



HAL
open science

PID1 is associated to a respiratory endotype related to occupational exposures to irritants

Andrianjafimasy Miora, Olivier Laurent, Margaritte Jeannin Patricia, Mohamdi Hamida, Demenais Florence, Le Moual Nicole, Matran Regis, Zerimech Farid, Dumas Orianne, Marie Hélène Dizier, et al.

► To cite this version:

Andrianjafimasy Miora, Olivier Laurent, Margaritte Jeannin Patricia, Mohamdi Hamida, Demenais Florence, et al.. PID1 is associated to a respiratory endotype related to occupational exposures to irritants. *Free Radical Biology and Medicine*, 2021, 172, pp.503-507. 10.1016/j.freeradbiomed.2021.05.038 . hal-03401805

HAL Id: hal-03401805

<https://hal.science/hal-03401805>

Submitted on 2 Aug 2023

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 4.0 International License

2

3 **Title: *PID1* is associated to a respiratory endotype related to occupational exposures to**
4 **irritants**

5 Andrianjafimasy Miora *et al.*

6

7

8 **Acknowledgements**

9 **EGEA cooperative group members are as follows:**

10 **Coordination:** V Siroux (epidemiology, PI since 2013); F Demenais (genetics); I Pin (clinical
11 aspects); R Nadif (biology); F Kauffmann (PI 1992-2012). **Respiratory**

12 **epidemiology:** *Inserm ex-U 700, Paris:* M Korobaeff (Egea1), F Neukirch (Egea1); *Inserm*

13 *ex-U 707, Paris:* I Annesi-Maesano (Egea1-2); *Inserm U 1018, Villejuif:* O Dumas, F

14 Kauffmann, N Le Moual, R Nadif, MP Oryszczyn (Egea1-2), R Varraso; *Inserm U 1209*

15 *Grenoble:* J Lepeule, V Siroux. **Genetics:** *Inserm ex-U 393, Paris:* J Feingold; *Inserm UMR*

16 *1124, Paris:* E Bouzigon, MH Dizier, F Demenais; *CNG, Evry:* I Gut (now CNAG,

17 Barcelona, Spain), M Lathrop (now Univ McGill, Montreal, Canada). **Clinical**

18 **centers:** *Grenoble:* I Pin, C Pison; *Lyon:* D Ecochard (Egea1), F Gormand, Y Pacheco;

19 *Marseille:* D Charpin (Egea1), D Vervloet (Egea1-2); *Montpellier:* J Bousquet; *Paris*

20 *Cochin:* A Lockhart (Egea1), R Matran (now in Lille); *Paris Necker:* E Paty (Egea1-2), P

21 Scheinmann (Egea1-2); *Paris-Trousseau:* A Grimfeld (Egea1-2), J Just. **Data management**

22 **and quality:** *Inserm ex-U155, Paris:* J Hochez (Egea1); *Inserm U 1018, Villejuif:* N Le

23 Moual, L Orsi; *Inserm ex-U780, Villejuif:* C Ravault (Egea1-2); *Inserm ex-U794, Evry:* N

24 Chateigner (Egea1-2); *Inserm UMR 1124, Paris:* H Mohamdi; *Inserm U1209, Grenoble:* A

25 Boudier, J Quentin (Egea1-2).

27 aspects of the examinations involved: interviewers, technicians for lung function testing and
28 skin prick tests, blood sampling, IgE determinations, coders, those involved in quality control,
29 data and sample management and all those who supervised the study in all centres. The
30 authors are grateful to the three CIC-Inserm of Necker, Grenoble and Marseille who
31 supported the study and in which participants were examined. They are also grateful to the
32 biobanks in Lille (CIC Inserm), and at Annemasse (Etablissement Français du sang) where
33 biological samples are stored.

