



**HAL**  
open science

## Modeling the growth of tomato fruits based on enzyme activity profiles.

Daniel Jacob

► **To cite this version:**

Daniel Jacob. Modeling the growth of tomato fruits based on enzyme activity profiles.. [Research Report] INRAE. 2020. hal-02611223

**HAL Id: hal-02611223**

**<https://hal-cnrs.archives-ouvertes.fr/hal-02611223>**

Submitted on 19 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike| 4.0 International License

# Modeling the growth of tomato fruits based on enzyme activity profiles

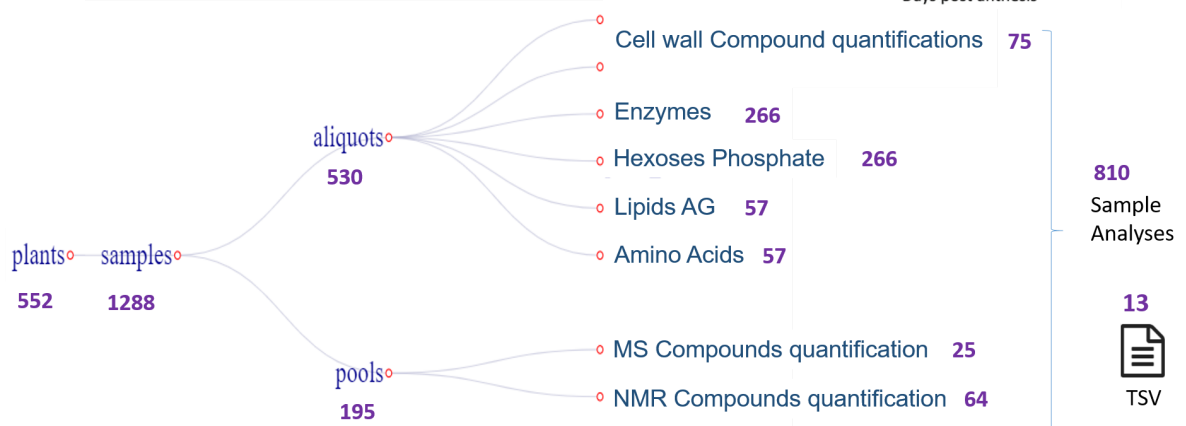
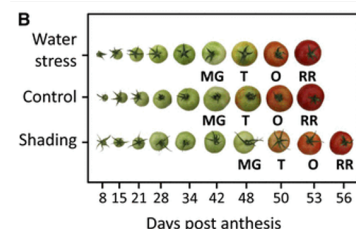
Daniel Jacob - UMR1332 Fruit Biology and Pathology & Bordeaux Metabolome Facility at INRAE

## FRIM data analysis interfaced by ODAM

To assess the influence of the environment on fruit metabolism, tomato (*Solanum lycopersicum* 'MoneyMaker') plants were grown under contrasting conditions (optimal for commercial, water limited, or shaded production) and locations. Samples were harvested at nine stages of development, and 36 enzyme activities of central metabolism were measured as well as protein, starch, and major metabolites, such as hexoses, sucrose, organic acids, and amino acids.

### FRIM1

- **580 tomato plants** were grown in a **greenhouse** in the southwest of France (Sainte-Livrade sur Lot - 2010)
- Fruits were harvested on **3 different trusses** (5–7) and at **9 stages of development**.
- Treatment - **3 levels** : **Control, Water stress, Shading**



## Online documentation

### Frim1 dataset

- References :
  - Biais B, Bénard C, Beauvoit B, Colombié S, Prodhomme D, Ménard G, Bernillon S, Gehl B, Gautier H, Ballias P, Mazat J-P, Sweetlove L, Génard M, Gibon Y. 2014. Remarkable reproducibility of enzyme activity profiles in tomato fruits grown

under contrasting environments provides a roadmap for studies of fruit metabolism. [Plant Physiology 164, 1204-1221](#)

