2015_IG01_Julio.pdf

Association mapping in a collection of tobacco reference cultivars.

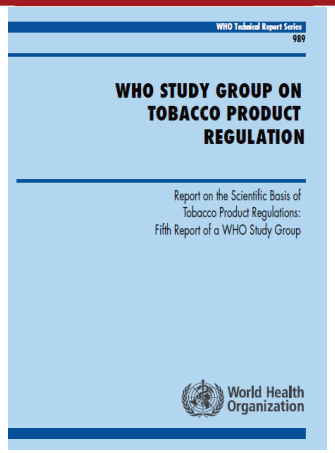
Step One: Variability of smoke constituents

Emilie JULIO, Thomas VERRON, Xavier CAHOURS, François DORLHAC de BORNE and Stéphane COLARD

AP2015 - Document not peer-reviewed by CORESTA

Context

WHO Study Group on Tobacco Product Regulation (5th report, 2015)



« Non-exhaustive priority list of tobacco and smoke constituents»

The list of tobacco contents and emissions of cigarette smoke was drawn up on the basis of the following criteria:

- the presence of specific chemicals in cigarette smoke at levels that are toxic for smokers as determined by well-established scientific toxicity indices;
- variations in concentrations among cigarette brands that are substantially greater than the variation in repeated measurements of the toxicant in a single brand;
- the existence of technological routes to mitigate these yields.

Are there constituents that can be reduced by modifying the tobacco plant ?

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

Are there constituents that can be reduced by modifying the tobacco plant ?

- **Many constituents in tobacco and tobacco smoke are influenced by the genetics of the tobacco plant**, or might be affected by processes occurring in the plant during growing, senescence, curing, as well as blending and pyrolysis.
- Molecular markers to identify and follow these genes could be used in traditional breeding in order to develop new plants that result in lower “toxicants” both in tobacco and in cigarette smoke.

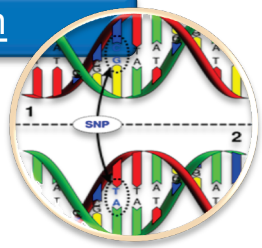
Objective

Identify the **direct links between gene variability and chemical variability** observed in tobacco and/or in tobacco smoke using **genetic association methodology .**

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

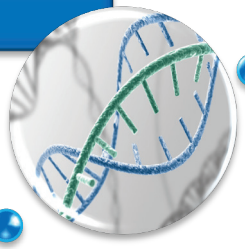
Methodology of association mapping

SNP detection

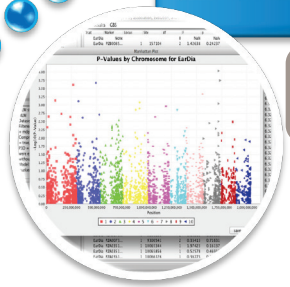


mRNA sequencing

Extraction of mRNA from leaves, stems and roots



Association mapping



Statistical analysis
GLM/MLM

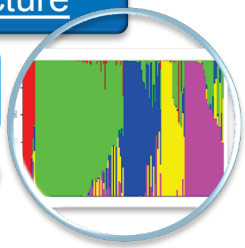
The germplasm group

162 Tobacco varieties

69	Dark air-cured
24	Burley
46	Flue-cured
16	Oriental
6	Experimental
1	« Ancestral »

Population Structure

Structure + Kinship (SSR/SNP)



Smoke analysis

- **Aromatic amines**
  (1-AMN, 2AMN 3-ABP, 4-ABP)
- **BaP**
- **Volatiles**
  (1-3 butadiene, isoprene, acrylonitrile, benzene, toluene)
- **Carbonyles**
  (Formaldehyde, acetaldehyde, acetone, acrolein, crotonaldehyde, butyraldehyde, propionaldehyde)
- **Hydrogen cyanide**
- **Phenols**

Tobacco analysis

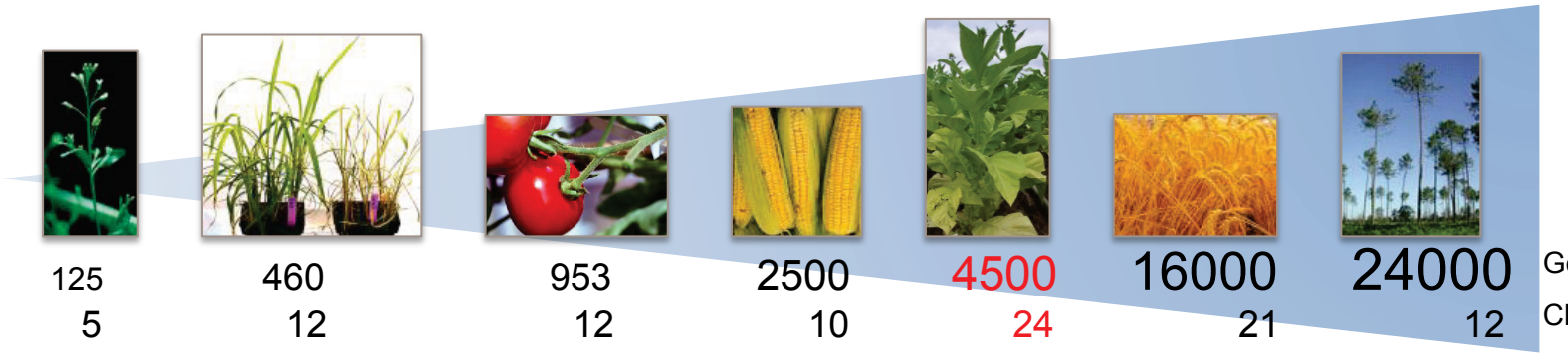
Cigarette making

Hand made cigarettes
Tobacco weight: 750 mg +- 5mg.

2015_IG01_Julio.pdf
AP2015 - Document not peer-reviewed by CORESTA

Strategy for tobacco: RNA-Seq

Tobacco has a large genome compared to other species



- Whole genome sequencing would be too expensive !
- Advantages of RNA-Seq:
  - most repetitive regions are ignored: cost efficiency
  - Simultaneous SNPs (Single Nucleotide Polymorphism) detection and expression profile analysis
  - Also useful for candidate gene approach
- Pre-requisite
  - Reference transcriptome available
  - RNA preparation of multiple tissues at multiple time points

2015_IG01_Julio.pdf

AP2015 - Document not peer-reviewed by CORESTA

2015_IG01_Julio.pdf

Step One: Variability of smoke constituents on a panel of Flue-cured.

