradish peroxidase (HRP) catalyses the oxidation of luminol. The reaction is accompanied by emission of light.





DNA combing technique:

- Replication can be followed by the incorporation of Biotin-dUTP;
- The pH dependent interaction between DNA and the hydrophobic coverslip and the air-solution meniscus allows the stretching of the fibers across the glass surface;
- The biotin labelling is realized by a succession of five incubations alternating between Streptavidin Alexa Fluor 594 and biotinyled anti-Streptavidin antibodies;
- Totality of DNA is labelled with human anti-DNA antibody followed by Alexa Fluor 488 anti-mouse and anti-rabbit .

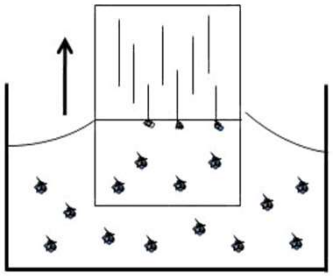
1. DNA labelling *in vitro*

S phase Xenopus egg extract
+ sperm nuclei

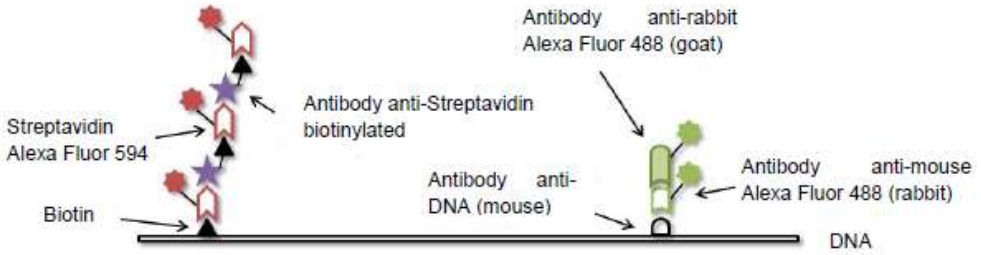
Pulse of Biotin-dUTP → Replication

Newly synthetized DNA is labelled

2. DNA fiber stretching

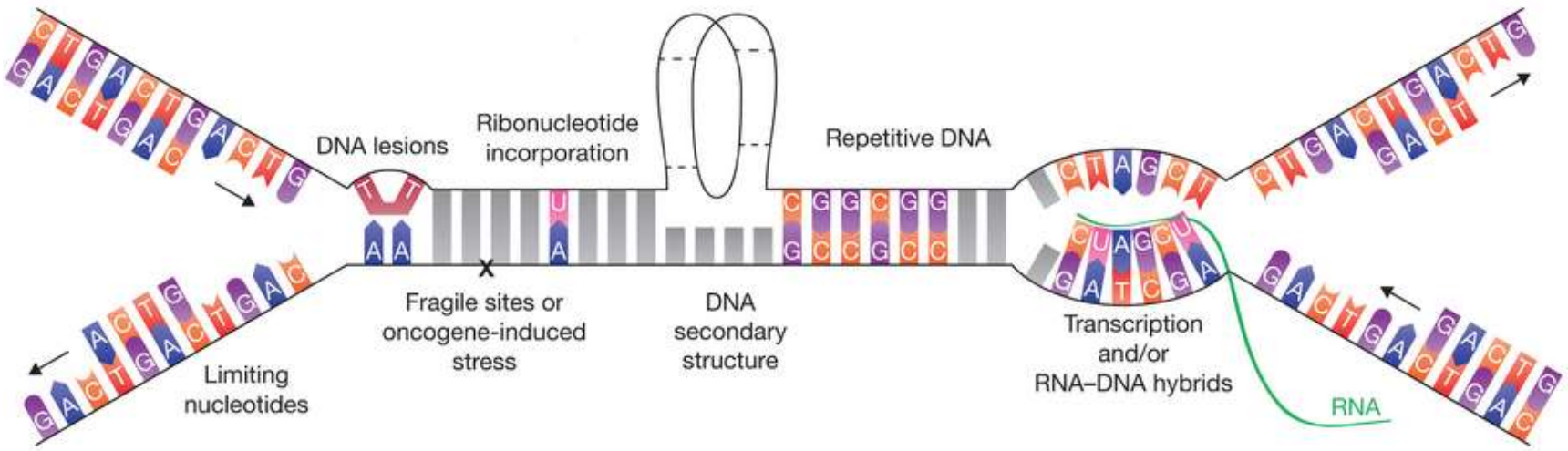


3. DNA labelling

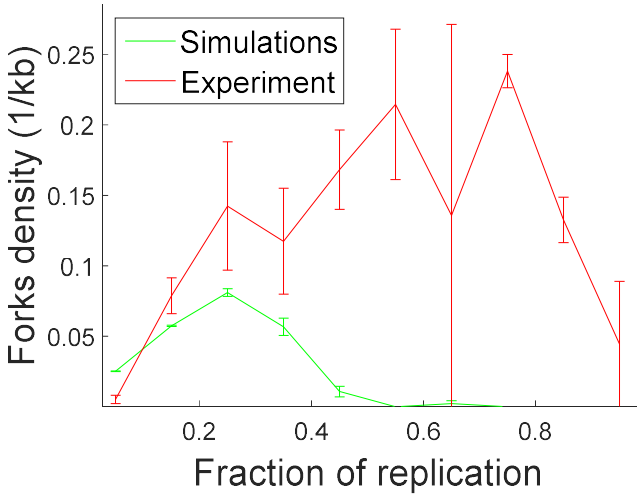


Causes of replication stress:

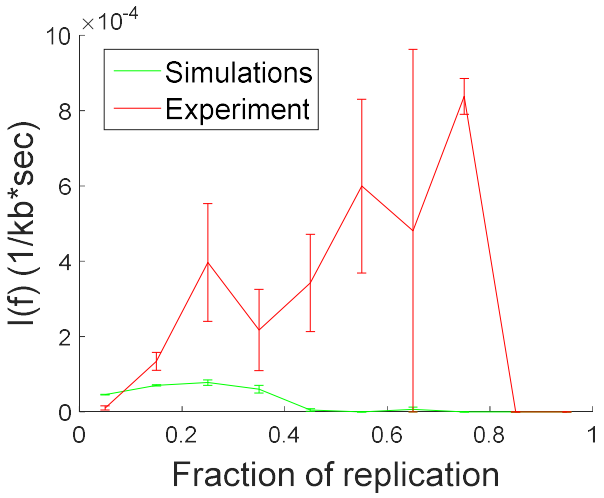
Internal replication stress is generally present in the replication process:



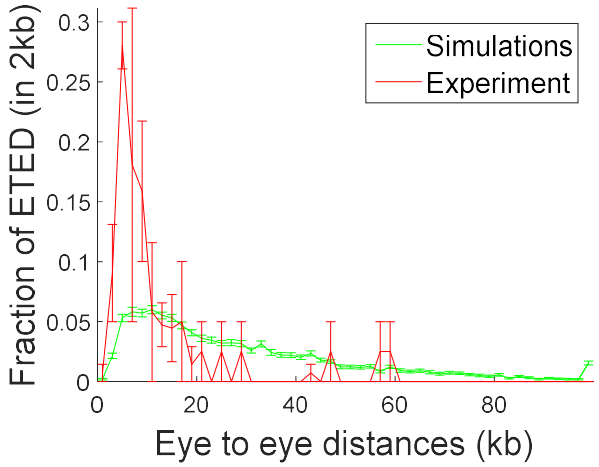
Scenario: Random initiation by limiting factor