- Dataset DOI : [doi:10.15454/95JUTK](https://doi.org/10.15454/95JUTK)
- Dataexplorer : <https://pmb-bordeaux.fr/dataexplorer/?ds=frim1>

## ODAM

- <https://bio.tools/ODAM>
- <https://inrae.github.io/odam-docs/>

## Metabolism Team

- [https://www6.bordeaux-aquitaine.inra.fr/bfp\\_eng/Research/Metabolism-Team](https://www6.bordeaux-aquitaine.inra.fr/bfp_eng/Research/Metabolism-Team)

# Initialisation

```
[2] source("Functions.R")
```

- Init ODAAM object : retrieve the dataset structure along with their metadata

```
[3] # FRIM1 dataset managed by ODAAM
dh <- new('odamws', 'https://pmb-bordeaux.fr/getdata/', 'frim1')

# Show dataset structure
options(width=128)
show(dh)
```

		levelName	SetID	Identifier	WSEntry
Description Count					
1 plants			1	PlantID	plant
Plant features	552				
2 °--samples			2	SampleID	sample
Sample features	1288				
3  --aliquots			3	AliquotID	aliquot
Aliquots features	530				
4    --cellwall_metabo			4	AliquotID	aliquot Cell wall
Compound quantifications	75				
5    --cellwall_metaboFW			5	AliquotID	aliquot Cell Wall
Compound quantifications (FW)	75				
6    --activome			6	AliquotID	aliquot
Activome Features	266				
7    --plato_hexosesP			10	AliquotID	aliquot
Hexoses Phosphate	266				

```

8      | |--lipids_AG      11 AliquotID aliquot
Lipids AG      57
9      | °--AminoAcid  12 AliquotID aliquot
Amino Acids    69
10     °--pools      7   PoolID    pool
Pools of remaining pools  195
11     |--qMS_metabo  8   PoolID    pool      MS
Compounds quantification  25
12     °--qNMR_metabo 9   PoolID    pool      NMR
Compounds quantification  64

```

- List of 'samples' factors

```
[4] dh$getWS('(samples)/factor')
```

A data.frame: 3 × 6

Subset	Attribute	Description	Type	CV_Term_ID
<chr>	<chr>	<chr>	<chr>	<chr>
plants	Treatment	Treatment applied on plants	string	<a href="http://www.ebi.ac.uk/efo/EFO_C">http://www.ebi.ac.uk/efo/EFO_C</a>
samples	DevStage	fruit development stage	string	<a href="http://purl.obolibrary.org/obo/F">http://purl.obolibrary.org/obo/F</a>
samples	FruitAge	fruit age (dpa)	string	<a href="http://purl.obolibrary.org/obo/F">http://purl.obolibrary.org/obo/F</a>

- List of 'samples' quantitative variables

```
[5] dh$getWS('(samples)/quantitative')
```

A data.frame: 5 × 6

Subset	Attribute	Description	Type	CV_Term_ID
<chr>	<chr>	<chr>	<chr>	<chr>

Subset	Attribute	Description	Type	CV_Term_ID
<chr>	<chr>	<chr>	<chr>	<chr>
samples	FruitDiameter	Fruit diameter (mm)	numeric	http://aims.fao.org/aos/aξ
samples	FruitHeight	Fruit height (mm)	numeric	http://aims.fao.org/aos/aξ
samples	FruitFW	Fruit Fresh Weight(g)	numeric	http://aims.fao.org/aos/aξ
samples	FruitDW	Fruit Dry Weight(g)	numeric	http://aims.fao.org/aos/aξ
samples	DW	Percentage of dry matter (% DW), measured after lyophilisation	numeric	

## Modelisation of the plant growth

Analysing and modelling plant growth is an important interdisciplinary field of plant science. The use of relative growth rates, involving the analysis of plant growth relative to plant size, has more or less independently emerged in different research groups and at different times and has provided powerful tools for assessing the growth performance and growth efficiency of plants and plant populations.

To measure the bioproductivity of a plant, the component of immediate interest is the net primary production or total yield. The plant weight — usually the dry weight - is a needed measurements for growth analysis. First we have to model the plant growth. The optimization parameters are based on a [Levenberg-Marquardt algorithm](#).