34 They are indebted to all the individuals who participated, without whom the study would not
35 have been possible.

36

38 The EGEA study and inclusion criteria have been described in detail previously (1). The
39 EGEA is a French case–control and family-based study started in 1990s with two follow-ups
40 over 20 years. The first EGEA survey (EGEA1) included cases with asthma recruited in five
41 chest clinics, their first-degree relatives and population-based controls (n=2047). A first
42 follow-up of the participants was completed in 2003-2007 (EGEA2), including 1602 subjects,
43 almost exclusively adults (98%) with complete examination including lung function tests,
44 skin prick test to 12 aeroallergens, and blood samples. At each survey, all participants
45 answered a questionnaire based on international standardized tools to evaluate asthma, and to
46 determine respiratory and allergic symptoms, treatments, occupational history and
47 environmental exposures. Occupational exposure to irritants during the current or last held job
48 was assessed by the updated Occupational Asthma-specific Job-Exposure Matrix (three levels
49 of exposure: no/medium/high) (2). Ethical approval was obtained from the relevant
50 institutional review board committees (Cochin Port-Royal Hospital and Necker-Enfants
51 Malades Hospital, Paris). Participants signed a written informed consent.

52

53 **Respiratory endotypes**

54 Five respiratory endotypes, previously described in detail by Nadif *et al.* (3), have been
55 identified using a cluster analysis based on demographic, clinical, functional and biological
56 characteristics: two endotypes among never-asthmatics (NA): NA1 (n=166) predominantly
57 asymptomatic; NA2 (n=489) predominantly with respiratory symptoms, and three endotypes
58 among current asthmatics (CA): CA1 (n=53) predominantly characterized by adult-onset
59 asthma with poor lung function, use of asthma treatments, cough and phlegm, asthma
60 exacerbations, high neutrophil count and high FLOPs level; CA2 (n=219) predominantly with

62 asthma, allergic sensitization and high IgE level.

63

64

65 **Genotyping**

66 The EGEA participants were genotyped using Illumina 610 Quad array at the Centre National
67 de Génotypage (CNG, Evry, France) as part of the European Gabriel consortium asthma
68 GWAS (4). Stringent quality control criteria were used to select both individuals and
69 genotyped SNPs for analysis as previously described (4). Five hundreds and nine participants
70 out of the 542 participants from NA1 and CA1 endotypes (94%) were genotyped. Imputations
71 of non-genotyped SNPs were performed using HapMap2 CEU haplotypes as reference panel.
72 Imputed SNPs were kept for analysis if their imputation quality score (rsq) was greater than or
73 equal to 0.9.

74

75

76 **Selection of genes**

77 The gene selection was based on the first step of the pathway-based approach described
78 previously (5). We selected from the Gene Ontology (GO) database 602 genes involved in the
79 “response to oxidative stress” pathways (GO: 0006979) that we have enriched by biological
80 knowledge and literature reviews, resulting in a list of 625 genes. After excluding duplicates,
81 597 genes and 6265 corresponding SNPs located within these genes were selected.

82 To keep the number of tests reasonable and thus increase the power, we examined only the
83 SNPs located within the limits of the genes.

84

85

87

88 *Analysis with genotyped data*

89 To study the associations between SNPs located in genes involved in the response to
90 oxidative stress and the endotype CA1 (*vs.* NA1), we performed logistic regression using
91 STATA v14.1 with the cluster (within family) and robust variance option of the logit
92 function, to take into account familial dependency between individuals. All models included
93 the principal components (PCs) accounting for within-France diversity. After selection of
94 SNPs with size of affected and unaffected subjects per genotype (effective ≥ 10), 4715 SNPs
95 belonging to 422 genes were analyzed.

96 To account for multiple testing, the Bonferroni corrected significance p-value threshold
97 applied to the Meff (effective number of independent tests after discarding dependence due to
98 linkage disequilibrium (LD) between the SNPs (6)) was calculated for the 4715 SNPs. The
99 Meff was estimated to 2185 and thus the significance p threshold equal to $0.05 / 2185 = 2.3E-$
100 05.

101 To check the robustness of the results regarding the possible bias due to small sample size (for
102 some SNPs, the size per genotype expected under the null hypothesis of no association was
103 too small (<5) in CA1), we conducted different tests specific to small sample size including
104 exact logistic regression and Firth logistic regression, an approach equivalent to penalizing the
105 likelihood by the Jeffreys invariant prior (7,8). Because with these tests, familial correlation
106 was not accounted for, we compared these results with those obtained by the previous
107 regression logistic, but without cluster and robust variance option.

108 To check the specificity of the associated SNPs with the endotype CA1, we also investigated
109 the associations between these SNPs and the other endotypes CA2 and CA3 (*vs.* NA1).

