#### OptKnock Tutorial

Author: Sebastián N. Mendoza. Center for Mathematical Modeling. University of Chile. snmendoz lifuc.cl

Reviewer(s): Thomas Pfau, Sylvian Arreckx, Laurent Heirendt, Ronan Fleming, John Sauls, Anne Richelle

Optionox is an algorithm suggesting the genetic manipulation that lead to the overproduction of a specified metabolite [1]. Opinoxic pinpoints which set of reactions

to to remove (i.e. deletion of the general associated to these reactions) from a metabolic network to obtain a mutant that will produce a particular farout of interest at For example, imagine that we would like to increase the production of succinate or lactate in Eacherichia coil. Which are the knock-outs needed to increase the

MATERIALS & EQUIPMENT 1. MATLAR

2. A solver for Mixed Imager Linear Programming (MLP) problems Use changeCobraGolver to choose the solver for MILP problems (e.g., Gurobi).

#### The proceduce consists on the following steps:

# 3) Define the number of reactions to be deleted, the target reaction and some constraints to be acc

TIMING: This task should take from a few seconds to a few hours depending on the size of your reconstruction

Verify that cobratoobox has been initialized and that the solver has been set global THTOREM\_INST\_CR If -iconpty(TUTURIN\_INST\_CR) && TUTURSAL DAIT CR--1

changeCobratioliver('gurobi', 'all');

Load the model of E. coli US.

modelEirectory = getEistributedModelFolder(modelFileName); %Look up the falder for the distributed Models. modelFileName |modelDirectory filesep modelFileName); % Get the full path. Necessary to be sure, that the right model is loaded sodel = readCtModel(sodelFileNtane);

blommer = "8109855\_6c\_1301366\_core\_53p850";

threshold = 5;

selectedResist = ("Successory "Successory "Missing "Missi "RESPECT; "RECT; "REST; "RET; "REST; "RET; "RELT; "RELT; "RELT; "REST; "

Constraint the model with biological assumptions

model = changeRunBounds(model, 'EX\_glc\_D\_e', -18, 'b');

```
Bounds+[8;-1888;-1886;-1888];
model = changeRunBounds(model, Exchange, Bounds, 'l');
Exchanges (MEC, acuery MEC, as 2, ery MEC, et al., ery MEC, for year y MEC, fact, at year y MEC, as course, ery later.
Bounds+[1000;1000;1000;1000;1000;1000];
model = changeRxnBounds(model, Exchange, Bounds, 'v');
* Cabiffain Toe prospect materials system.
sodel = changekonBuunds(sodel, 'GLCatopp', -1888, 'l');
sodel = changekonBuunds(sodel, 'GLCatopp', -1888, 'k');
sodel = changekonBuunds(sodel, 'GLCatopp', 1888, 'k');
sodel = changeRunBsundc(model, 'GLCptcpp', 1000, 'u');
sodel = changeRunBsundc(model, 'GLCt2pp', 0, 'b');
Then, calculates the production of metabolites before running optionock
 fbawY = aptimizeCbModel(model);
 succFluoNT = fbaNT.x(strcmp(sodel.rxns, 'EX_succ_e'));
etohFluxWT = fbsWT_x(strcmp(model_rxns, 'EX_stan_s'));
 formFluoNT = fooMT.x(strcmp(model.rxns, 'EX_for_e'));
 lactFluoWT = floaMT.x(strcmp(model.rxns, 'EX_lac_D.e'));
 acetFluxWT = fouWT.x(strcmp(model.rxns, 'EX_ac_e'));
growthRateMT = fbaWT.f;
 fariatf('The production of succinate before optimization is w.if \s', succFloat()
 Sprintf("The growth rate before optimization is %.1f 's', growthRateMT);
          'acetate are %.1f, %.1f, %.1f and %.1f, respect
etablishing, formFluodT, lactFluodT, acetFluodT);
B EXAMPLE 1 : SUCCINATE OVERPRODUCTION
Aim: finding cotKnock reactions sets of size 2 for increasing production of succinate
.... EXPECT to Finding outliness sets of size 2 or bess....
constript = struct('runList', {(biomass)), 'values', 0.5=fbmff.f, 'conse', '6');
sites - 1:
while niter < threshold
         aptWrackSol = DutWrackSodel, selectedNowList, outlons, constrbut);
       gatKnackiel = dutKnackisodel, selectedMusiist, outlons, constrbut, previousielutions, 1);
     succFlusM1 = outMockfol_fluxes(strompleodel_rxms, 'EX succ e'));
      growthWateM1 = optKnockSal.flaxec(strcmp(model.rxms, biomass));
      setMi = cotMnockSol_ronList;
      # -icensty(setMi)
          aneviausiolutions(contPreviousiolutions) = setM1:
                   forietf('%', setti());
                 forietf(' and %s', septi(s));
                  fprintf(', wa', setMS(j));
```