### NLS problems with the Levenberg-Marquardt algorithm

- The purpose nlsLM is to minimize the sum square of the vector returned by the function fn, by a modification of the Levenberg-Marquardt algorithm
- <https://www.rdocumentation.org/packages/minpack.lm/versions/1.2-1/topics/nlsLM>

### Models

- Models : Weight = f(time-dependent .i.e Day After Anthesis)

## Model 1 - Single Sigmoidal Model

$$\bullet f(t) = d + \frac{a}{1+e^{-b(t-c)}}$$

## Model 2 - Sum of two Sigmoidal Model

$$\bullet f(t) = d + \frac{a}{1+e^{-b(t-c)}} + \frac{f}{1+e^{-g(t-h)}}$$

- Performance of the modelisation

```
[6] set.seed(1674)

# Sample subset name
setName <- 'samples'

# Time variable for applying the modelisation
Tname <- 'FruitAge'

# Y variable for applying the modelisation
Yname <- 'FruitDW'

# Additional condition
condition <- 'treatment/Control'

# Performance of the modelisation
system.time( fitObj <- fitSigmoid(dh, setName, Tname, Yname,
condition, model=1, info=TRUE) )
```

Getting ... Fitting ... R2 = 0.9496223 OK

```
user  system elapsed
0.30   0.00   1.75
```