110

112 For each of the 422 genes, associations with the endotype CA1 (vs. NA1) were also
113 investigated at the gene level by using the versatile gene-based test (Versatile Gene-based
114 Association Study (9)). The gene-based statistic was defined as the best SNP test statistic (or
115 min p-value) using the results of logistic regression. The correlation between these statistics
116 due to the LD between SNPs is taken into account by computing an empirical p-value through
117 Monte-Carlo simulations using the linkage disequilibrium pattern of HapMap CEU (Utah
118 residents with ancestry from northern and western Europe) reference sample. The empirical p-
119 values were then adjusted for multiple testing using the Bonferroni correction (significance p
120 threshold = $0.05/422 = 1.2E-04$).

121

122 *Analyses with imputed data from HapMap phase 2 population CEU*

123 The regions around the SNPs showing significant association with the endotype CA1 (vs.
124 NA1) were more finely studied using the imputed SNPs (imputation from HapMap 2 CEU)
125 spanning 5kb on each side of each associated SNP.

126

127 *Stratification on occupational exposure to irritant agents*

128 For the 10 top SNPs associated with CA1, we conducted association analyses separately in
129 two groups: in participants who had high occupational exposure to irritants and in those who
130 had no exposure. ~~Performing formal interaction test was not possible because of too small
131 sample size after stratification in two groups.~~

132

133 *eQTLs, functional annotations and chemical-gene/protein interactions*

134 We investigated whether the SNPs (or their proxies, $r^2 \geq 0.8$) associated with the endotype CA1
135 were cis-expression quantitative trait loci (cis-eQTLs). We used the eQTL browser

137 databases like the Genotype-Tissue Expression project (GTEx) and includes e-QTL data from
138 many tissues (10,11). No eQTL was found.

139 Furthermore, functional annotations of these SNPs (or proxies) were done using the HaploReg
140 v4.1 tool (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>). HaploReg
141 annotates SNPs in terms of colocalisation with regulatory elements, such as promoter and
142 enhancer marks, DNase I hypersensitivity sites and transcription factor (TF) and protein-
143 binding sites, based on Roadmap Epigenomics data and Encyclopedia of DNA Elements data
144 (Table E2).

145 Finally, curated (chemical-gene interactions | chemical-disease | gene-disease) data were
146 retrieved from the Comparative Toxicogenomics database (CTD) (12), MDI Biological
147 Laboratory, Salisbury Cove, Maine, and NC State University, Raleigh, North Carolina. Word
148 Wide Web (URL: <http://ctdbase.org/>) [May, 2019]). CTD is a database that aims to advance
149 understanding about how environmental exposures affect human health (Table E3).

150

151

153 1. Kauffmann F, Dizier MH, Pin I, Paty E, Gormand F, Vervloet D, Bousquet J, Neukirch
154 F, Annesi I, Oryszczyn MP, Lathrop M, Demenais F, Lockhart A FJ. Epidemiologic
155 Study of the Genetics and Environment of Asthma, Bronchial Hyperresponsiveness,
156 and Atopy. Phenotype Issues. *Am J Respir Crit Care Med.* 1997;156:s123–9.

157 2. Le Moual N, Zock J-P, Dumas O, Lytras T, Andersson E, Lillienberg L, et al. Update
158 of an occupational asthma-specific job exposure matrix to assess exposure to 30
159 specific agents. *Occup Env Med.* 2018;75(7):507–14.

160 3. Nadif R, Febrissy M, Andrianjafimasy MV, Le Moual N, Gormand F, Just J, et al.
161 Endotypes identified by cluster analysis in asthmatics and non-asthmatics and their
162 clinical characteristics at follow-up: the case-control EGEA study. *BMJ Open Resp
163 Res.* 2020;7:632.

164 4. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A Large-
165 Scale, Consortium-Based Genomewide Association Study of Asthma. *N Engl J Med.*
166 2010;363:1211–21.

167 5. Rava M, Ahmed I, Demenais F, Sanchez M, Tubert-Bitter P, Nadif R. Selection of
168 genes for gene-environment interaction studies: a candidate pathway-based strategy
169 using asthma as an example. *Environ Heal.* 2013;12:56.