1. Experimental design
2. Statistical approach
3. Preliminary results
4. Next steps

The germplasm group

162 Tobacco varieties

69	Dark air-cured
24	Burley
46	Flue-cured
16	Oriental
6	Experimental
1	« Ancestral »

Tobacco analysis

Cigarette making

Hand made cigarettes
Tobacco weight: 750 mg +- 5mg.
120 cig./sample

Smoke analysis

- **Aromatic amines**
  (1-AMN, 2AMN 3-ABP, 4-ABP)
- **BaP**
- **Volatiles**
  (1-3 butadiene, isoprene, acrylonitrile, benzene, toluene)
- **Carbonyles**
  (Formaldehyde, acetaldehyde, acetone, acrolein, crotonaldehyde, butyraldehyde, propionaldehyde)
- **Hydrogen cyanide**
- **Phenols**

AP2015 - Document not peer-reviewed by CORESTA

Experimental Design

Flue-cured (46 varieties)

3 blocks

3 squares per block

3 bands per square

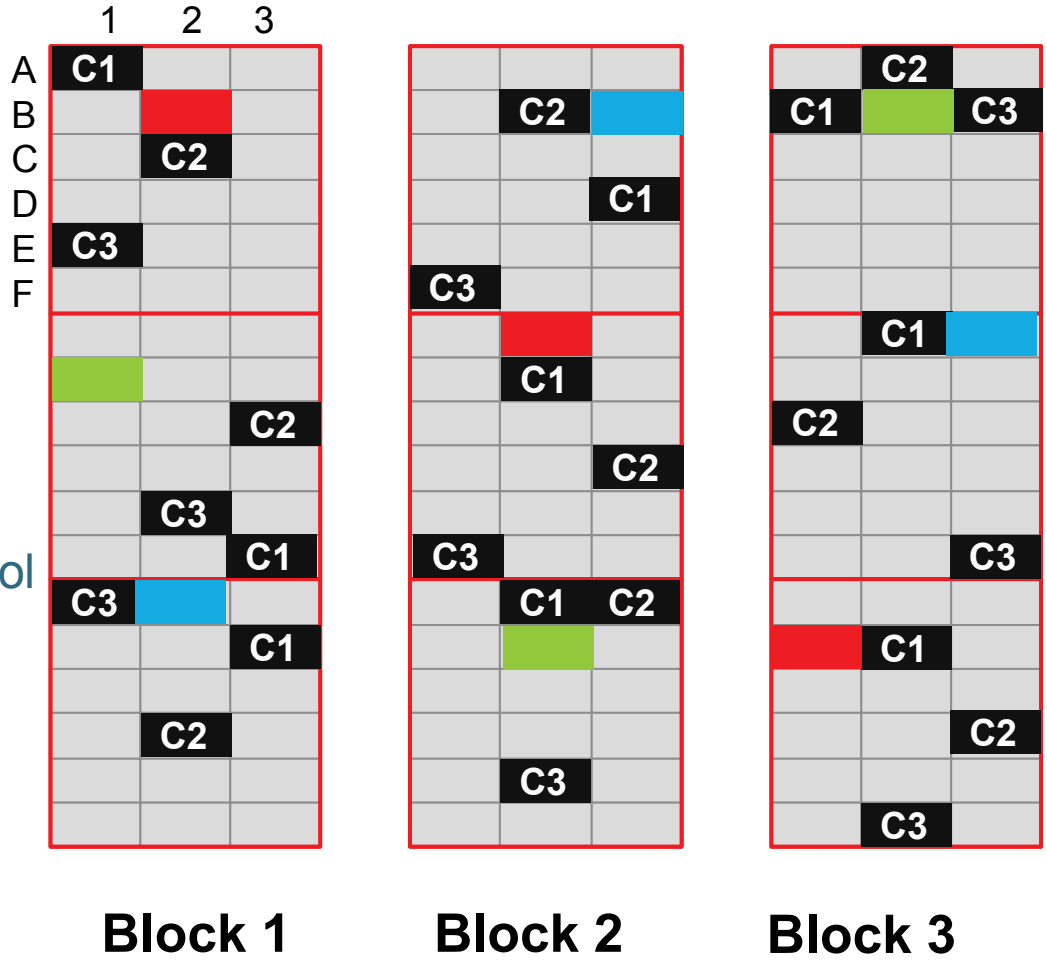
6 lines per square

To take into account environmental factors:
3 varieties are used as control

3 controls per square

C1 C2 C3

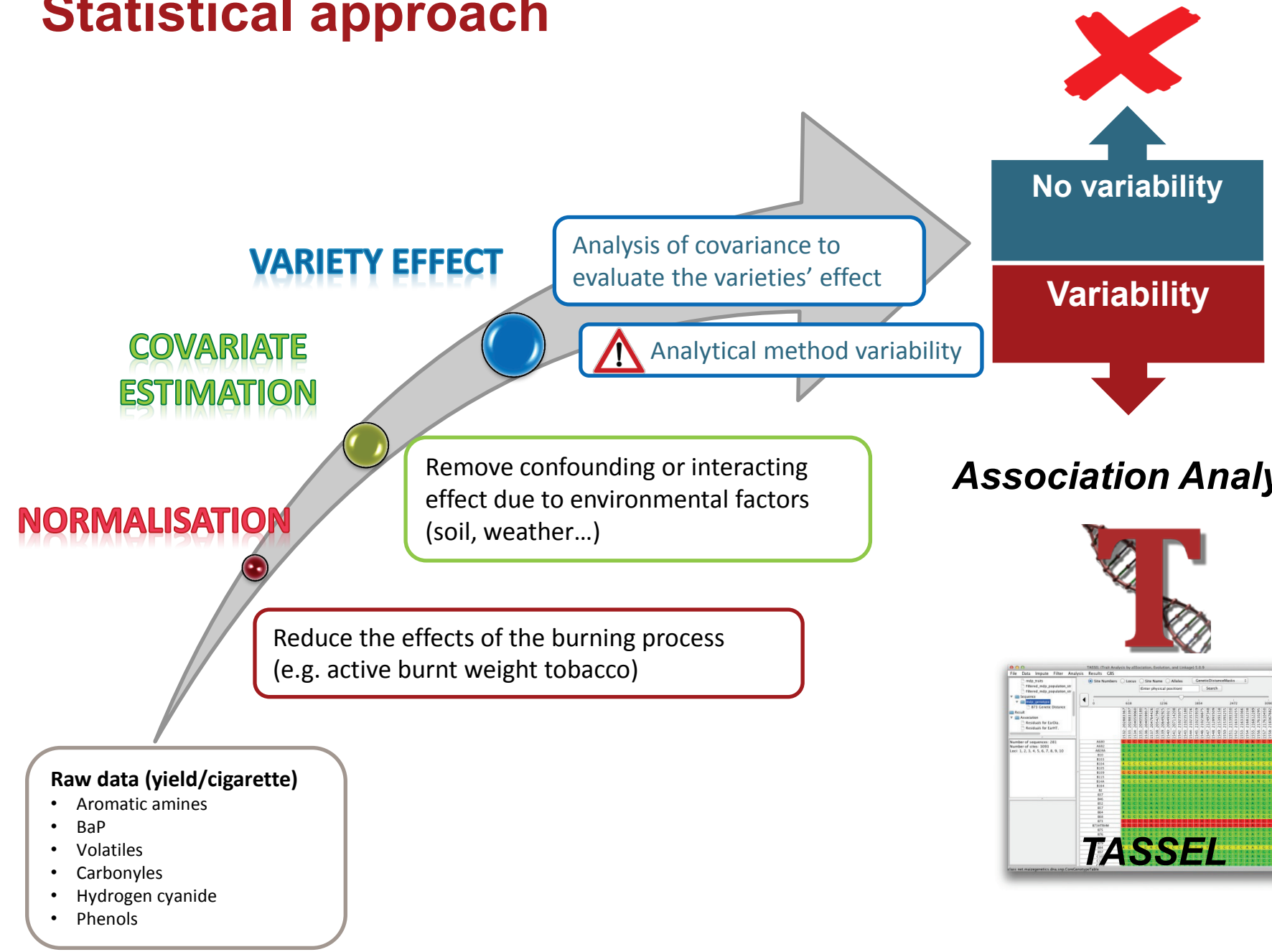
3 replicates for the other varieties (1 per block)



2015_IG01_Julie.pdf