- Print the coefficients of the model

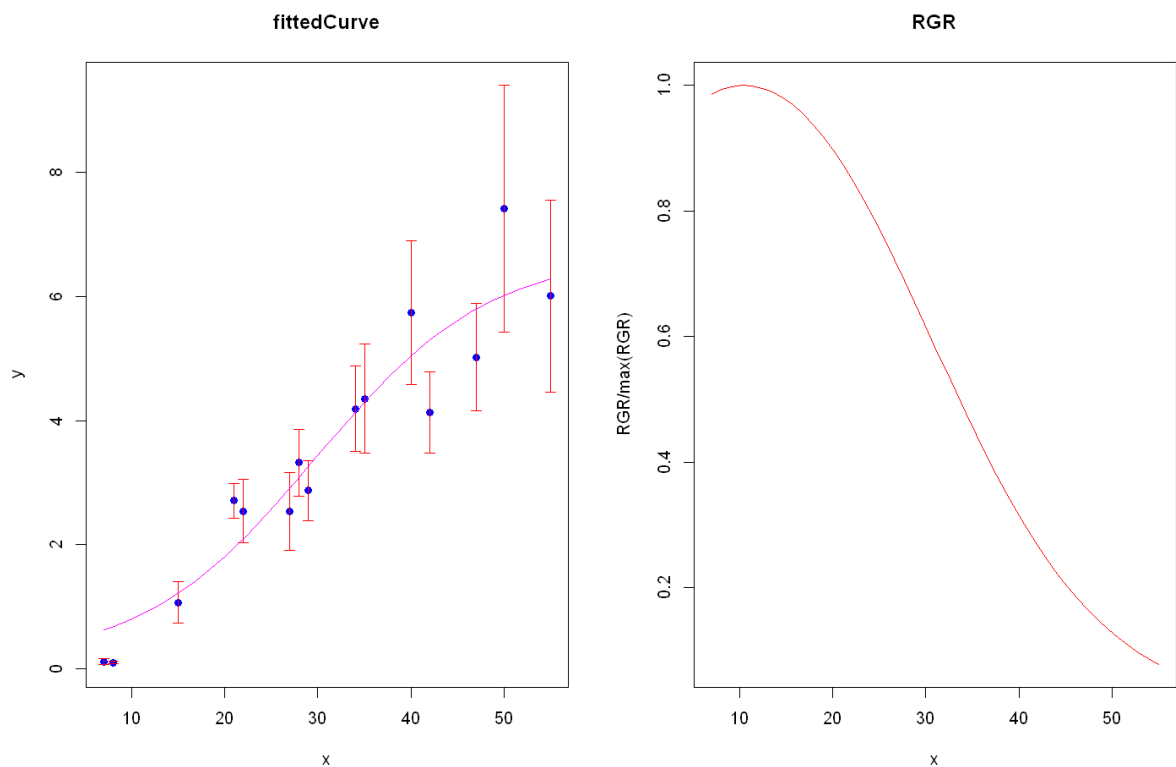
```
[7] print_fittedParams(fitObj)
```

Single Sigmoid

a                    b                    c                    d

- Plot both 'RGR' & 'Weigth' curves

```
[8] options(repr.plot.width=12, repr.plot.height=8)
par(mfrow=c(1,2))
plot_fittedCurve(fitObj)
plot_RGRCurve(fitObj)
```



## Explanation of the Relative Growth Rate (RGR)

Relative growth rate is a standardised measure of growth with the benefit of avoiding, as far as possible, the inherent differences in scale between contrasting organisms so that their performances can be compared on an equitable basis

- The **Relative Growth Rate (RGR)** of a plant at an instant in time (t) is defined as **the increase of plant material per unit of material present per unit of time.**

- The equation is written as:  $RGR = \frac{1}{W} \frac{dW}{dt}$  - W stands for Weight
- To calculate the RGR curve, we will leverage on the previous modelisation of the Weight curves (i.e. fitObj object obtained by the fitSigmoid function)

- Then we will **apply linear modeling** using the 'cv.glmnet' in order to **determine which variable(s) could explain the fruit growth.**
- **References**
  - Arne Pommerening & Anders Muszta (2015) Methods of modelling relative growth rate, Forest Ecosystems - [doi.org/10.1186/s40663-015-0029-4](https://doi.org/10.1186/s40663-015-0029-4)
  - Beauvoit et al (2014) Model-Assisted Analysis of Sugar Metabolism throughout Tomato Fruit Development Reveals Enzyme and Carrier Properties in Relation to Vacuole Expansion, Plant Cell, American Society of Plant Biologists, epub ahead of print. 10.1105/tpc.114.127761 [hal-01058814](https://doi.org/10.1105/tpc.114.127761)

## Linear Modeling with cv.glmnet

**run\_cvglmnet** :Application of linear modeling using the 'cv.glmnet' function of the R glmnet package

- for the Relative Growth Rate (RGR)
- using the models previously fitted by the fitSigmoid function (fitObj)
- by selecting some one or more data subsets (setNameList - the data subsets will be merged if several)
- $\alpha=1 \Rightarrow$  lasso penalty,  $\alpha=0 \Rightarrow$  ridge penalty,  $0 < \alpha < 1 \Rightarrow$  elastic-net penalty

### GLMNET

- <https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>
- <https://cran.r-project.org/web/packages/glmnet/vignettes/glmnet.pdf>

```
[9] # List of data subset names for explaining the RGR
    setNameList <- c('activome') # qNMR_metabo

    system.time( fitRGR <- run_cvglmnet(fitObj, dh, Tname,
    setNameList, alpha=1, info=TRUE) )
```

```
Getting & Merging ...
Fitting ... R2 = 0.9994468 OK

    user  system elapsed
    0.39   0.01   2.13
```

