

# Rapport EI Exposome

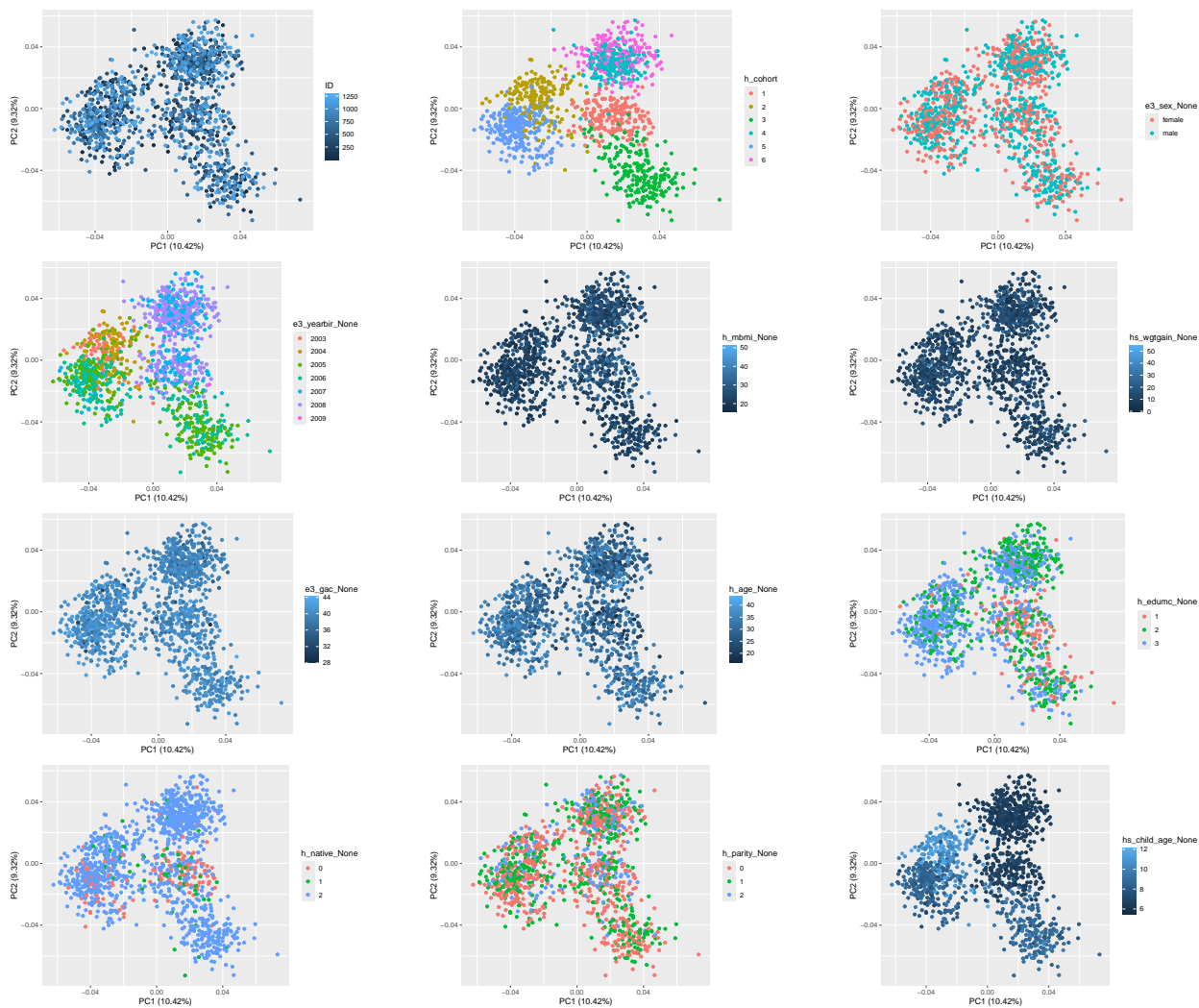
Groupe 5

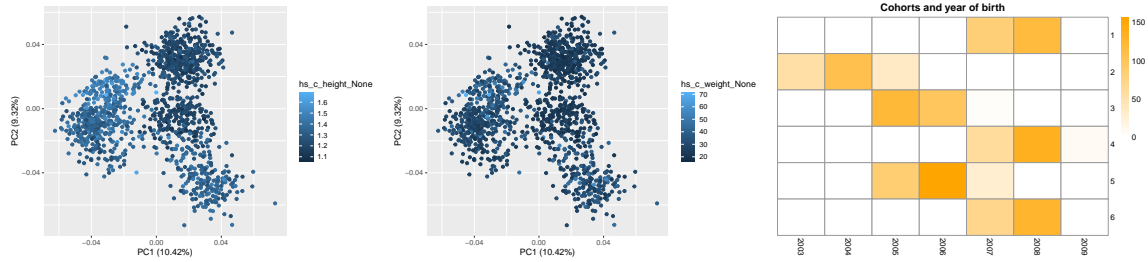
2024-06-06

## Détection des covariantes principales influant le modèle

Dans un premier pas, on s'est concentré sur les valeurs numériques de l'exposome.