R= -0,25

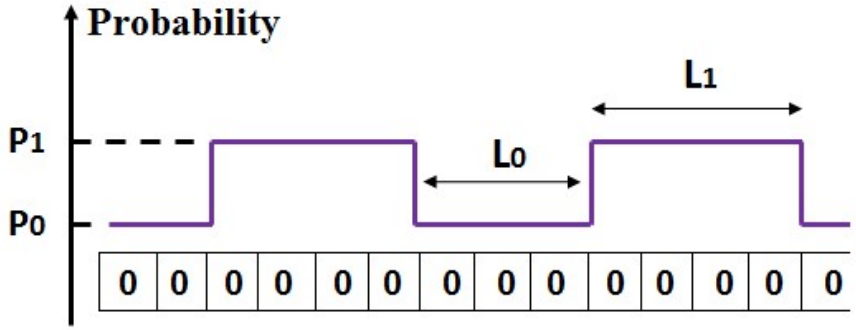


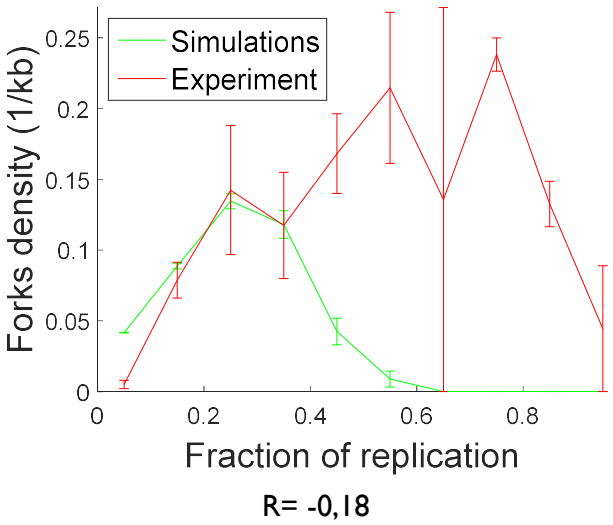
R= -0,26



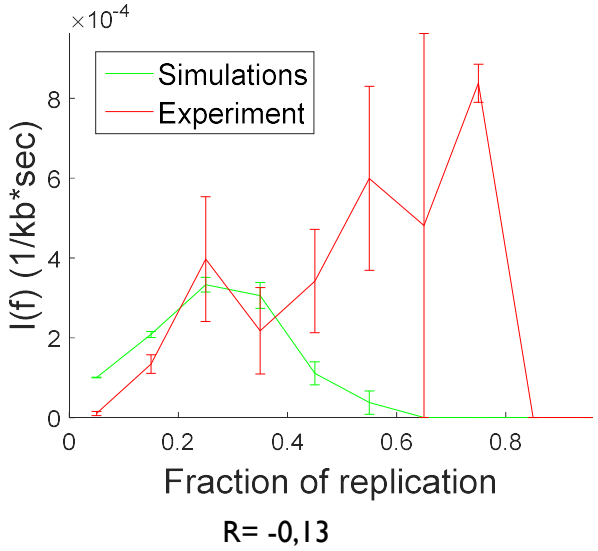
R=0,65

Scenario: Random initiation by limiting factor on preferential genomic loci

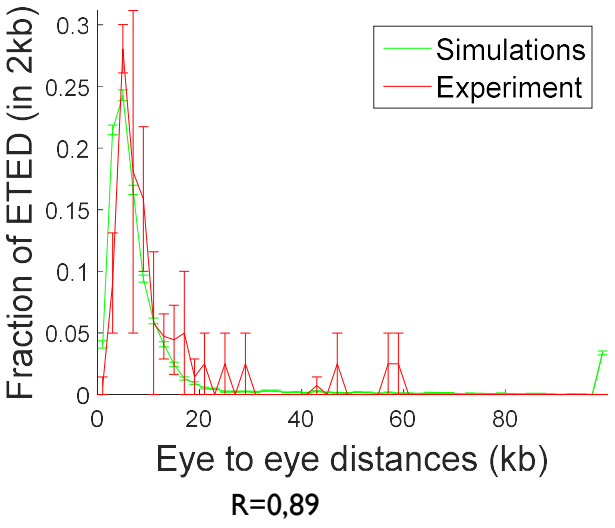




R= -0,18

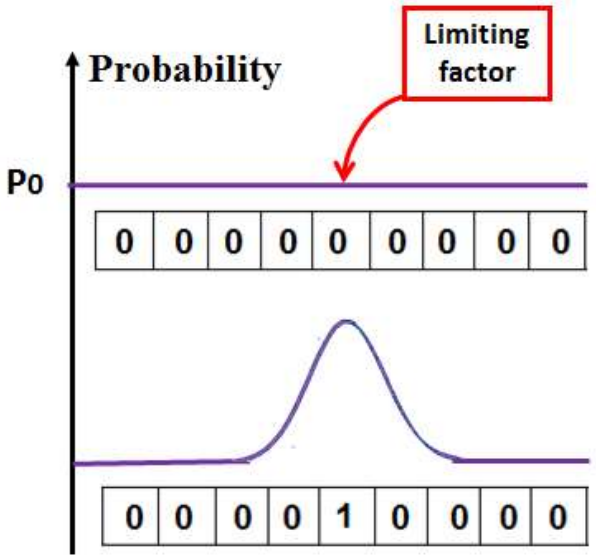


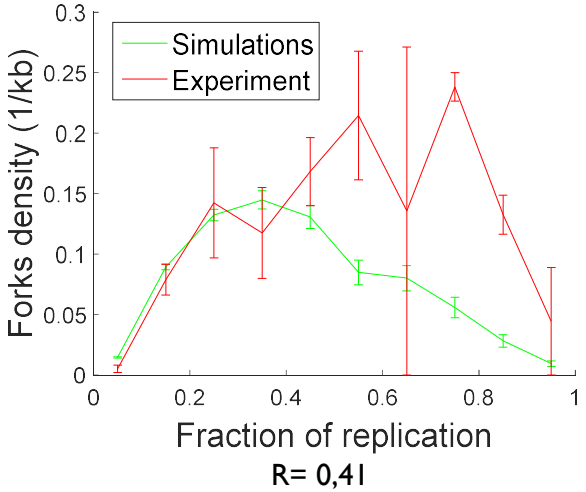
R= -0,13



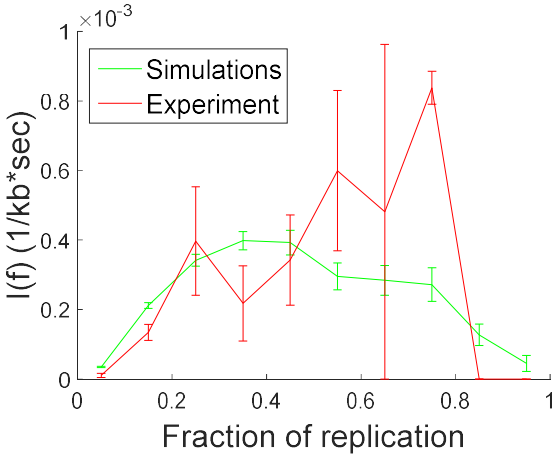
R=0,89