- **CV Glnet : RGR = F('activome')**

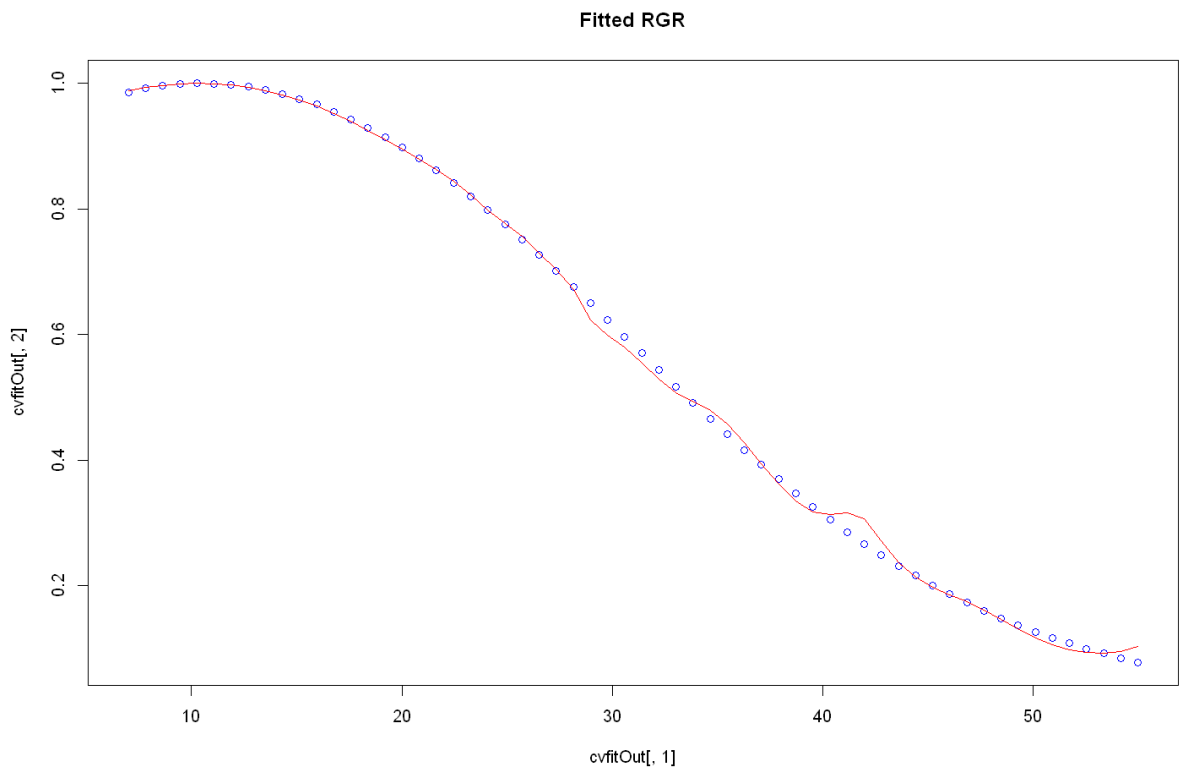
```
[10] print.table( as.matrix(fitRGR$cvfitList), digits=4,
    zero.print=".", na.print='- ' )
```



	1
(Intercept)	5.730e-01
PGM	-1.040e-01
cFBPase	2.321e-03
PyrK	.
CitS	.
PFP	.
Aconitase	.
PFK	-6.313e-03
FruK	.
pFBPase	.
GluK	1.687e-01
NAD_ISODH	.
Enolase	2.971e-01
NADP_ISODH	-9.027e-05
PEPC	.
Aldolase	-3.589e-02
Succ_CoA_ligase	2.045e-01
NAD_MalDH	1.836e-01
AlaAT	1.827e-03
Fumarase	.
AspAT	.
NADP_GluDH	.
NAD_GAPDH	.
NADP_GAPDH	-5.692e-03
NAD_GluDH	-1.480e-01
TPI	.
PGK	.
Neutral_Inv	-1.623e-02
Acid_Inv	3.795e-02
G6PDH	-2.069e-01
UGPase	.
SuSy	.
NAD_ME	1.495e-02
ShiDH	.
NADP_ME	.
PGI	-9.186e-02
StarchS	6.404e-03
AGPase	.
SPS	-8.475e-02