Exchange+(https://profitting.org/https://profitting.org/https://profitting.org/https://profitting.org/https://

```
Special Test is produce on the continuous in Californ's produced in Californ's produced in Californ's produced in Californ's Califor
```

forietfi'The production of succionse after actinization to N.25 Not. succFluggit

nîter = nîter + 1;

optionick found a optionick set of large 2 composed by

and RPI

The production of socionate after optimization is 2.75
The growth rate after optimization is 0.15
The production of other products contained to the product of other products of the product of the products o

country country products than an economic, namely, calcular an enting country grantly country one economic

imum growth rate given the uptCnack set is 8.12 imum and minimum production of socionate given the uptCnack set is 8.86 and 2.31, resp

...Performing options analysis... SILP problem with ESPE constraints ES integer variables and ESSE continuous variables Dotable a model with ESPE rows. ESSE columns and SESSE negatives

as a model with EDER room, EXES columns and SETT name for types: EXES continuous, EX lateger (EX binary) const statistics: is range (2x-08, 2x-08) cities come (2x-08, 2x-08)

range (1e-62, 1e-65) by removed METO rous and SEES calamis

Presidenti 3329 raus, 5609 columns, 25627 nomieros Sariable typec: 5612 continuous, 37 integer (37 binary)

Northele Type: 1812 continuous, 37 integer (37 binary)

Root relaxation: objective 1.238325+81, db64 iterations, 8.66 seconds

Nodes | Current Node | Dijective Bounds | Mark

| Sept. Marrier | Sept. Section | Localization | Section | Localization | Localiz

aptored 379 modes (3522 simples iderations) in 3.75

Solution count to 1.67928 1.66239 1.28237 ... 0.6923893

Optimal colution found (talerance 1.68e-12) Bect sépection 1.5753EE39E55e-88, bect bound 1.6752E633EE56+88, gap 8.6888 coffices from 8 coffices and 6 talera 2 commoned by

and 9072

the production of vaccinarie after optimization is 1.65 the quants rate after optimization is 6.22 the production of other products such as otherwis, formule, textste and acetate arcid.5, 6.8, 6.8 and -6.8, respe

...Performing coupling analysis... The solution is of twee prouth coupled non union The maximum growth rate given the optimist set is 0.18 The maximum and minimum production of socialize given the optimist set is 0.00 and 1.61, respec-...Performing optimist analysis...

Matria range (30-06, 10-05) Spiritur range (30-06, 10-05) Spiritur range (30-06, 10-06) Shound range (30-06, 10-05) MS range (30-07, 10-05)

N range (1e-82, 1e-83) olve removed MFP rook and SSE1 columns

Precolved: 3338 raws, 5600 calamos, 25650 nonzeros Nariable types: 3622 continuous, 37 integer (37 binary

Next relaxation: objective 1.200325e=81, 5100 iterations, 0.

| Marke | Current Base | Diportice Baseds | Mork Engl Deeps! | Diportice Baseds | Mork Engl Deeps! | Diportice Baseds | Mork | Diportice Baseds | Diportice Baseds | Diportice Baseds | Diportice Based | Diportic

291 18 21 1.442987 12.2778 782
 Explored STS modes (SSST9 simples identions) in 3.08 seconds

Explored STS nodes (SSST9 simples iterations) in 3.68 seconds Thread count was 8 (of 8 available processors)

Solution count S: 1.46299 1.36292 1.36392 ... 6.6923890 Pool signification bound 1.46298

Optimal solution found (tolerance 1.00=-12)
Best significe 1.66299373957=-00, best bound 1.66299373957=+00, gap 6.00004
optimick found a optimick set of large 2 composed by

and RPE

he production of succionse after optimization is 1.66 he growth rate after optimization is 6.22

Performing coupling analysis...

ne tailing at it types graces toopies out unappe to maximus growth rate given the uptChock set is 8.38 he maximus and minimus production of taclinate given the uptChock set is 8.86 and 1.22, res

...Performing optional analysis...
MILP problem with BDMS constraints EE integer variables and EETS continuous variables optimize a madel with BDMS room, EETS columns and SDETY nanoerom thrubble types: EETS continuous, EE integer (EE binary)

Dent statistics: x range (3e-66, 2e-65) Elve range (3e-66, 2e-65) k range (3e-66, 2e-65)

Preside removed METO rous and EXII calams Preside time: 6.0% Presided: EXII rows, METO calams, 20021 nonzeros

Presided: 333 raws, 5800 calumns, 25021 nonzeros Sariable types: 3812 cantinuous, 27 integer (27 binary)

Node: | Current Node | Dijection South Conf. | Mark.
Engl Desgt | On Depth Inite! | Dijection South Con | Inite.

Explored 390 codes (X2027 simples iderations) in 3.50 seconds Thread count was E (of E available projectors)

Solution court & 1,26592 6,289687 6,265365 6.6

Optimat solution found (taterance 1.880-13) Marriang man constraint validation (1.3830-88) wasseds tolerance Best sejective 1.78091783887+88, best bound 1.38291783867+88, gap 8.88884

4900

III EXAMPLE 2: SUCCINATE OVERPRODUCTION

Aim: finding gotKnock reactions sets of size 2 for increasing production of succinate 

... EXPEND to Finding outpools sets of size 3....

constitut = struct('conlist', ((biomass)), 'values', #,5+fbmW.f. 'conce', '6'); elter = 1: while niter < threshold

agromackiel = dythock(model, selectedhunist, agrions, constrbpt, previousiolations);

succFlusM1 = outMockfol\_fluxes(strompleodel\_rxms, 'EX succ e')); growthWateM1 = optKnockSal.flaxec(strcmp(model.rxms, biomass));

setMi = cotMnockfol\_ronList;

aneviausiolutions(contPreviousiolutions) = setM1: forietf('cottlock found a cottlock set of large nd companed by '.length(cottl));

forletf('%',setMt(i));

forietf(' and %',setMp(13); forietf(', %', septi(i));

forietf('The production of succinate after actinization is %.2f \s', succFluxts); fgrietf('The growth rate after optimization is w.2f \s', growthkateMi);

"N.15, N.15, N.15 and N.15, respectively, No'l, etobylasts, foreylasts, lactylasts, acetylasts); [type, maximust, maxProd, minProd] = analyzedotKnock/model, setMi, "Mi succ e")u forietf('The solution is of type; NCAC, type);

singleProductionSovelaps(model, setts, 'Expact\_s', biomass, 'savePlot', 1, 'showPlat', 1, ...
'filetone', Prace and ' magnetrisiteril, 'naturalelder', 'becomeshowite');

if olter == 1

olter = olter + 1;

, LDK D

, LDI,D



### II) EXAMPLE 3 : LACTATE OVERPRODUCTION

Aim: finding optKnock reactions sets of size 3 for increasing production of lactate forintfills...EXMPLE 1: Finding outdoork sets of size 2...

```
... DODDER In Finding options note of size 3...
options = struct('targethus', 'EX_lac_0.e', 'numbel', 3);
constript = struct('restist', ((biomass)), 'values', 0.5*fbmW.f, 'sense', '6');
previousiolutions = cell(100, 1);
while niter < threshold
    if isempty(previousholutions(1))
       agronackiel = Sythock(model, selectedHxxList, options, constrbpt);
       aptKnackSol = BytKnack(model, selectedMxxList, aptions, constrbyt, previousSolutions);
    lactFluxMS = optNnockSol.fluxes(strcmp(model.rxms, 'EX_lac_0.e'));
    growthMateM1 = optKnockSul.flaxes(strcmp(model.rxms, bGomass));
    etohFluxMS = optNnockSol.fluxes(strcmp(model.rxms, 'EX_etoh_e'));
    formFluxMS = optNockSol.fluxes(stromp(model.rxms, 'EX for a'));
    succFluxMS = optNnockSol.fluxes(strcmp(model.rxms, 'EX succ.e'));
    acetFluxMi = optNockSol.fluxes(stromp(model.rxms, 'EX.ac.e'));
    setMS = optWnockSol.rwnList;
        previous@olutions(contPrevious@olutions) = setM1;
        forietf('optNock found a optNock set of large nd composed by ',length(setMI));
        for j = 1:length(cattl)
            elseif j == length(court)
              fprintf(' and we', capti(j));
                fprintf(', wa', setMS(j));
        forietf('The oranth rate after optimization is %.2f \s', growtHateMI);
                's.if, s.if, s.if and s.if, respectively. \n'], etabFluxEs, formFluxEs, succFluxEs, acetFluxEs);
        [type, maximumth, maxProd, minProd] = analyzedptKnock(model, setMI, "EX_loc_D_o");
        forietf('The colution is of type: %c(a', type);
        foristf('The maximum growth rate given the optKnock set is %.2f(n', maximumth);
                "w.2f and w.2f, respectively (40,0"), minProd, maxProd);
        singleProductionEnvelope(mode), cmtf1, 'EX_lac_D.e', biomass, 'cavePlot', 1, 'chouPlat', 1, ...
                                 "filetone", ['lact_est_' sumbtroafter)], 'outor
        if alter == 1
 The production of other products such as ethanot, formate, succinate and acetate are 8.8, 8.8, 8.8 and 8.8, respectively
```

The production of lactate after optimization is 18.13 The growth rate after optimization is 6.12

...Performing coupling analysis...
The saturion is of types non-unique
The nazimum neuthrists given the appEnnix set is 6.88

The maximum growth rate given the uptionic set is in The maximum and minimum production of lactate given

...Performing optional analysis... options found a options set of large 3 compos

, PEI

The conduction of factors offer enthaltering to the

he growth rate after optimization is 0.12

The production of other products such as ethan ...Performing coupling analysis...

e maximum growth rate given the optionix set is 8.89 e maximum and minimum production of lactate given the optionix set is 8

active conditions and a spit active conditions are active conditions and active conditions are active conditions and active conditions are active conditio

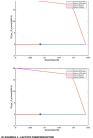
o PCE production of lactate after optimization is 18.85

ground rate after optimization is 6.12 products cach as ethanol, formate, succinate and acetate are 6.8, 6.6, 6.8 and 6.1, referring coupling analysis...

he maximum growth rate given the optionic set is 8.89







#### Aim: finding optKnock reactions sets of size 6 for increasing production of lactate

## fprintf('...SAMPLE 3: Finding optKnock sets of size 6...\n')

....DOPPLE In Finding optimate sets of size 6....

when operate options will be the objective of the outer problem options of increase will be the objective of the outer problem options of extract("inspectar", "including of options," (subset), "including options of the outer options of the

siter = 1; unite siter < threshold

sempty(previous(o)utions(1))
sprknack(o) = bytknack(sade), selectedRunlist, options, constrbyt);
los

agtKnackiol = OptKnack|model, selectedRexList, options, constrUpt, previousbolutions);

lactiage = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o''))
gouchMatest = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o'''))
gouchMatest = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o'''))
gouthMatest = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o'''))
forsilant = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o'''));
forsilant = optmockiol.Thans(strong)socia.rem; 'Nc(ac\_o'''));
socialism; optmockiol.Thans(strong)social.rem; 'Nc(ac\_o'''));
socialism; optmockiol.Thans(strong)social.rem; 'Nc(ac\_o'''));

if -icempty(setMi) previousSolutions(contPreviousSolutions) = setMi;

\*marketing results fighterf(\*opthock found a opthock set of large \*d composed by \*, length(set#1)); for j = likeght(set#1)

if j == 1
 fprintf('%c', catM1{j});
elseif j == leogth(catM1)

```
The Control of the Co
```

The production of other products can be related, fromter, consistent and metrics are i.e., E.C., E.A. and E.L., respectively,
including complete productions.

The desiration product of prince the explication of i.e. EAT

The desiration prince of the explication of i.e. EAT

The desiration prince of the explication of i.e. EAT

The desiration prince of the explication of i.e. EAT

The desiration prince of i.e. EAT

The desiration of i.e. EAT

The

e production of lacture after optimization in 17.00 opcome fine after optimization in 6.32 e-production of the production on actional, forestle, occinate and actists are 6.8, 6.7, 6.8 and 6.1, respect Performing compling analysis... existing in 60 person occupies on scalage.

The satisfactor is of Types non-unique
The maximum operator are given by spotents set is 0.89
The maximum and minimum production of latitate given the optionick set is 0.00 and IE-02, respectly
.../Performing optionic maniprist...
optionics found a spotionic set of large d composed by

, PEE , PEE and TOTS

forietf(' and M', septi());

The production of loctate after optimization is 17.00 The quantity rate after optimization is 6.12 The production of other products can be of behavio, foreste, occinate and acetate are 6.6, 6.7, 6.8 and 6.1, respectively.

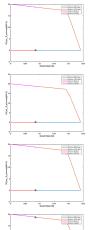
The salelies is of type non unique
The maximum growth rate given the applicanck set is 6.89

The maximum and minimum production of lactate given the optionick set is 8.88 and 18.82, respective...Performing optionic analysis... optionick Toward a optionick set of large 6 composed by

, MCEEN , GLCHRON , PEE , PIK by promotions of interest enter optimizations in 15-70 to growth risk after optimization in 16-12 to growth risk after optimization is 6-12 to product of other products such as ethnost, formate, succinate and acetate are 6.6, 8.7, 8.8 and 8.1, respectively.

The solution is of type: grawth coupled non unique The maximum growth rate given the optionics set is 0.00

The maximum and minimum production of lactate given the optionick set is 38.38 and 38.82, respectively





THERE

Example 1 – 1-2 minutes
 Example 2 – 1-2 minutes

## 4. Example 4 – 1-2 minutes TROUBLESHOOTING

1) If the argoment steer a long time to find a solution, it is possible that the search space is too large. You can reduce the search space using a smaller set of reactions in the input variable. \*search-space using a smaller set of reactions in the input variable. \*Search-space using a smaller set of reactions used to input variable of search space using a smaller set of reactions used to input variable of search space using a smaller set of reactions.
2) The default mumber of deletions used by optimize it is if the apportion is returning more deletions that what you want, you can change the number of deletions.

3) opticionic could find a solution that in is not useful for you. For example, you may think that a solution is very obvious or that it breaks some important biological contraints. If opticionic bound a solution that you don't want to find, we the input variable "previousment" to prevent that solution to be found.

constants, regenerate municipal sections may gue continue to most, use the special variables previouslation for prevent that column to the studies.

The opposition skippithms will find use of mentions that, when sectioned from the most skip interpretable production of the mentionist of interpretable of interesting or, succional and management of the mentionist interpretable or interesting or accordance and management of the mention is destinated for the colorable consistent format of the mentionist format formation or not formed the measure. In this shadow, one conditions that such doubtless indeed reference to produce consistent shadows in the mention of the mentionist formation or mentionist or mention or mention is destinated formation.

scenery or restruction, usual polymonal, i.e., the production side will increase as biomass formation increases. For these kind of mactions a piot will be generated using the function single/Production/Energie and will be assent in the folder administration/Energy production/Energy and will be assent in the folder administration/Endoccidentals.

When source of a sold door with Collection via obtained above the formation and anterium conduction are wiston the function analized/beforces.

en you find a solution with Optificati, you should always verify the minumum and maximum production rate using the function analize/Optification. Renercoss

[1] Burgard, A. P., Phariqu, P. & Moranas, C. D. (2000). Cyptiolock: A Billiveli Programming Plasmework for Identifying Gene Woodcoot Stranges for Microbial Strain Cyptionalization. Educationalogy and Biologicaering, 84(8), 647–647. https://doi.org/10.1006/strainide.
30 (2007). A. D. Common, T. M. N. J., Lammar, J. A. Naus, N., Felder, M. M., & Pallison, G. (2017). A comprehensive genome-scale inconstruction of Euchedolai

[9] Oth, J. D., Connad, T. M., Nu, J., Leman, J. A., Nam, H., Feler, A. M., & Palsson, B. G. (2011). A comprehensive genome-scale reconstruction of Eucheschia coli metabolism—2011. Molecular systems biology, 7(1), 525.