AP2015 - Document not peer-reviewed by CORESTA

Statistical approach



Association Analysis

2015_IG01_Julio.pdf

AP 2015 - Document not peer-reviewed by CORESTA

Normalisation

To eliminate variation related to burning process

Active burnt weight tobacco estimation*

Smoking time

Smoking time SR1 = 315.9 s
Puff number SR1 = 10.95
Smoking time SR2 = 405.8 s
Puff number SR1 = 7.25

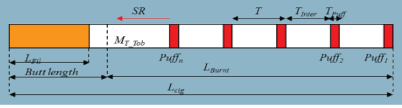
Puff number

Puff duration SR1 = 2s
Puff duration SR2 = 2s
Puff freq SR1 = 30s
Puff freq SR2 = 60s
Puff Volume SR1 = 55 mL
Puff Volume SR2 = 35 mL

Cig. param.

Butt length = 27 mm
Filter lenght = 15 mm
Tip length = 24 mm
Tob weight = 750 mg
Filter ventilation = 0

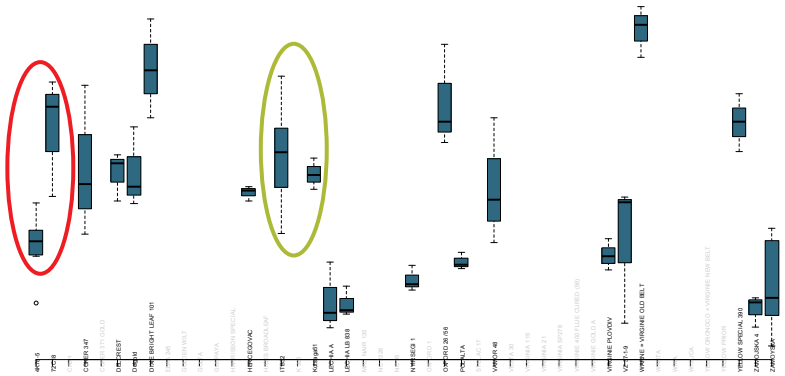
Burning model ISO/HCI



Active burn weight tobacco
SR1: 253.6 mg
SR2: 132.5 mg

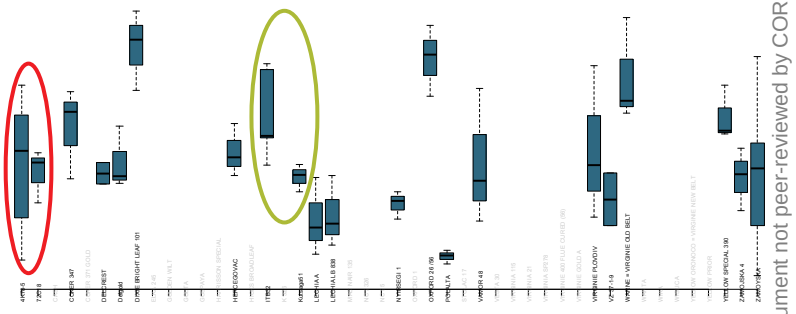
*S. Colard. 2015.
Prediction Model for Cigarette Yields Derived from Data Obtained under Two Different Machine Smoking Regimes.
Beitrage zur Tabakforschung International/Contributions to Tobacco Research. issue No. 7 of Vol. 26.