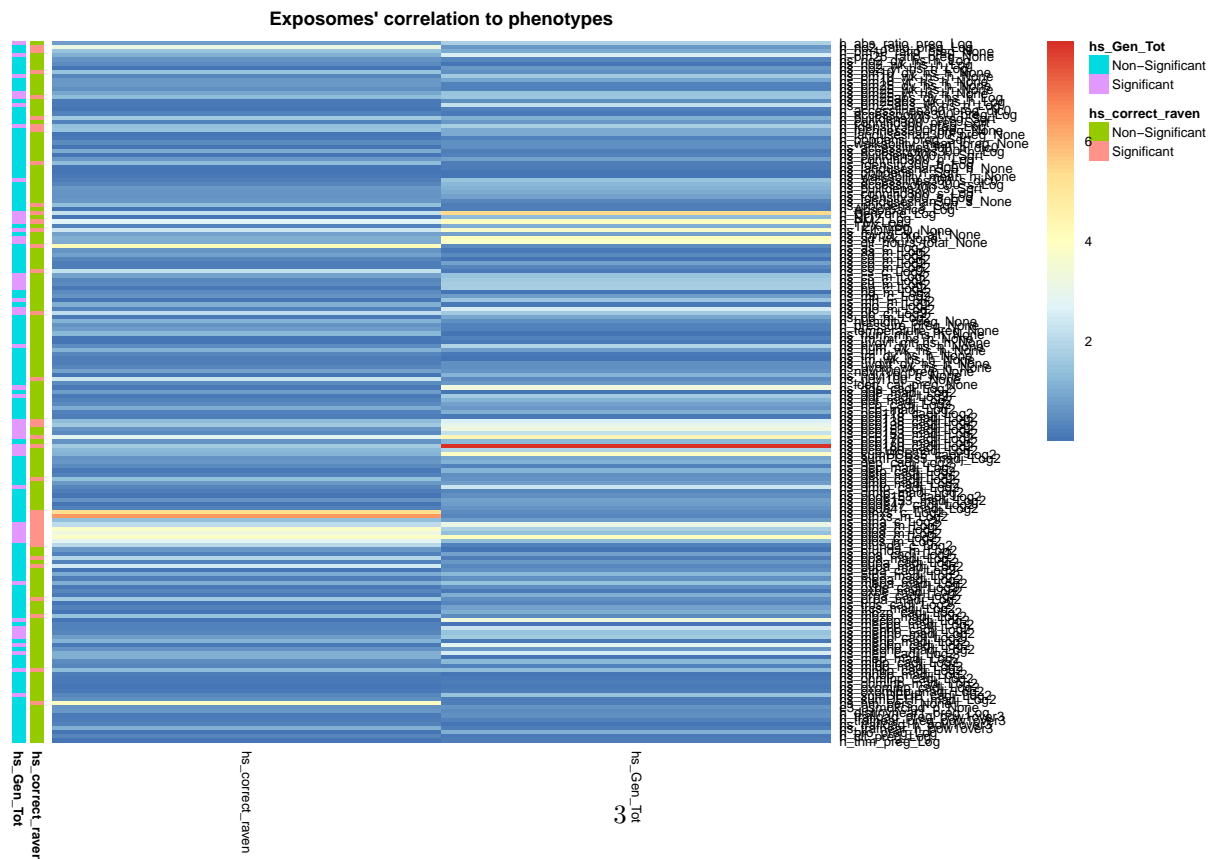
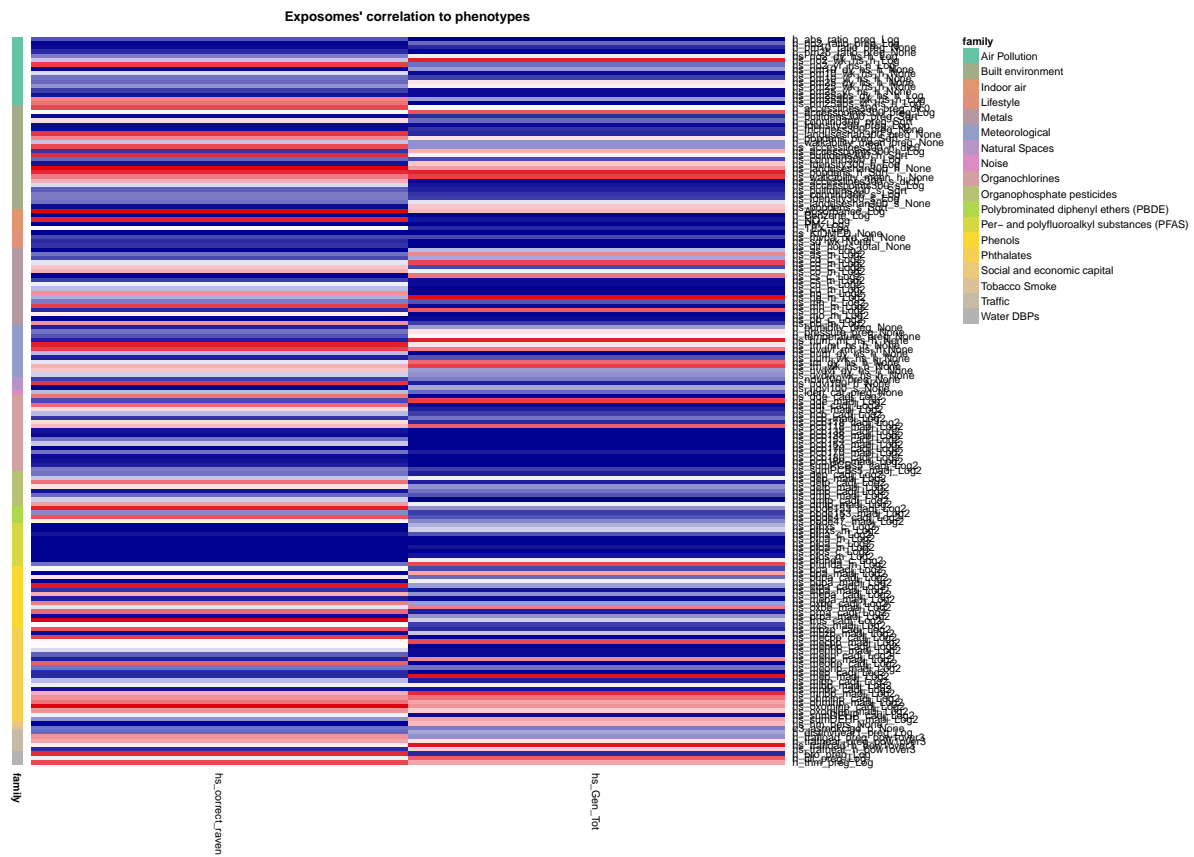
Ceci a été fait pour pouvoir effectuer une analyse de composantes principales sur les données obtenues et identifier les covariantes influant les différentes expositions pour savoir desquels il faut faire attention dans les analyses suivants.





Dans l'analyse des composantes principales on peut trouver des fortes corrélations liées à la variable `h_cohort` qui représente le pays auquel les données ont été collectées et au `e3_yearbir_None` qui représente l'année de naissance des enfants. Quand-meme, pour la suite du projet, nous ne prendrons en compte que la variable `h_cohort`. Apparemment, les cohortes ne sont pas créées en même temps car l'année de naissance est liée visiblement à la cohorte (chaque cohorte a certaines années de naissance spécifiques).