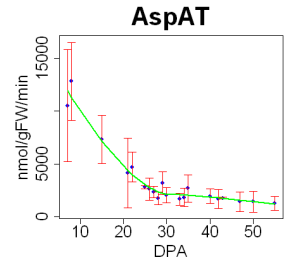
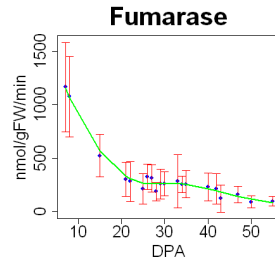
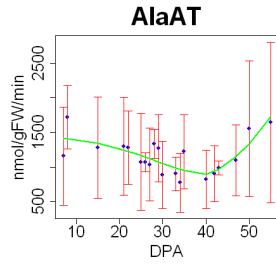
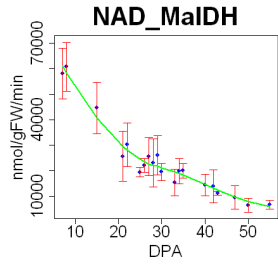
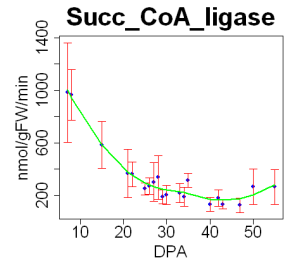
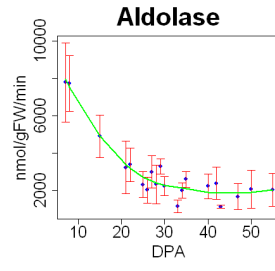
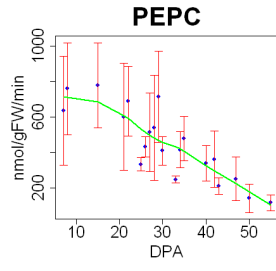
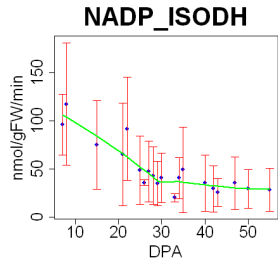
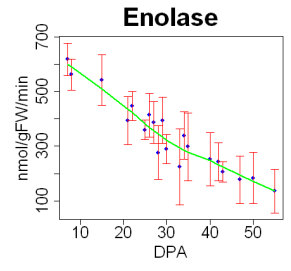
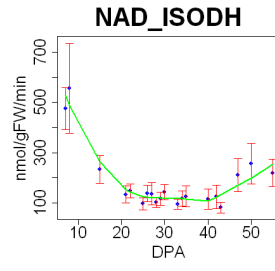
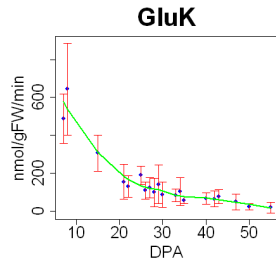
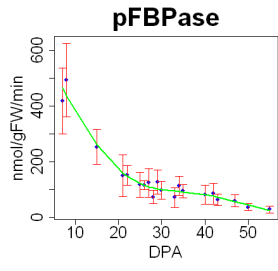
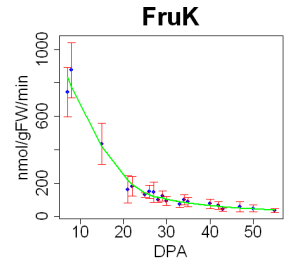
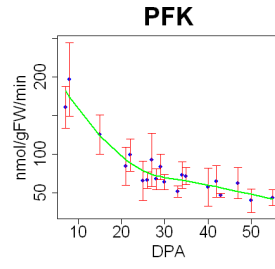
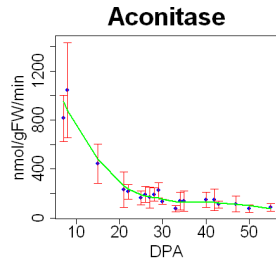
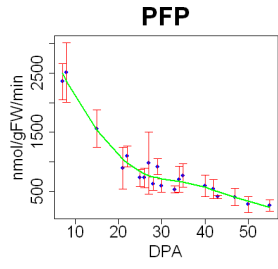
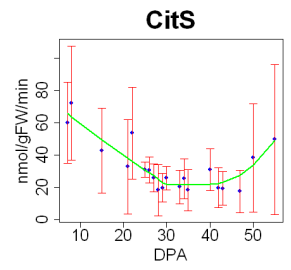
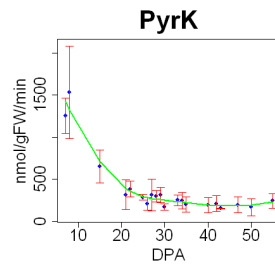
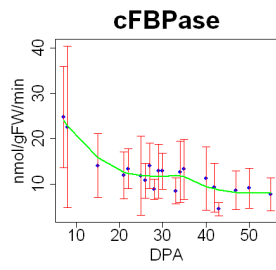
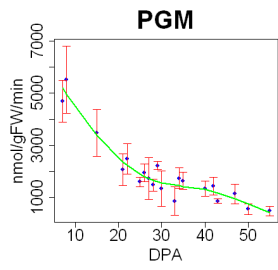
- 'RGR' fitted curve - activome