Scenario: Random initiation by limiting factor whose number is enhanced near activated origins

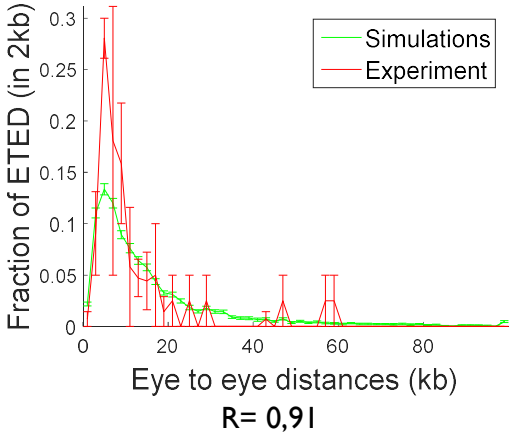




R= 0,41



R= 0,60

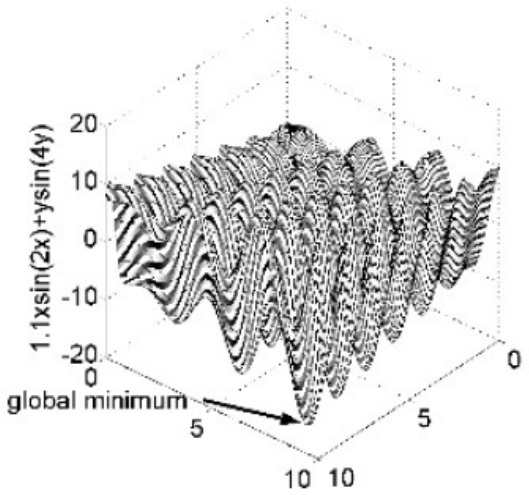


R= 0,91

GENETIC ALGORITHM

It is an optimization and search technique based on the principles of genetic recombination and natural selection.

EXAMPLE:
Optimization of a two variables problem.



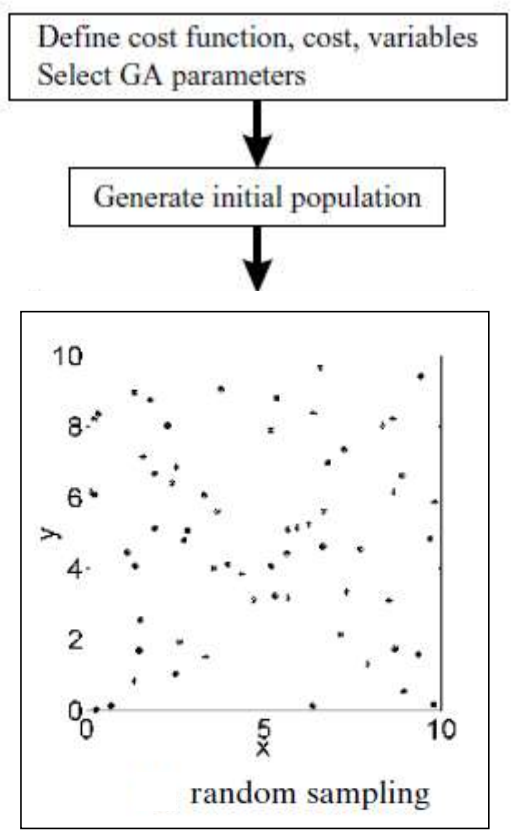
Variables: [x,y]
Cost function: $1{,}1xsen(2x) + ysen(4y)$