BaP (raw data)



Normalization

BaP (Normalized)
= BaP / actBurntW

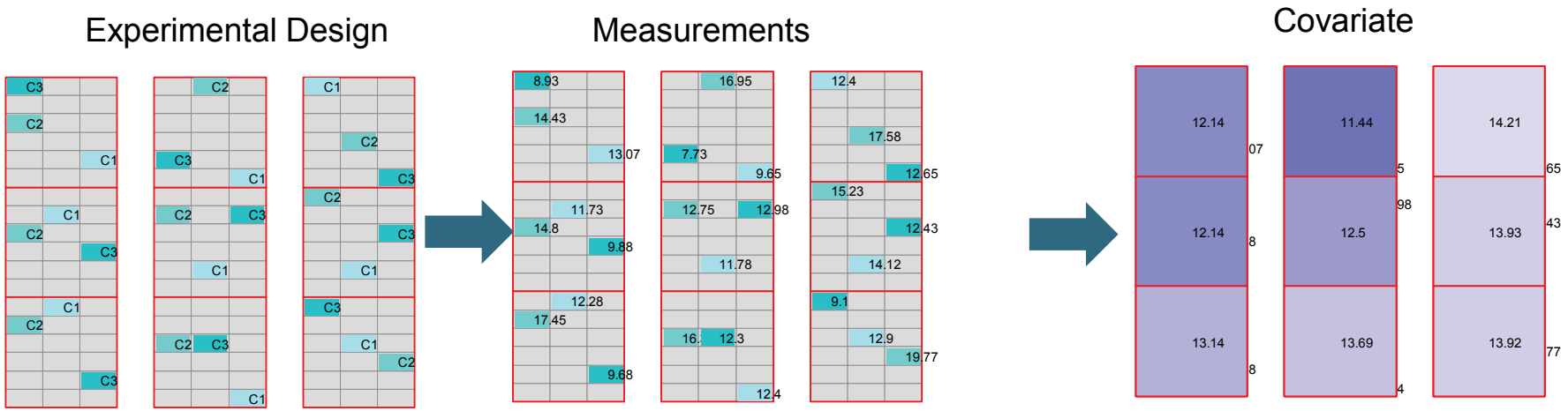


2015_IG01_Julio.pdf

AP 2015 - Document not peer-reviewed by CORESTA

Covariate Estimation

Remove confounding effect due to environmental factors



3 Controls:

C1
C2
C3

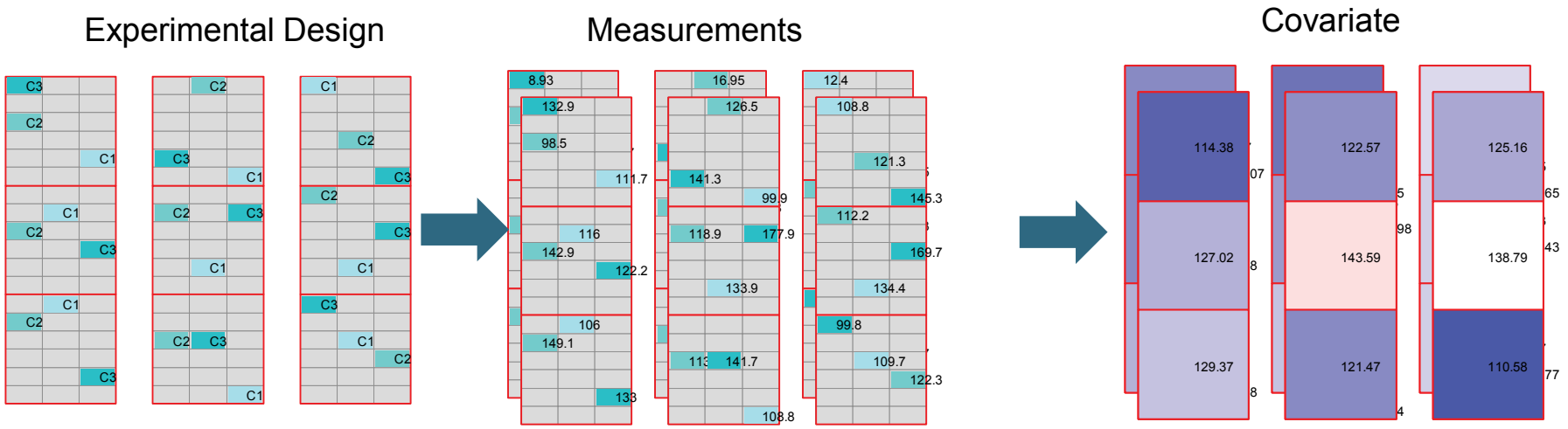
Base line effect:

Differences observed are considered to be the consequence of environmental factors

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

Covariate Estimation

Remove confounding effect due to environmental factors



3 Controls:

C1
C2
C3

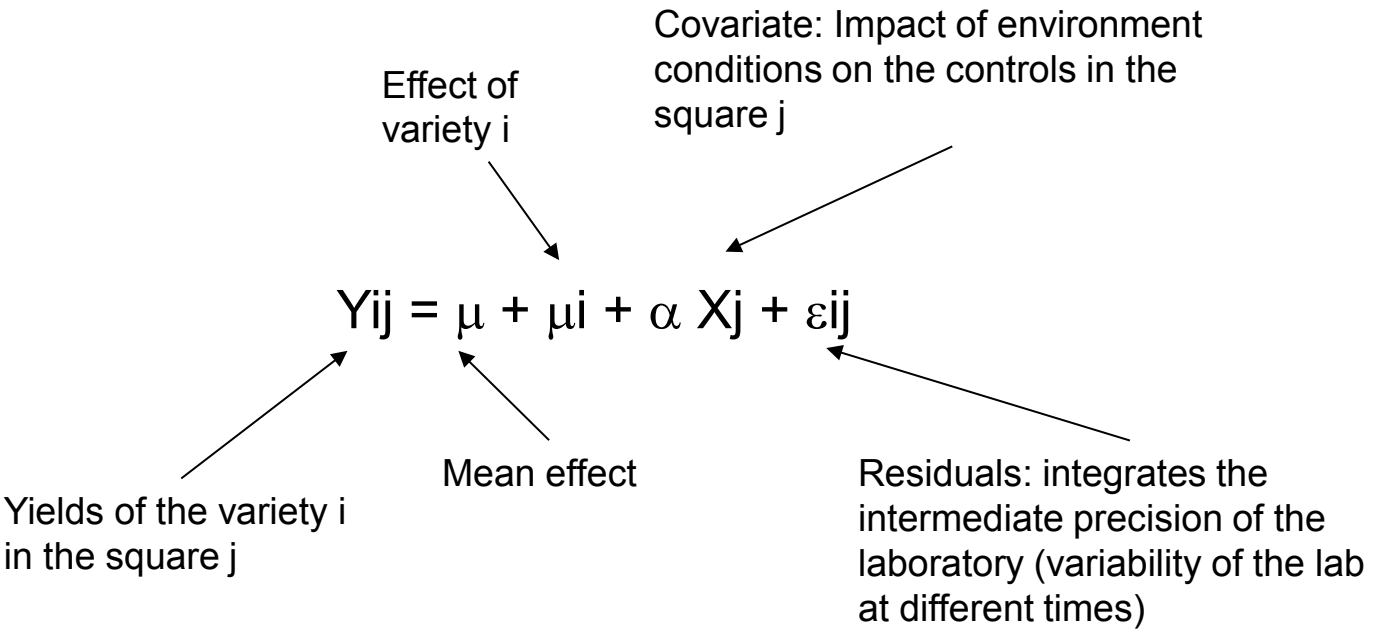
Base line effect:

Differences observed are considered to be the consequence of environmental factors

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

Variety effect: ANCOVA

Evaluate the varieties' effect while statistically controlling for the covariable effect



	DoF	Sum of sq.	Mean Sq.	F value	P-value
Covariable	1	1394.9	1394.88	7.2102	0.011005
Variety	17	11355.0	667.94	3.4526	0.000942
Residuals	35	6771.1	193.46		

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

Preliminary results

21 flue-cured varieties

21 varieties, 3 repeats + control
= 81 samples of cigarettes

ANALYSES
Aromatic amines 1-aminonaphtalene (1-AMN), 2-aminonaphtalene (2-AMN), 3-aminobiphenyl (3-ABP), 4-aminobipheny (4-ABP)
Carbonyles Formaldehyde, acetaldehyde, acetone, acrolein, crotonaldehyde, butyraldehyde, propionaldehyde,)
Benzo[a]Pyrene (BaP)
Volatiles 1-3 butadiene, isoprene, acrylonitrile, benzene, toluene
Hydrogen cyanide
Phenols Hydroquinone, resorcinol, catechol, phenol, gaiacol, m/p/o cresol
TNCO

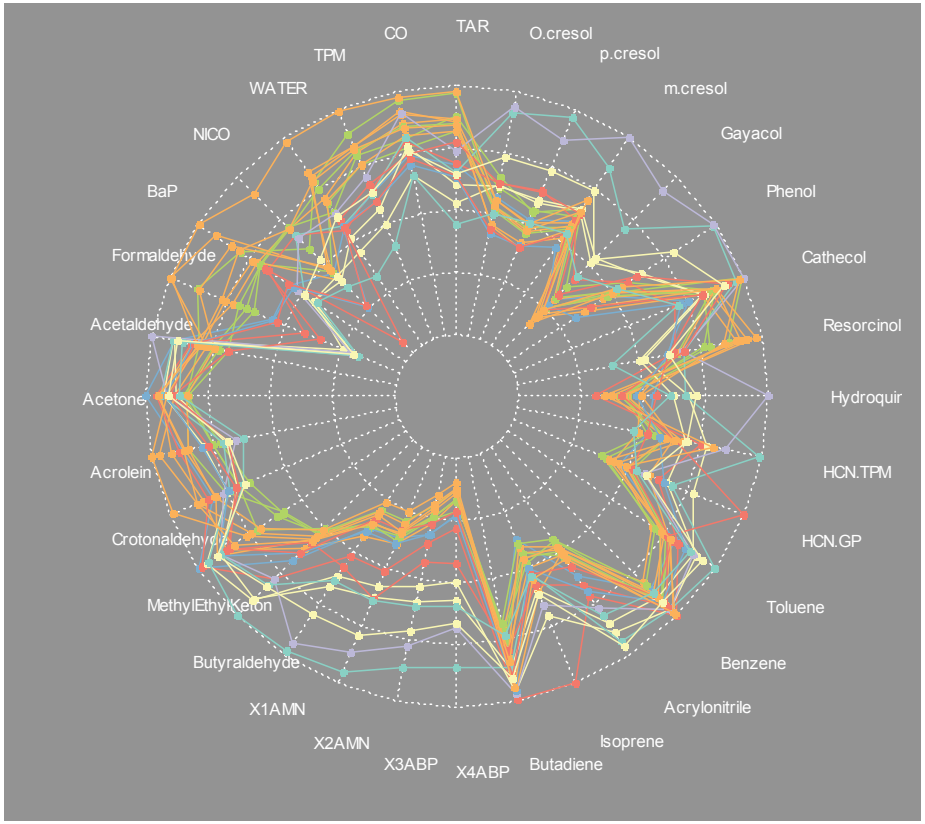
~ 6000 cigarettes

2015_IG01_Julio.pdf

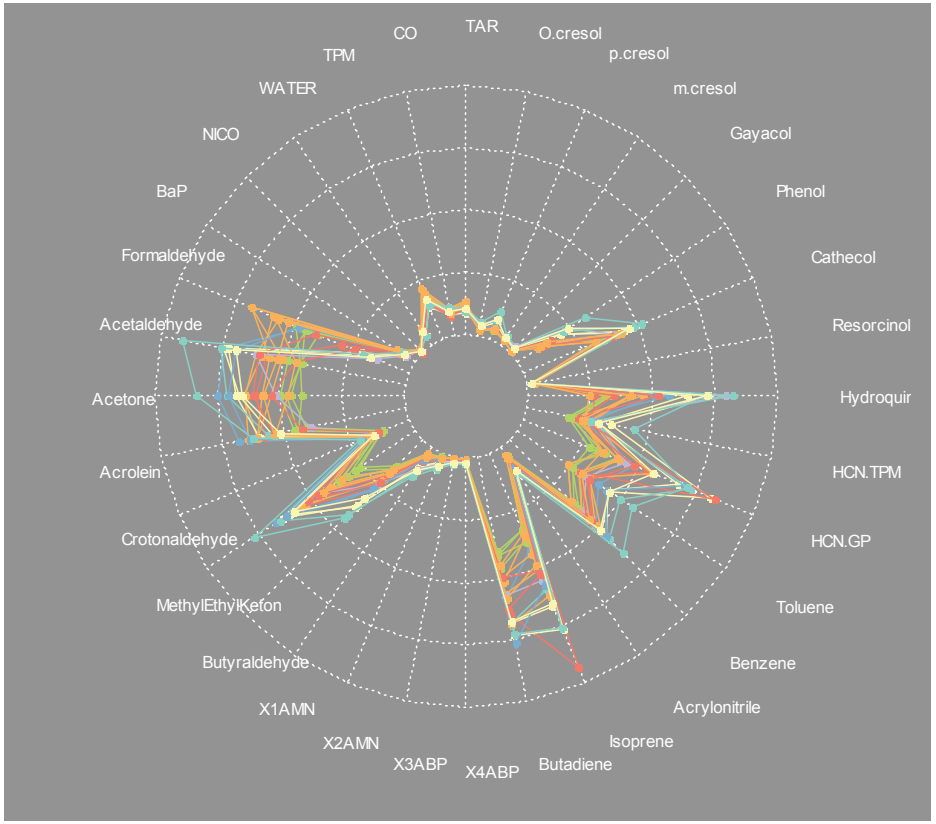
AG2015 - Document not peer-reviewed by CORESTA