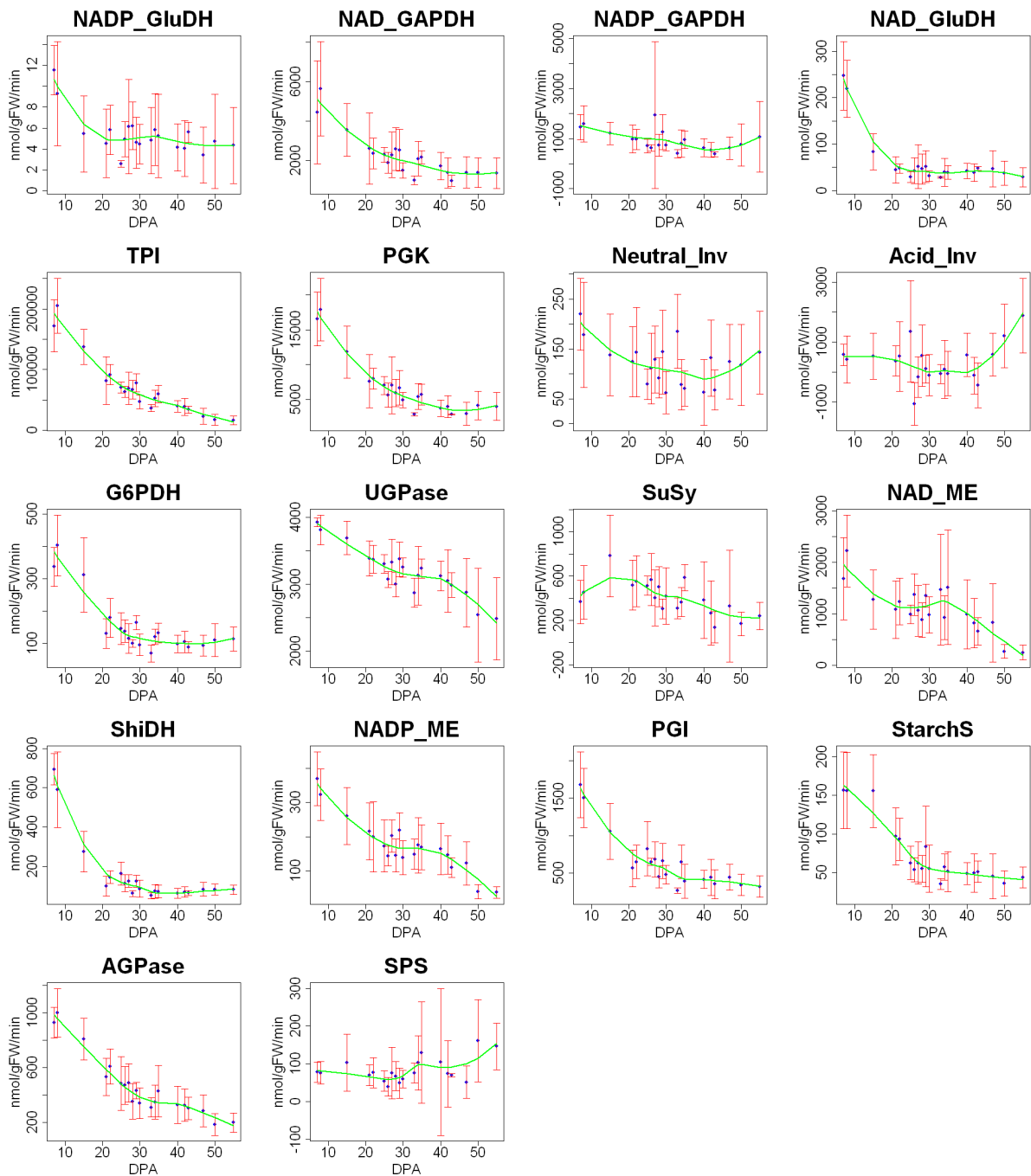
```
[11] options(repr.plot.width=12, repr.plot.height=8)
      plot_fittedRGRCurve(fitRGR$cvfitOut)
```