OBJECTIVE:
Find the set of variables that minimize the cost.

Set of **variables** → Simulation with a certain model → **Cost**: Compare experimental and simulation results

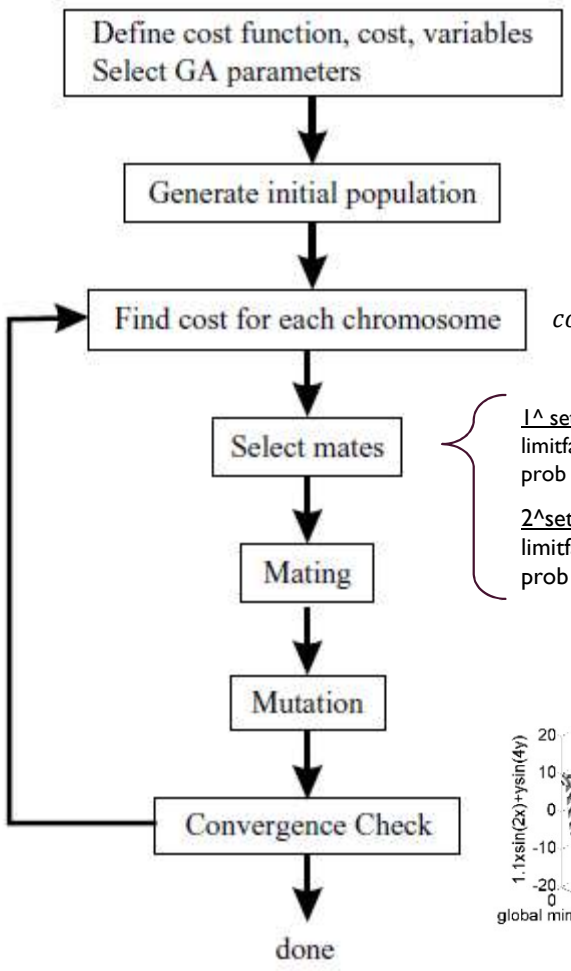
$$cost = \sum \frac{(exp - theo)^2}{\overline{exp}^2}$$

GENETIC ALGORITHM



Set of 11 variables (**chromosome**):
limitfactor max:788375, ratej:2870.327, prob init:0.39965, perc probinit:0.8825, initn chk1:575553,
perc chk1:6.761, prob pchk1:0.046096, prob box:0.88395, dist box:14, prob polo:0.40167, perc polo:0.45698

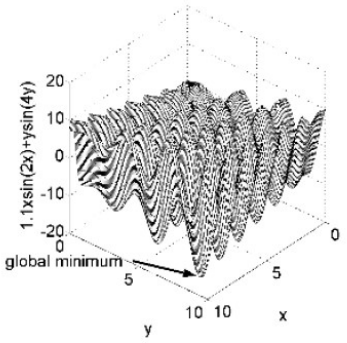
GENETIC ALGORITHM



$$cost = f(chromosome) = \sum \frac{(exp - theo)^2}{\overline{exp}^2}$$

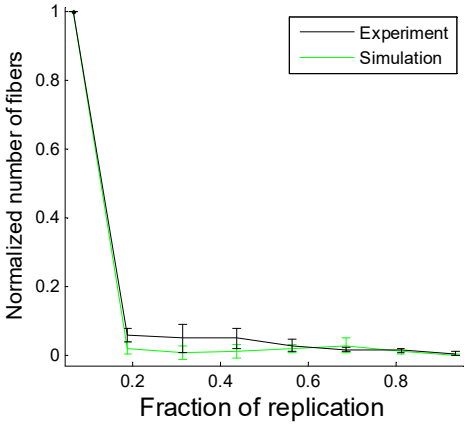
1^ set of variables:
limitfactor max:788375, ratej:2870.327, prob init:0.39965, perc probinit:0.8825, initn chk1:575553, perc chk1:6.761, prob pchk1:0.046096, prob box:0.88395, dist box:14, prob polo:0.40167, perc polo:0.45698

