

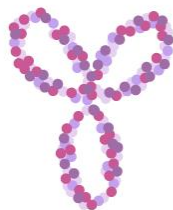
Biosimilars: A new innovative approach to optimize the biosimilar cell-clone selection using high-throughput methods

Thomas de Marchin | Associate Director Statistics & Data Science

Prepared for NCS 2022
20OCT2022

What is a biosimilar?

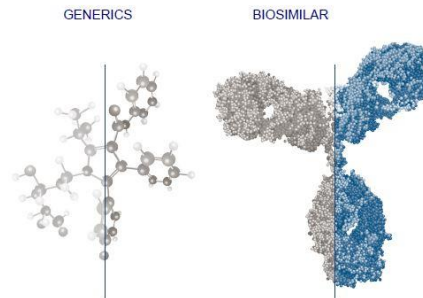
- ▶ A Biosimilar is a biologic drug (hormone, cytokine, antibody...) that is almost an identical copy of an original product that can be marketed when the original product's patent expires. Substantially cheaper as they do not need to be taken through the discovery and phase I+II clinical stages.



Reference product



Biosimilar

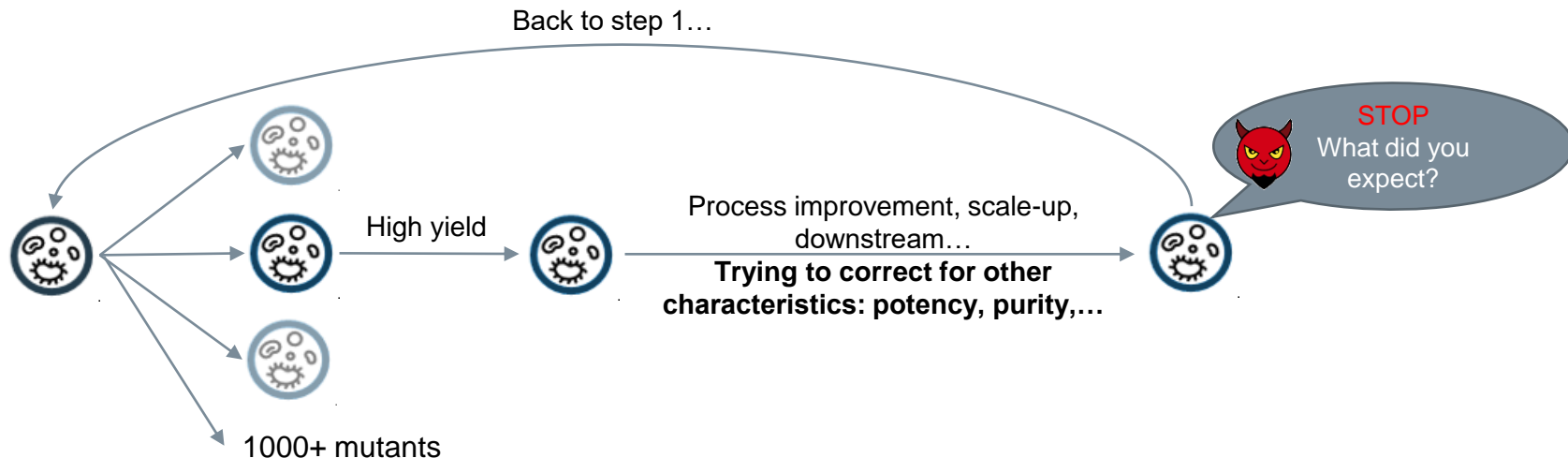


- ▶ In contrast to generics, biosimilars are not identical copies of the original product (made from living organism, complex structure, folding, post translation modifications...)



Picking the best clone

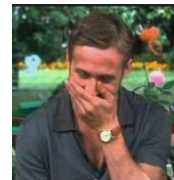
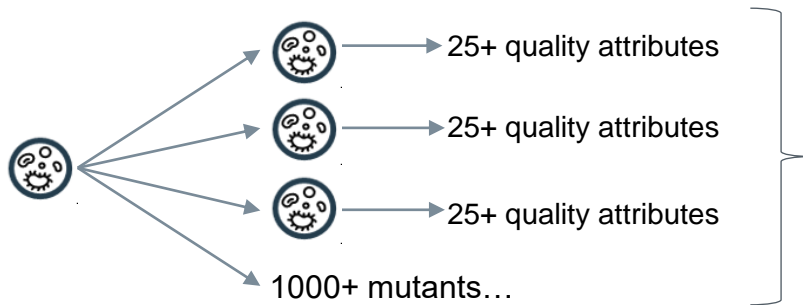
- ▶ To demonstrate that the biological compound is highly similar to the reference product: need to show the equivalence of Critical Quality Attributes (CQAs) of both products → Biosimilar manufacturers must carefully select and/or develop further their own cell lines (clones).



- ▶ There are cases whereby some characteristics of biosimilarity cannot be achieved without significant detriment to the commercial viability of the program → back at square one (re-selecting the cell-clone)

Picking the best clone

- High-throughput methods are more and more used to ensure the cell-clone selected not only accounts for the yield but also for other characteristics



- Biosimilar companies are still missing a smart, robust and yet adaptive selection approach:
 - How to process overwhelming breadth of data (e.g. 1000+ clones)?
 - How to choose the best clone (i.e. the one which will meet all criteria jointly)?
 - How to take into account the analytical method variability?



Introducing the CellClone app

Upload an infinite number
of QA and clones

Impute missing values

Define targets and
specifications for
each CQA and
compute the
desirability values for
each clone

Rank clones from
best to worst, based
on probability to
meet the target

CellClone

Introduction

Data

Missing values

Desirability

Ranking

About/Tool description

Generate report

Introduction

The aim of this application is to help selecting the best clone in a biosimilarity context. A biosimilar is a biologic drug or medical product that is almost an identical copy of an original product that can be produced in a different way and be marketed when the original product's patent expires. Generics alike, biosimilars can be marketed substantially cheaper as compared to the original product as they do not need to be taken through the discovery stage and phase I-II clinical stage and subsequently the development cost is quite reduced. To demonstrate that the biological compound is highly similar to the reference product, analytical studies are performed to show the equivalence of Critical Quality Attributes (CQAs) of both products. Unlike small white molecules (drugs usually produced by chemical synthesis), biosimilars are produced by living cells due to their larger size. Their more complex structures and activity depends on the folding of the protein and associated modifications such as glycans produced by the cell. Hence, the biosimilar manufacturers must carefully select and/or develop further their own cell lines (clones) to produce the molecule of interest (hormone, cytokine, antibody). The goal of this application is to standardize the clone selection process and to facilitate the decision making. This involves:

1. Definition of targets and specifications to achieve for each CQA.
2. Computation of the desirability values for each attribute based on a set of desirability functions.
3. Definition of the relative weight of each CQA and computation of the overall desirability (probability to achieve all targets jointly).
4. Ranking of the clones from best to worst, based on their probability to meet the target.

Quality attributes (by method)

Identity

Efficacy of active ingredient

Production / effectiveness / yield

Impurities / side products

Safety

Production input

Stability

... 25+ quality attributes

Different clones + production processes

Cell clone candidate for production

... 1000+ unique clones

Lab analysis

Regional HA requirements

Biosimilars expertise

GMP

Trust in method

Weighting the quality attributes

Scoring / profiling of cell-clones

Simulations and decision support

Contact:

Bruno Boulanger

Thomas de Marchin

Peter Leister

Peter Laurencic

PHARMALEX

CONFIDENTIAL BEYOND COMPLIANCE

The data

CQAs

clones

ID	G0_F	G1_F	Man5	Acidic_peak	Main_peak	Basic_peak	Binding	FCRN_Binding	RIIA_binding	Main_purity	Aggregates	Degradants	Yield
1	21	11	0	30	40	9	80	100	80	99	1	1	0.5
2	23	13	1	29	43	15	81	88	87	98	1	1	0.5
3	25	15	2	28	46	15	82	98	120	98	2	0	0.7
4	27	16	3	27	48	15	83	87	100	97	2	1	0.6
5	26	17	4	26	50	14	84	86	101	96	3	1	0.7
6	33	8	5	25	52	13	100	85	103	96	4	0	0.8
7	36	7	6	24	48	12	112	100	104	96	?	4	1.5
8	23	6	7	23	46	11	105	103	98	95	?	2	1.6
9	25	5	8	29	55	10	103	102	70	96	?	1	1.7
10	34	10	9	29	51	14	102	104	60	95	3	2	1.8
11	40	11	10	28	43	15	103	105	50	94	4	2	2.2
12	60	15	2	28	43	19	89	106	40	93	7	0	2.5
13	10	20	5	33	43	20	89	107	98	80	10	10	6
14	22	18	8	34	45	20	87	108	96	97	1	2	6.5
15	23	21	6	35	45	20	86	109	95	96	2	2	6.5
16	24	9	5	40	45	21	85	110	94	97	2	1	4
17	27	8	3	38	48	14	86	11	96	98	1	1	4.2

Impute missing values (Predictive Mean Matching for the moment, to be explored...)



Target, specifications and desirability

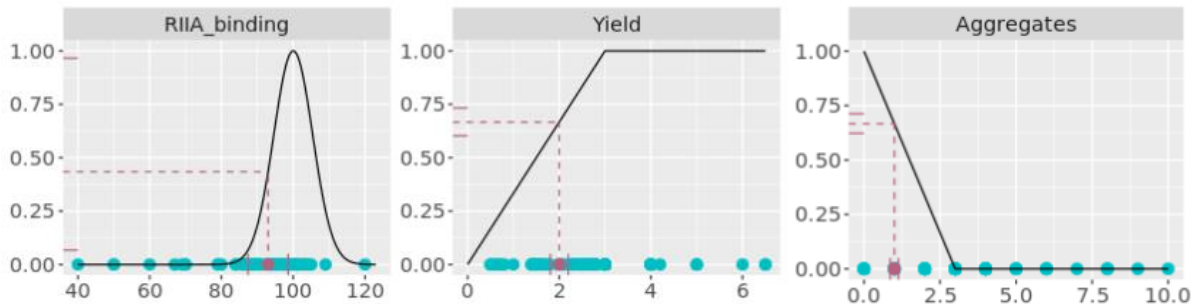
- Coloured points = clones (red = selected one)
- X-axis = value (measurement)
- Y-axis = desirability
- Black line = desirability function
- An overall desirability is computed for each clone:

$$\exp\left(\frac{\sum_{i=1}^n w_i \ln x_i}{\sum_{i=1}^n w_i}\right)$$



Target, specifications and desirability

- Desirability function is calculated from
- Specifications (limits) should be derived from the reference product
- Relative weights can be defined for each CQA
- Intermediate precision is used to simulate plausible “true values” for the measurement → allows to compute a desirability CI by simulation



CQA

RIIA_binding ▼

Lower limit:

77

Upper limit:

123

Calculation method:

Gaussian distribution ▼

Weight:

8

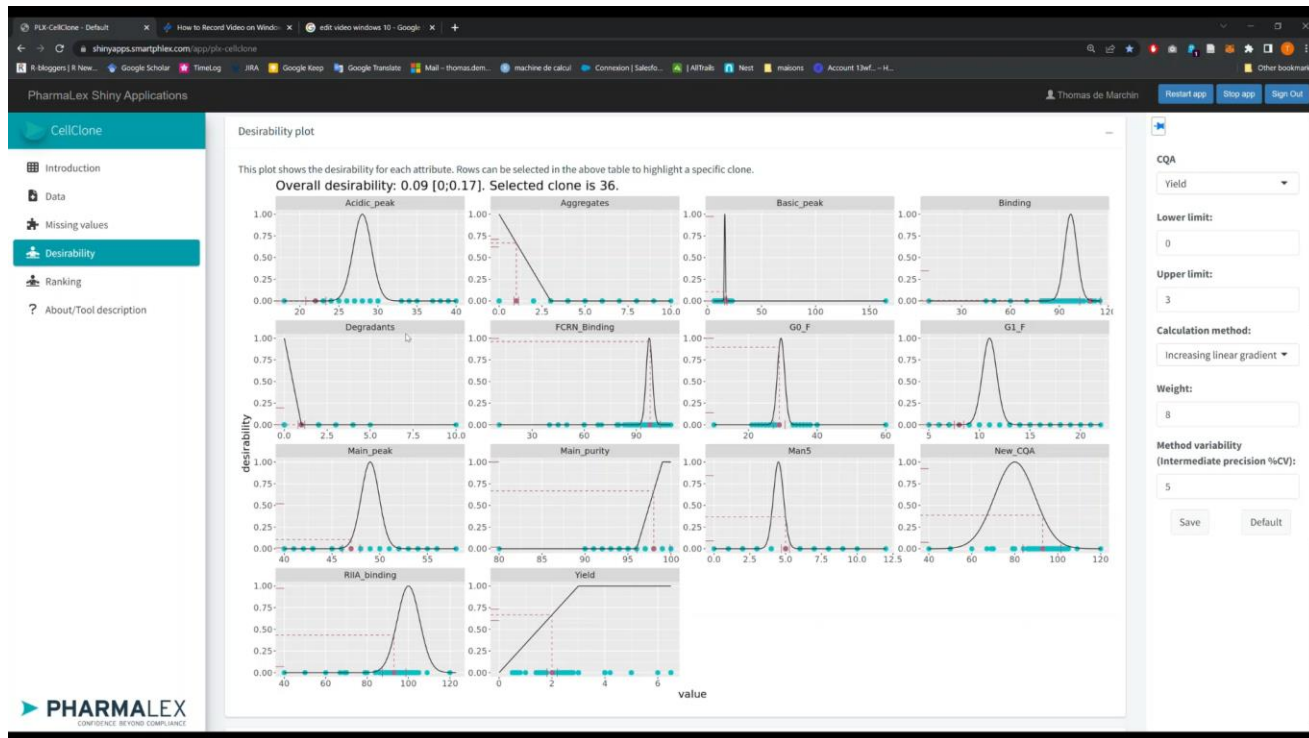
Method variability (Intermediate precision %CV):

3

Save Default

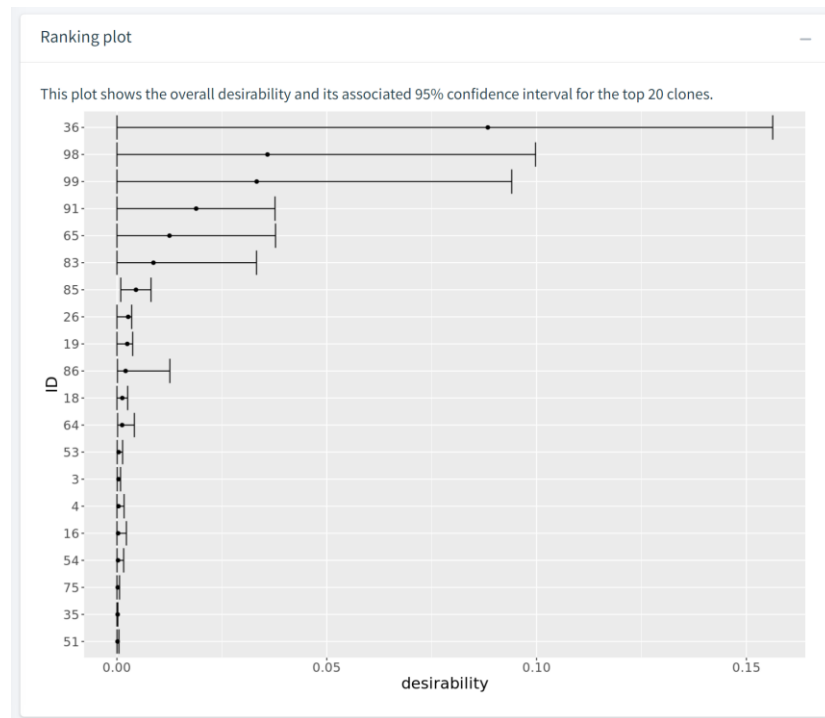
Target, specifications and desirability

- We provide default values, and you can adapt interactively



Clone ranking

- Ranking of the clones from best to worst
- Importance of the measurement uncertainty → CI!
- Final outcome is a "basket" of clones with high probability of future success



Summary

- ▶ We propose:
 - Approach which optimizes and standardizes the clone selection process
 - Includes the measurement uncertainty!
 - App easy to use, especially for non-statisticians
 - Aimed at biosimilar developers but might also be attractive for clients with innovative biopharmaceuticals (emphasis on clone quality also rising).
- ▶ Next step:
 - Investigate the contribution measurement variability (IP) of each CQA to the uncertainty of the global desirability → which analytical method should be improved?
 - Investigate missing value imputation

<https://shinyapps.smartphlex.com/app/plx-cellclone>



Thomas de Marchin

Associate Director, Statistics & Data Science

Thomas.demarchin@pharmalex.com



<https://www.linkedin.com/in/tdemarchin/>

Peter Lavrencic

Manager, Consultant DCS/BPE

peter.lavrencic@biopharma-excellence.com



Bruno Boulanger

Senior Director, Global Head Statistics and Data Science

Bruno.boulanger@pharmalex.com



All content is based on the current science and best practices and is subject to change. This document is property of the PharmaLex Group and is intended for the client only. Unauthorized use, disclosure, or copying of the information contained in this document is prohibited.

List of suggested default quality attributes

- ▶ Yield
- ▶ Charge heterogeneity
 - Acidic
 - Basic
 - Main
- ▶ Potency
 - FC RN binding
 - RIIA binding
- ▶ Aggregation
 - Aggregation
 - degradants
- ▶ Glycans
 - G0F
 - G1F
 - Mannose



Appendix

Desirability table

This table lists the overall desirability value as well as the desirability for each attribute.

Desirability functions are normal distributions whose parameters are derived from the expected ranges observed from the reference.

Individual desirabilities are calculated by projecting the value onto their desirability function. Overall desirability (chance to achieve all targets jointly) is calculated as the weighted geometric mean of individual desirabilities.
clones are ranked from best to worst.

	ID	Overall desirability	Acidic_peak	Aggregates	Basic_peak	Binding	Degradants	FCRN_Binding	G0_F	G1_F	Main_peak
1	36	0.09 [0;0.16]	0 [0;0]	0.67 [0.62;0.71]	0.11 [0;0.97]	0.01 [0;0.33]	0 [0;0.21]	0.96 [0;1]	0.89 [0.12;1]	0 [0;0]	0.11 [0.01;0.11]
2	98	0.04 [0;0.1]	0 [0;0.08]	0 [0;0.13]	0.11 [0;0.96]	0.87 [0.16;1]	0 [0;0.21]	0.14 [0;1]	0.37 [0.01;0.99]	0.02 [0;0.12]	0.01 [0;0.09]
3	99	0.03 [0;0.09]	0 [0;0.01]	1 [1;1]	0.11 [0;0.95]	0.28 [0.01;0.98]	0 [0;0.2]	0 [0;0.86]	0.06 [0;0.71]	1 [0.58;1]	0 [0;0]
4	91	0.02 [0;0.04]	0 [0;0.08]	0.67 [0.62;0.71]	0.99 [0.1;0.99]	0.28 [0.01;0.97]	0 [0;0.2]	0.7 [0;1]	0.89 [0.15;1]	0.02 [0;0.1]	0 [0;0.01]
5	65	0.01 [0;0.04]	0.7 [0.11;1]	0.33 [0.23;0.42]	0 [0;0]	0.97 [0.19;1]	0 [0;0.2]	0 [0;0.82]	0 [0;0.17]	0.02 [0;0.13]	0.01 [0;0.08]
6	83	0.01 [0;0.03]	0.7 [0.09;1]	0 [0;0.15]	0 [0;0.05]	0.42 [0.01;0.99]	0 [0;0.2]	0 [0;0.87]	0 [0;0]	0.37 [0.06;0.91]	0.57 [0.11;0.99]
7	1	0 [0;0]	0.24 [0.01;0.98]	0.67 [0.62;0.71]	0 [0;0]	0 [0;0.01]	0 [0;0.2]	0.37 [0;1]	0 [0;0]	1 [0.61;1]	0 [0;0]

Showing 1 to 10 of 100 entries

Previous

1

2

3

4

5

...

10

Next

Appendix

Parameters

These are the limits, calculation method and method variability specified for each CQA. Default values are provided for some CQAs. These values can be adjusted using the right sidebar menu.

	CQA	Lower limit	Upper limit	Calculation method	Weight	Method variability (Intermediate Precision % CV)
1	Acidic_peak	23	33	Gaussian distribution	1	3
2	Main_peak	45	53	Gaussian distribution	3	1
3	Basic_peak	14	18	Gaussian distribution	1	3
4	Binding	81	113	Gaussian distribution	8	3
5	FCRN_Binding	90	105	Gaussian distribution	8	3
6	RIA_binding	77	123	Gaussian distribution	8	3
7	Main_purity	96	99	Increasing linear gradient	5	1
8	Aggregates	0	3	Decreasing linear gradient	5	7
9	Degradants	0	1	Decreasing linear gradient	5	10
10	Yield	1	4	Gaussian distribution	8	5

Showing 1 to 10 of 14 entries

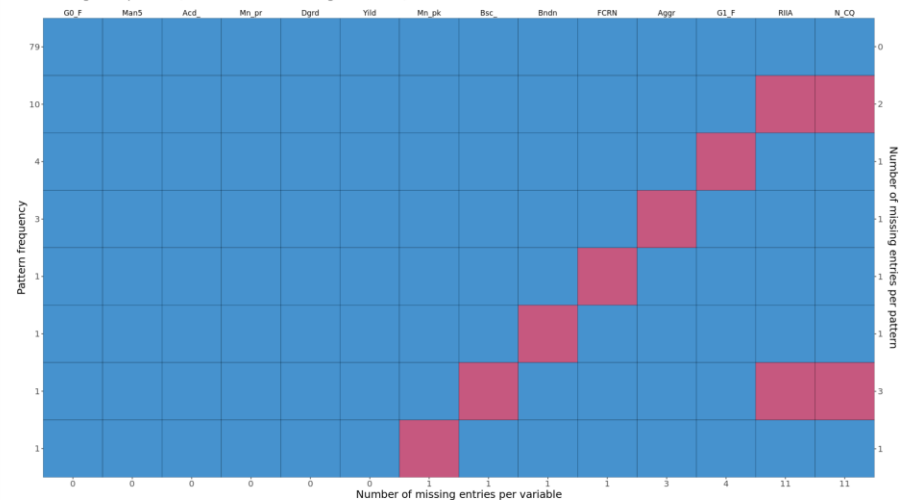
Previous **1** 2 Next

Appendix

Missing value pattern

This plot shows the missing data pattern. Blue is observed, red is missing.

Missing data pattern (total number of missing cells = 33)



Impute missing value

Impute missing values

This plot shows the distribution of the values. Grey is observed, red has been imputed.