Normalisation

Raw data



Normalized data



Normalisation lowers variability between varieties and allows accurate comparison based on mg of tobacco burnt

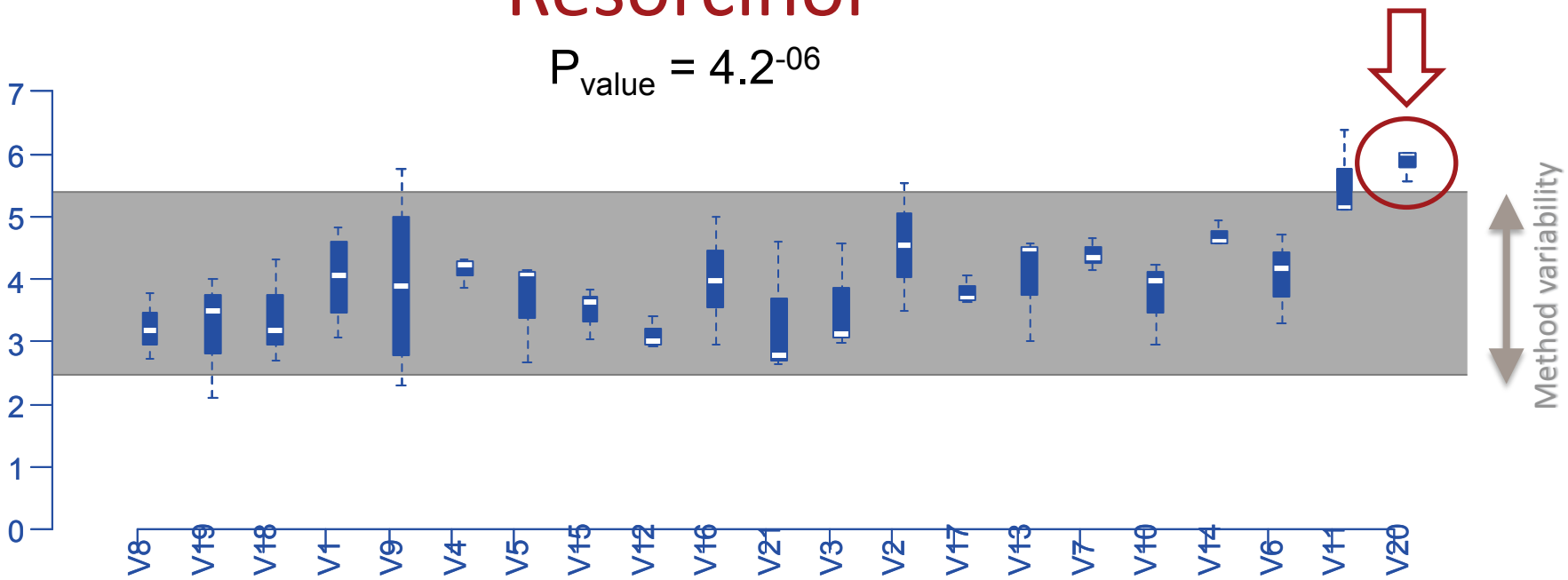
2015_IG01_Julio.pdf

AT 2015 - Document not peer-reviewed by CORESTA

Variability of smoke constituents

Resorcinol

$P_{value} = 4.2^{-06}$



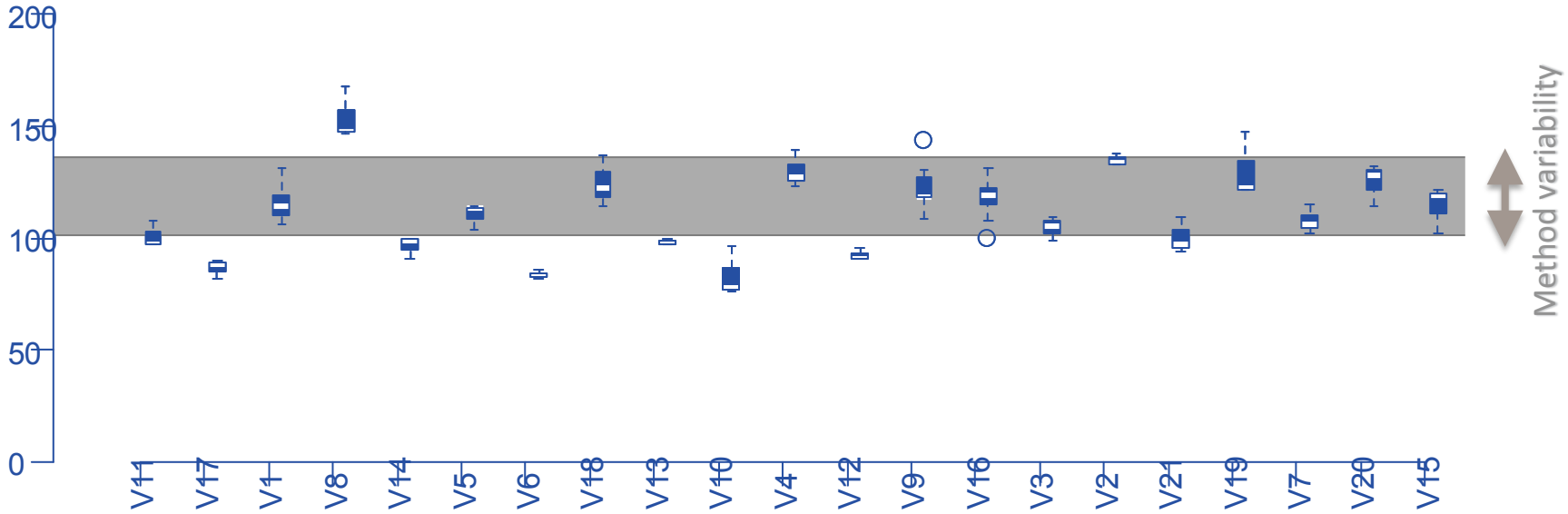
= no variability between varieties

2015_IG01_Julio.pdf

AP 2015 - Document not peer-reviewed by CORESTA

Variability of smoke constituents

Benzene

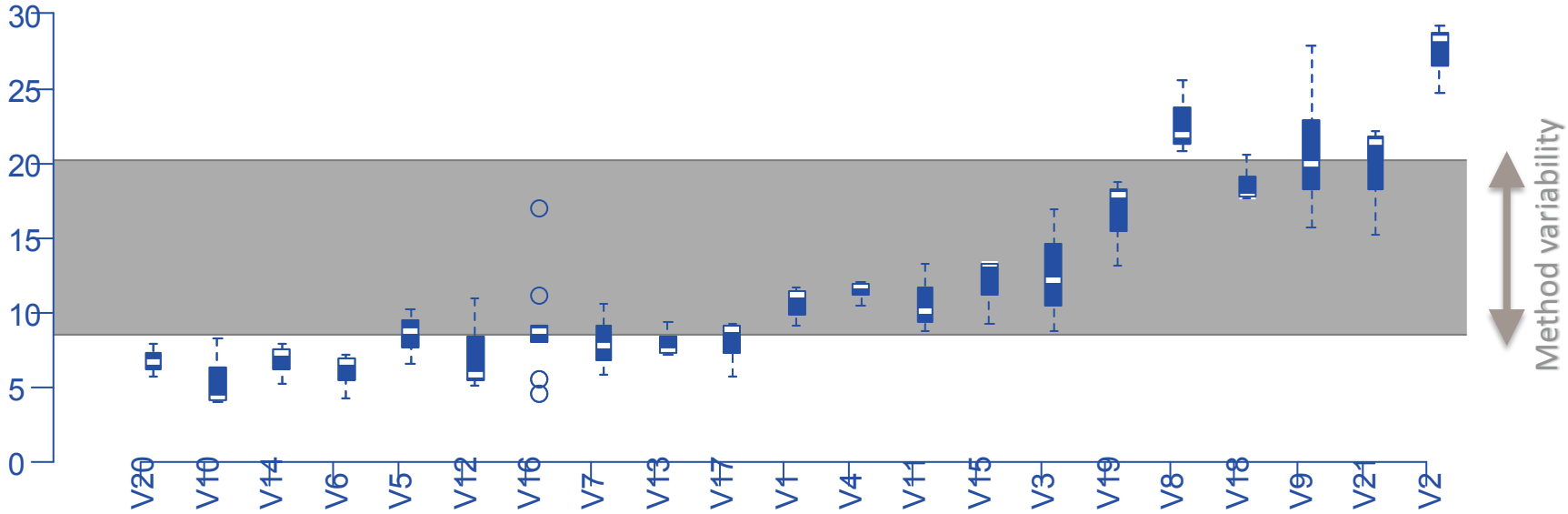


= low variability between varieties

2015_IG01_Julio.pdf
AP 2015 - Document not peer-reviewed by CORESTA

Variability of smoke constituents

1-aminonaphthalene



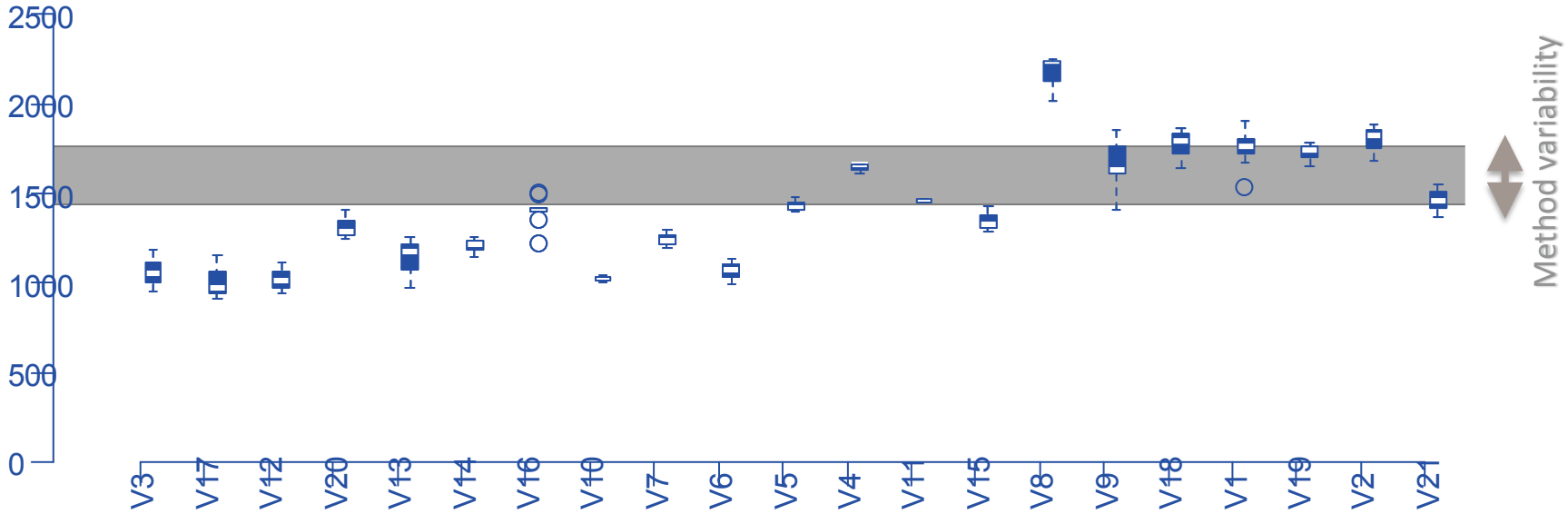
= Medium variability between varieties

2015_IG01_Julio.pdf

AP2015 - Document not peer-reviewed by CORESTA

Variability of smoke constituents

Acetaldehyde



= High variability between varieties

2015_IG01_Julio.pdf

AP 2015 - Document not peer-reviewed by CORESTA