170 6. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a
171 correlation matrix. *Heredity (Edinb).* 2005;95:221–7.

172 7. Firth D. Bias Reduction of Maximum Likelihood. *Biometrika.* 1993;80(1):27–38.

173 8. Wang X, Zhang Q, Liu D. Firth logistic regression for rare variant association tests.
174 *Front Genet.* 2014;5(187).

175 9. Liu JZ, Mcrae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A Versatile
176 Gene-Based Test for Genome-wide Association Studies. *Am J Hum Genet.*

- 178 10. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al.
179 PhenoScanner V2: an expanded tool for searching human genotype-phenotype
180 associations. *Bioinformatics*. 2019;35:4851–3.
- 181 11. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al.
182 PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics*.
183 2016;32:3207–9.
- 184 12. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wieggers J, Wieggers TC, et al.
185 Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic Acids Res*.
186 2021;49:D1138–43.

NA1)

SNP	A1/A2	EAF	OR (95% CI)	z	p
rs4973171	A/G	0.51	0.72 (0.48-1.10)	-1.51	0.13
rs10171191	A/G	0.49	1.38 (0.91-2.10)	1.51	0.13
rs10171051	A/C	0.75	0.73 (0.43-1.25)	-1.16	0.25
rs10183571	C/T	0.49	1.38 (0.91-2.10)	1.51	0.13
rs2396617	A/G	0.15	0.35 (0.22-0.56)	-4.40	1.10E-05
rs2396618	A/G	0.15	0.35 (0.22-0.56)	-4.40	1.10E-05
rs10804349	G/T	0.11	1.00 (0.57-1.77)	0.01	1.00
rs1419958	G/T	0.85	0.35 (0.22-0.56)	-4.40	1.10E-05
rs1419959	G/T	0.40	0.62 (0.41-0.94)	-2.26	0.02
rs12477036	A/G	0.11	1.54 (0.75-3.18)	1.18	0.24
rs12694800	C/T	0.60	1.61 (1.07-2.44)	2.27	0.02
rs1419960	A/G	0.49	1.38 (0.91-2.10)	1.51	0.13
rs12470314	C/T	0.11	1.54 (0.75-3.18)	1.17	0.24
rs13035964	A/G	0.94	0.21 (0.03-1.48)	-1.57	0.12
rs17676334	C/G	0.11	1.53 (0.74-3.17)	1.14	0.26
rs1362256	C/T	0.44	0.63 (0.42-0.95)	-2.18	0.03

Significant result in bold.

SNP: Single nucleotide polymorphism; A1/A2: baseline/effect allele; EAF: Effect allele frequency estimated from the reference panel 1000G (European population); OR: odds ratio; CI: confidence interval.

Table E2: Regulatory elements of the three tops SNPs (or proxies) located in *PID1* gene associated to the CA1 endotype.

Chr	Gene	pos (hg19)	r ²	D'	SNP	Enhancer/Promotor histone marks	Protein bounds	Regulatory motifs altered
2		230030822	0.96	0.99	rs72617162			Osf2_2
2		230031022	0.96	0.99	rs147700606			EWSR1-FLI1. GR_disc6
2	PID1	230031593	0.96	0.99	rs2396617			Evi-1_2
2		230032055	0.96	0.99	rs2396618			AIRE_1
2		230033560	1	1	rs1419958	Fetal lung (H3K4me3_Pro)		COMP1. SIX5. SREBP
2	PID1	229116897	1	1	rs10490031			<u>HDAC2</u> . TCF4
2		229150560	0.82	0.93	rs7569303		FOXA1	STAT. p300
2		229156450	0.88	1	rs2396615			PRDM1
2		229156686	1	1	rs2396616			NRSF. p300
2		229157051	0.88	1	rs35419420			Pou1f1
2		229158359	0.88	1	rs13030636			PU.1
2		229159213	0.88	1	rs12463675			ZBRK1
2	PID1	229159594	0.88	1	rs12463823			Nanog, Nkx2, Pou2f2, Pou3f3, Pou5f1, TATA, Zfp740
2		229159794	0.9	1	rs17614196			<u>HDAC2</u> , Irf, p300, p53
2		229161497	0.88	1	rs17614232			Foxp1, Mrg_2, <u>NF-kappaB</u> , ZEB1
2		229161548	0.88	1	rs17676274			ATF3, E2F, Foxa, Jundm2, Maf, STAT
2		229162064	0.88	1	rs1362255			Pou3f3
2		229162503	0.88	0.99	rs6745401			Smad3
2		229162869	0.99	1	rs1419954			AFP1, AIRE_2, Evi-1, Foxa, Foxf1, Foxj1, Foxj2, Foxl1, Foxo, Foxq1, <u>HDAC2</u> , SIX5, Sox, Zfp105, p300

In bold: the three tops SNPs located in *PID1* gene associated to the CA1 endotype.

Chr : chromosome ; SNP : single nucleotide polymorphism

Table E3: Chemical retrieved from the Comparative Toxicogenomics Database (CTD, URL: <http://ctdbase.org/>, Davis et al. 2021 (12)) and interacting with *PID1* gene.

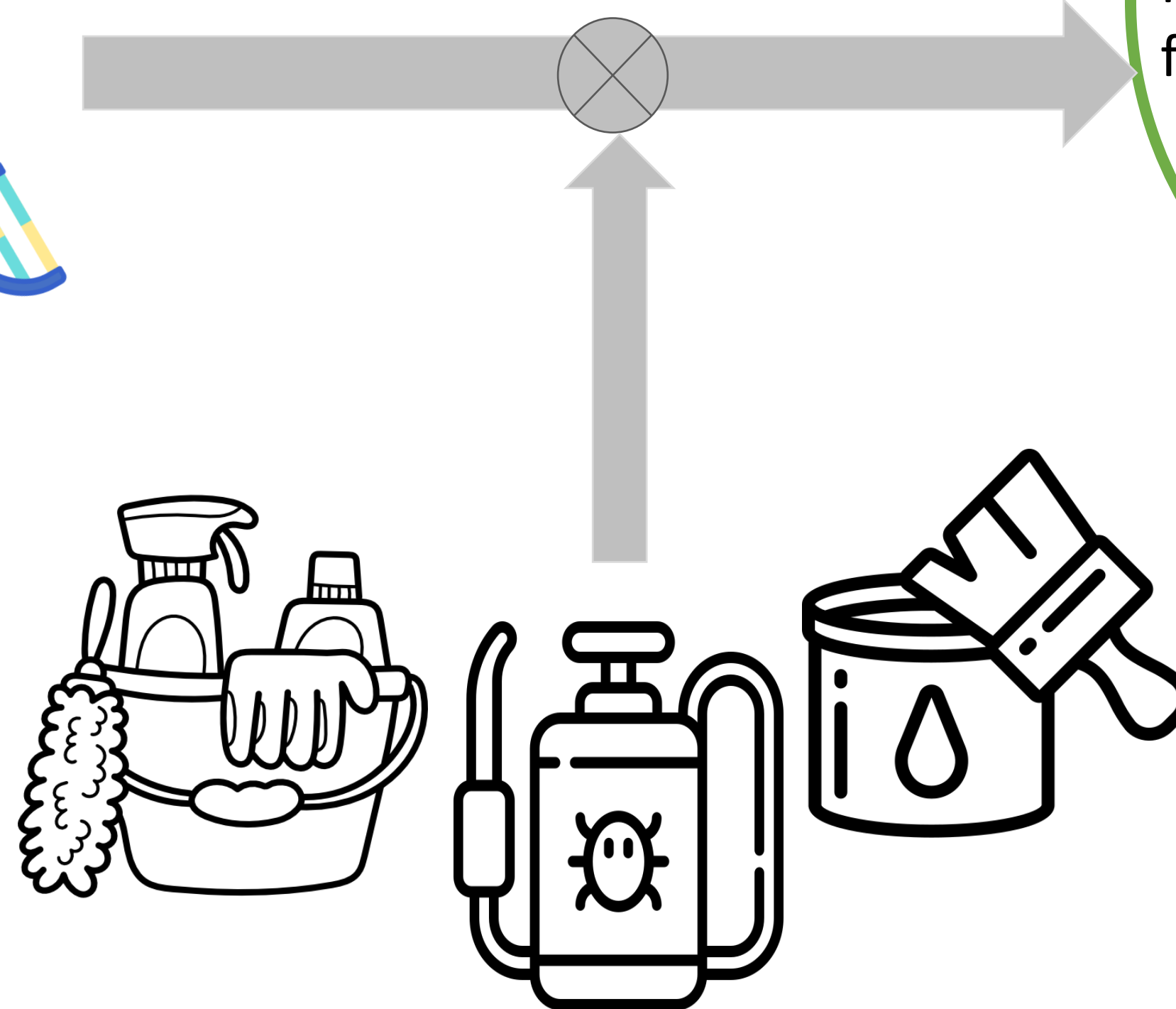
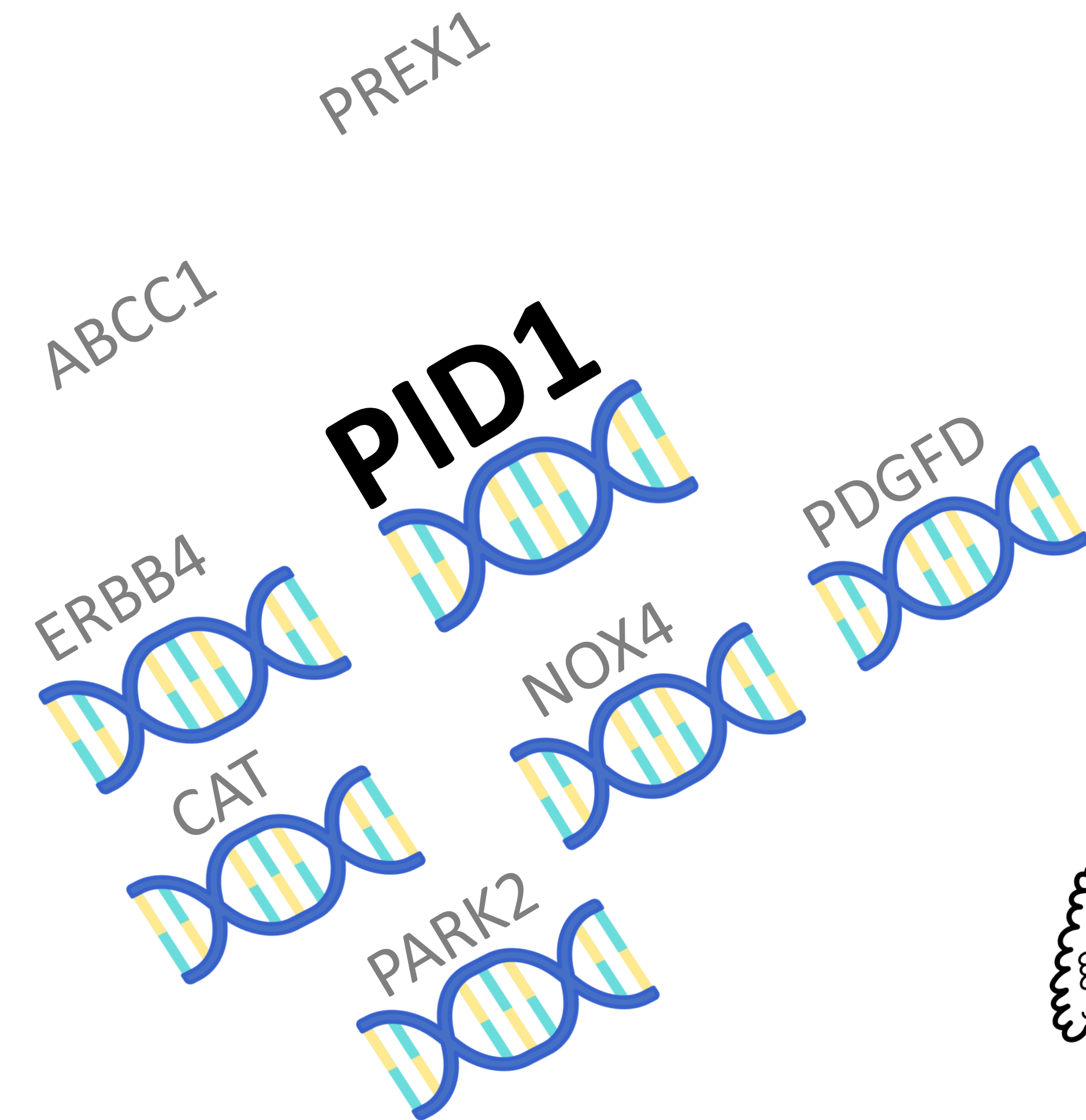
Chemical name	Gene	Interactions	Organism	References
Aldehydes	<i>PID1</i>	Aldehydes results in increased expression of PID1 mRNA	Homo sapiens	Song MK, et al. Transcriptome Profile Analysis of Saturated Aliphatic Aldehydes Reveals Carbon Number-Specific Molecules Involved in Pulmonary Toxicity. <i>Chem Res Toxicol.</i> 2014 Jun 25 ;27:1362-70.
Butyraldehyde	<i>PID1</i>	Butyraldehyde results in increased expression of PID1 mRNA	Homo sapiens	Song MK, et al. Integrated analysis of microRNA and mRNA expression profiles highlights aldehyde-induced inflammatory responses in cells relevant for lung toxicity. <i>Toxicology.</i> 2015 Aug 6 ;334:111-21.
Propionaldehyde	<i>PID1</i>	Propionaldehyde results in increased expression of PID1 mRNA	Homo sapiens	Song MK, et al. Integrated analysis of microRNA and mRNA expression profiles highlights aldehyde-induced inflammatory responses in cells relevant for lung toxicity. <i>Toxicology.</i> 2015 Aug 6 ;334:111-21.
Carbon Tetrachloride	<i>PID1</i>	Carbon Tetrachloride results in decreased expression of PID1 mRNA	Rattus norvegicus	Zhang H, et al. Transcriptomic analyses reveal the molecular mechanisms of schisandrin B alleviates CCl ₄ -induced liver fibrosis in rats by RNA-sequencing. <i>ChemBiol Interact.</i> 2019 May 28 ;309:108675.
Methyl cellosolve	<i>PID1</i>	Methyl cellosolve results in increased expression of PID1 mRNA	Rattus norvegicus	Yuan X, et al. Finding maximal transcriptome differences between reprotoxic and non-reprotoxic phthalate responses in rat testis. <i>J Appl Toxicol.</i> 2011 Jun ;31(5):421-30.
Phthalic Acids	<i>PID1</i>	Phthalic Acids results in increased expression of PID1 mRNA	Rattus norvegicus	Yuan X, et al. Finding maximal transcriptome differences between reprotoxic and non-reprotoxic phthalate responses in rat testis. <i>J Appl Toxicol.</i> 2011 Jun ;31(5):421-30.
Pentanal	<i>PID1</i>	Pentanal results in increased expression of PID1 mRNA	Homo sapiens	Song MK, et al. Integrated analysis of microRNA and mRNA expression profiles highlights aldehyde-induced inflammatory responses in cells relevant for lung toxicity. <i>Toxicology.</i> 2015 Aug 6 ;334:111-21.
Ethanol	<i>PID1</i>	Ethanol affects the expression of and affects the splicing of PID1 mRNA	Mus musculus	O'Brien MA, et al. Ethanol-Induced Behavioral Sensitization Alters the Synaptic Transcriptome and Exon Utilization in DBA/2J Mice. <i>Front Genet.</i> 2018;9:402.
Vehicle Emissions	<i>PID1</i>	Vehicle Emissions results in increased methylation of PID1 gene	Mus musculus	Tachibana K, et al. Prenatal diesel exhaust exposure disrupts the DNA methylation profile in the brain of mouse offspring. <i>J Toxicol Sci.</i> 2015 Feb ;40(1):1-11.

In this table are indicated the irritant agents or environmental exposures known to contain compounds with irritant properties (vehicle emissions).

- *PID1* gene involved in oxidative stress associated with a distinct asthma endotype.
- ~~Potential interaction between *PID1* and occupational exposure~~ Association found in unexposed participants but not in those exposed to irritants.
- ~~Interest of asthma endotypes in genetic study to discover new susceptibility genes.~~

Genes involved in "Response to oxidative stress"

Asthma endotypes (CA)



Occupational exposure to irritant agents