## Plot variables

```
[12] options(repr.plot.width=16, repr.plot.height=18)
      plot_vars(dh, Tname, setNameList[1], smoothtype='lowess',
               Gmax=20, ncol=4, cex.axis=2, cex.lab=2, cex.main=3)
```





## Session Information

```
[13] options(width=80)
      sessionInfo()
```

```
R version 3.5.1 (2018-07-02)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 18363)
```

```
Matrix products: default
```

locale:

```
[1] LC_COLLATE=French_France.1252 LC_CTYPE=French_France.1252
[3] LC_MONETARY=French_France.1252 LC_NUMERIC=C
[5] LC_TIME=French_France.1252
```

attached base packages:

```
[1] parallel stats graphics grDevices utils datasets methods
[8] base
```

other attached packages:

```
[1] UpSetR_1.3.3 Rodam_0.1.6 RCurl_1.95-4.11
[4] bitops_1.0-6 glmnet_2.0-16 foreach_1.4.4
[7] Matrix_1.2-14 minpack.lm_1.2-1 limma_3.36.3
[10] gplots_3.0.1 pcaMethods_1.72.0 Biobase_2.40.0
[13] BiocGenerics_0.26.0
```

loaded via a namespace (and not attached):

```
[1] Rcpp_1.0.3 lattice_0.20-35 tidy_0.8.3
[4] visNetwork_2.0.4 gtools_3.8.1 zeallot_0.1.0
[7] assertthat_0.2.1 digest_0.6.22 IRdisplay_0.7.0.9000
[10] R6_2.3.0 plyr_1.8.4 repr_1.0.1.9000
[13] backports_1.1.5 evaluate_0.14 ggplot2_3.2.1
[16] pillar_1.4.2 rlang_0.4.1 lazyeval_0.2.1
[19] uuid_0.1-2 rstudioapi_0.10 gdata_2.18.0
[22] DiagrammeR_1.0.0 downloader_0.4 readr_1.1.1
[25] stringr_1.4.0 htmlwidgets_1.5.1.9000 igraph_1.2.2
[28] munsell_0.5.0 compiler_3.5.1 influenceR_0.1.0
[31] rgexf_0.15.3 pkgconfig_2.0.2 base64enc_0.1-3
[34] htmltools_0.4.0.9000 tidyselect_0.2.5 tibble_2.0.1
[37] gridExtra_2.3 codetools_0.2-15 XML_3.98-1.16
[40] viridisLite_0.3.0 crayon_1.3.4 dplyr_0.8.0.1
[43] grid_3.5.1 jsonlite_1.6 gtable_0.2.0
[46] magrittr_1.5 scales_1.0.0 KernSmooth_2.23-15
[49] stringi_1.4.3 viridis_0.5.1 vctrs_0.2.0
[52] brew_1.0-6 IRkernel_1.0.2.9000 data.tree_0.7.8
[55] RColorBrewer_1.1-2 iterators_1.0.10 tools_3.5.1
[58] glue_1.3.1 purrr_0.2.5 hms_0.4.2
[61] Rook_1.1-1 colorspace_1.3-2 caTools_1.17.1.1
[64] pbdZMQ_0.3-3
```



[ ]