Matrice de “corrélations” (p\_values) entre les expositions et les différents phénotypes



##	hs_Gen_tot	V2
## h_abs_ratio_preg_Log	Significant	Air Pollution
## h_pm25_ratio_preg_None	Significant	Air Pollution
## hs_pm10_wk_hs_h_None	Significant	Air Pollution
## hs_pm25_yr_hs_h_None	Significant	Air Pollution
## hs_pm25abs_dy_hs_h_Log	Significant	Air Pollution
## hs_pm25abs_yr_hs_h_Log	Significant	Air Pollution
## h_fdensity300_preg_Log	Significant	Built environment
## hs_accesslines300_s_dic0	Significant	Built environment
## h_Benzene_Log	Significant	Indoor air
## h_NO2_Log	Significant	Indoor air
## h_PM_Log	Significant	Indoor air
## hs_KIDMED_None	Significant	Lifestyle
## hs_sd_wk_None	Significant	Lifestyle
## hs_dif_hours_total_None	Significant	Lifestyle
## hs_cs_m_Log2	Significant	Metals
## hs_cu_c_Log2	Significant	Metals
## hs_cu_m_Log2	Significant	Metals
## hs_hg_c_Log2	Significant	Metals
## hs_mn_m_Log2	Significant	Metals
## hs_mo_m_Log2	Significant	Metals
## hs_pb_c_Log2	Significant	Metals
## hs_hum_dy_hs_h_None	Significant	Meteorological
## hs_dde_cadj_Log2	Significant	Organochlorines
## hs_ddt_cadj_Log2	Significant	Organochlorines
## hs_pcb138_cadj_Log2	Significant	Organochlorines
## hs_pcb138_madj_Log2	Significant	Organochlorines
## hs_pcb153_cadj_Log2	Significant	Organochlorines
## hs_pcb153_madj_Log2	Significant	Organochlorines
## hs_pcb170_cadj_Log2	Significant	Organochlorines
## hs_pcb180_cadj_Log2	Significant	Organochlorines
## hs_pcb180_madj_Log2	Significant	Organochlorines
## hs_sumPCBs5_cadj_Log2	Significant	Organochlorines
## hs_dmtp_cadj_Log2	Significant	Organophosphate pesticides
## hs_pfna_m_Log2	Significant	Per- and polyfluoroalkyl substances (PFAS)
## hs_pfoa_c_Log2	Significant	Per- and polyfluoroalkyl substances (PFAS)
## hs_pfoa_m_Log2	Significant	Per- and polyfluoroalkyl substances (PFAS)
## hs_pfos_c_Log2	Significant	Per- and polyfluoroalkyl substances (PFAS)
## hs_pfos_m_Log2	Significant	Per- and polyfluoroalkyl substances (PFAS)
## hs_mepa_madj_Log2	Significant	Phenols
## hs_mecpp_cadj_Log2	Significant	Phthalates
## hs_mehhp_cadj_Log2	Significant	Phthalates
## hs_mehhp_madj_Log2	Significant	Phthalates
## hs_mehp_cadj_Log2	Significant	Phthalates
## hs_meohp_cadj_Log2	Significant	Phthalates
## hs_mep_cadj_Log2	Significant	Phthalates
## hs_mnbp_cadj_Log2	Significant	Phthalates
## hs_sumDEHP_cadj_Log2	Significant	Phthalates

Les facteurs de confusions étant finalement déterminés, et pour mieux identifier les relations entre les différentes expositions et les phénotypes, une matrice contenant les p-values de la régression linéaire généralisée (glm) entre chaque exposome et chaque phénotype a été calculé. Ceci a été effectué sur les deux sous bases de données (numérique pendant la grossesse et numérique postnatale) . Grâce à ceci, on a pu décider le phénotype d'intérêt sur lequel on allait travailler (Développement neuronal chez l'enfant h\_Gen\_Tot) et les

familles d'expositions considérées selon les valeurs de p values les plus petites présentes dans la matrice.

Les familles d'exposition qui vont être considérés sont : Organochlorines, Air pollution, Phthalates, Lifestyle and PFAS.

## **Analyse des Exposomes de type Facteur**

### **Analyse de l'effet des exposomes sur les phénotypes**

#### **Prédiction: Methode LASSO**