Variability of smoke constituents

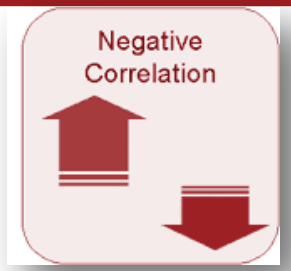
2015_IG01_Julio.pdf

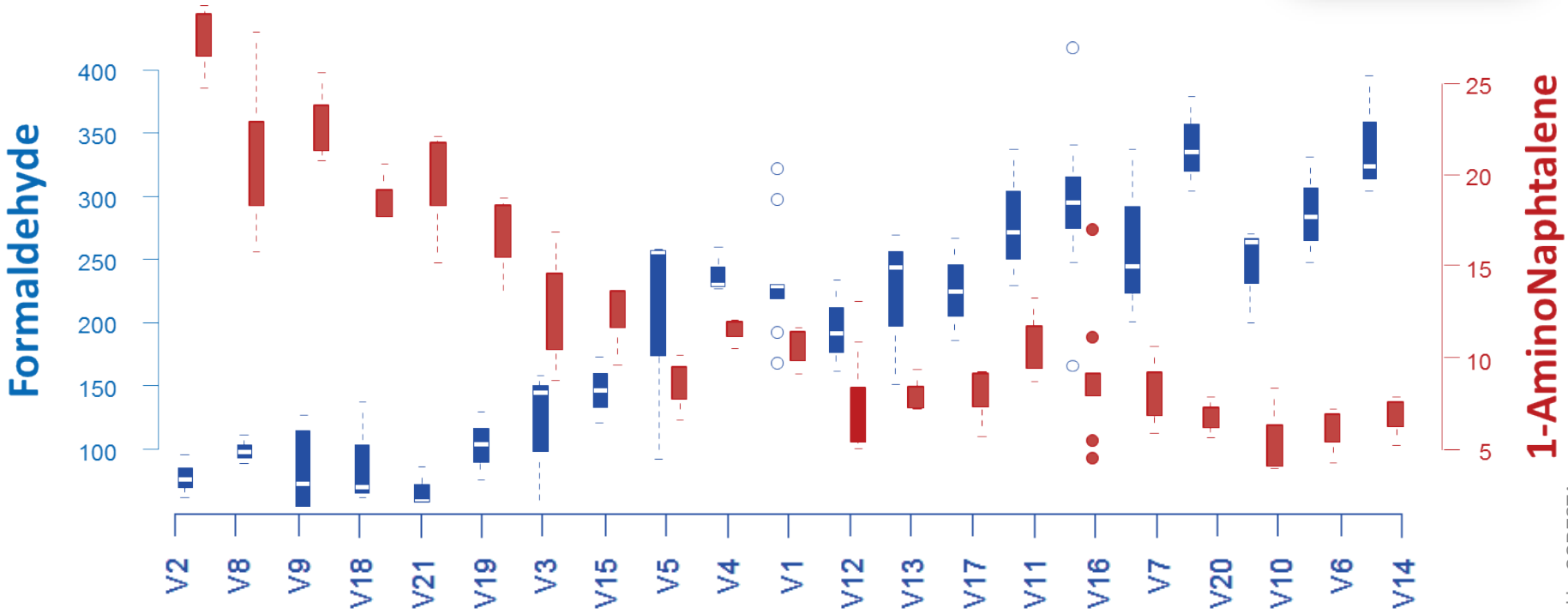
Phenols	Volatiles	Aromatic Amines	Volatiles	Carbonyles
CO	1,3-Butadiene	Isoprene	1AMN	Formaldehyde
HCN	Toluene	Benzene	2 AMN	Butyraldehyde
o, m, p.cresol		BaP	3ABP	Acetaldehyde
Cathecol	Phenol	Hydroquinone	4ABP	Acrylonitrile
Resorcinol	O.cresol	Gayacol	Acrolein	MethylEthylKeton
			Crotonaldehyde	
			Acetone	

LOW —— Variability between varieties —— **HIGH**

AT 2015 - Document not peer-reviewed by CORESTA

Negative correlation





In some cases lowering the yield of one smoke constituent may result in increasing the yield of another.

2015_IG01_Julio.pdf
A 2015 - Document not peer-reviewed by CORESTA

Conclusion

1. Normalisation using active burnt weight tobacco allows an accurate comparison between tobacco varieties.

2. To avoid misleading conclusion:
   - Confounding effect due to environmental factors (soil, weather…) is corrected using tobacco controls in the fields.
   - Variability of analytical methods has to be added in the comparison.

2015_IG01_Julio.pdf

AP2015 - Document not peer-reviewed by CORESTA

Conclusion

3. From the preliminary results

- Wide variability range for certain smoke constituents
  (e.g. Formaldehyde, acetaldehyde, etc.).
  = molecular markers can be expected

- Low variability range for some constituents
  (e.g. 1,3-butadiene, catechol, etc.).
  = This has to be confirmed with other varieties.

- Negative correlations exist.
  = It is very difficult to predict the ultimate impact that reducing, or selectively eliminating, a specific smoke constituent may have.

2015_IG01_Julio.pdf
AP2015 - Document not peer-reviewed by CORESTA

Perspectives

- Smoke analyses for the 140 additional varieties are in progress.
- RNA-seq is completed.
- At least 6,000 SNPs have been detected into the transcriptome.
- Association mapping will be presented next.

2015_IG01_Julio.pdf
AG2015 - Document not peer-reviewed by CORESTA

Thank you for your attention

www.imperialtobaccoscience.com

2015_IG01_Julio.pdf
AP2015 - Document not peer-reviewed by CORESTA