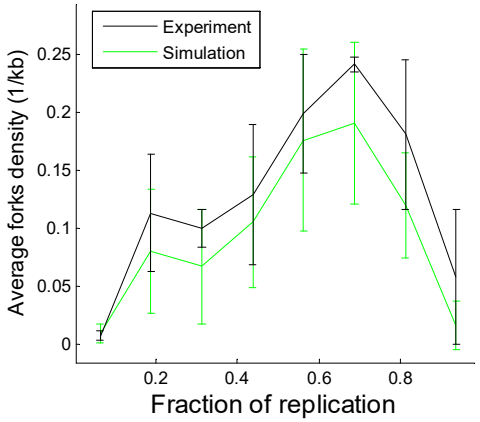
2^set of variables:
limitfactor max:415006, ratej:261.8735, prob init:0.018647, initn chk1:944780, perc chk1:3.0418, prob pchk1:0.99047, prob box:0.84097, dist box:52, prob polo:0.97729, perc polo:0.076676

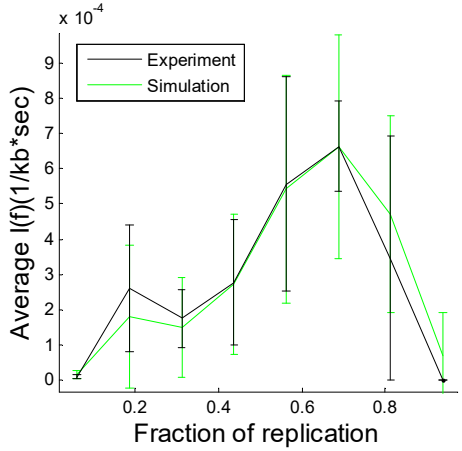


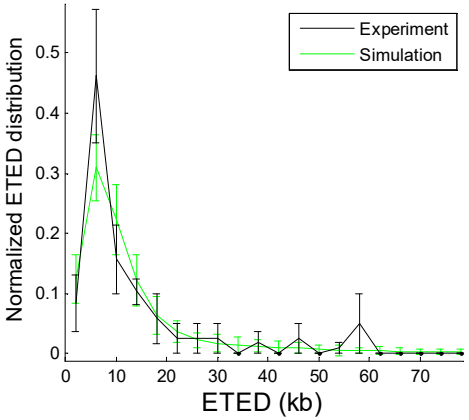
Results from best model (Absence of Chk1 inhibition):

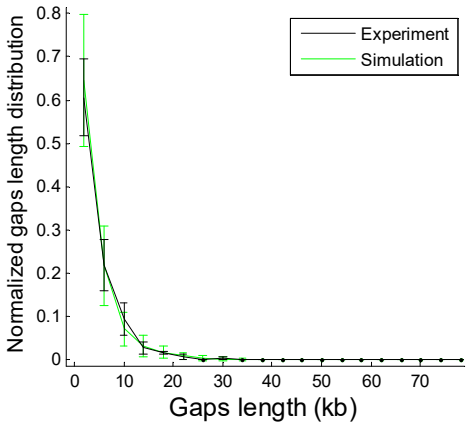
Analysis of early S phase:

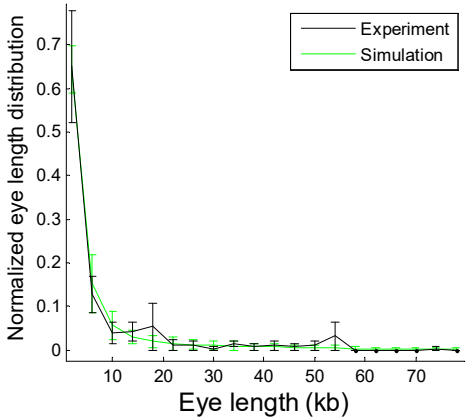


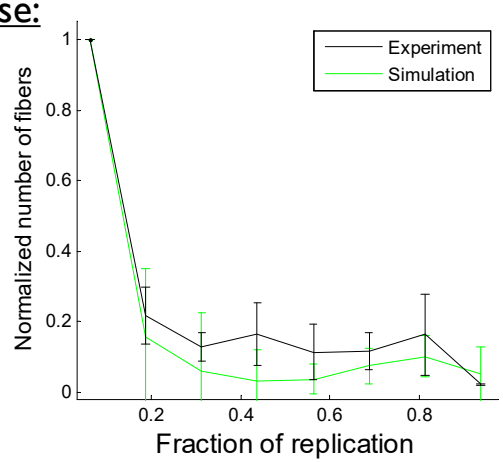
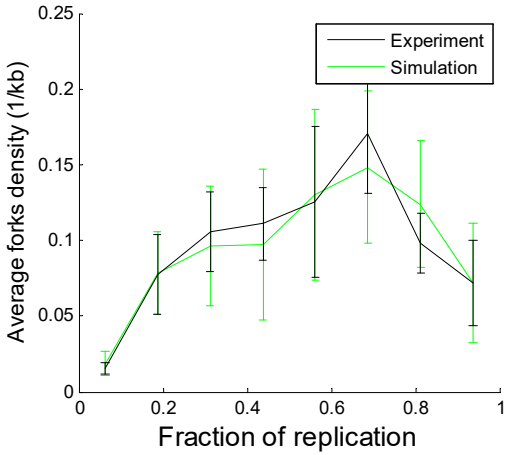
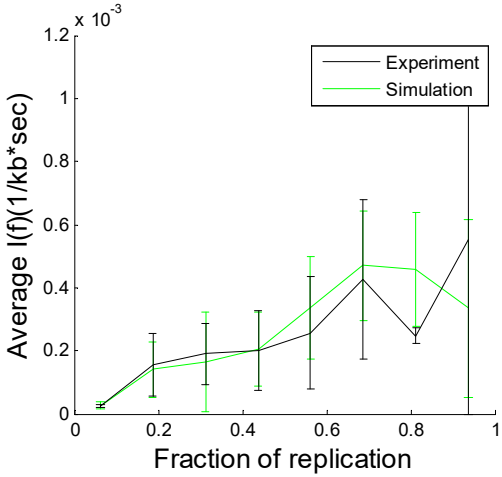
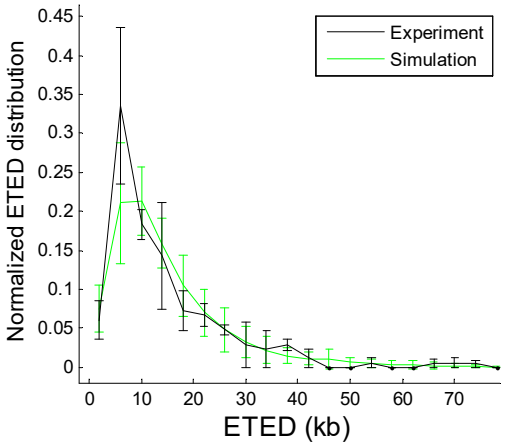
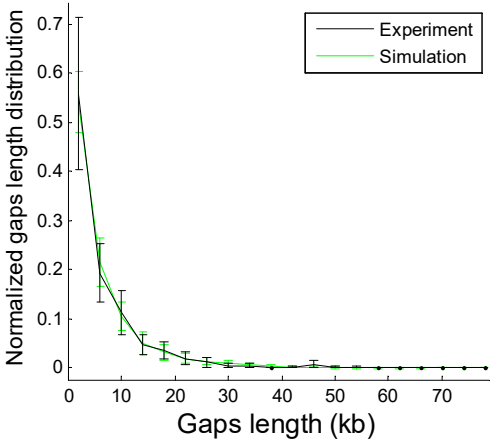
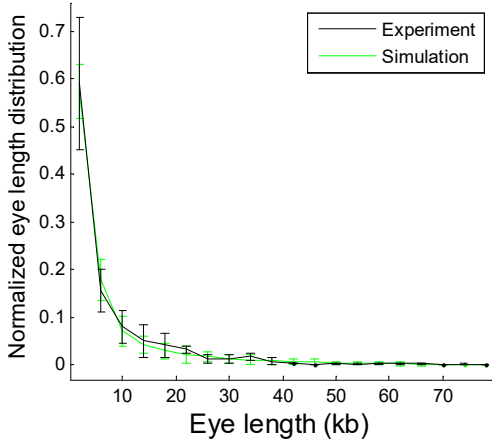










Results from best model (Chk1 inhibition):

Analysis of early S phase: