Supplementary Data

Embryonic exposure to an aqueous coal dust extract results in gene expression alterations associated with development and function of connective tissue and the hematological system, immunological and inflammatory disease, and cancer in zebrafish

Karina Caballero-Gallardo[a], Sara E. Wirbisky-Hershberger[b], Jesus Olivero-Verbel[a], Jesus de la Rosa[c], Jennifer L. Freeman[b*]

[a]Environmental and Computational Chemistry Group, School of Pharmaceutical Sciences, Zaragocilla Campus, University of Cartagena, Cartagena 130015, Colombia.
[b]School of Health Sciences, Purdue University, West Lafayette, IN 47907, USA.
[c]Associate Unit CSIC – University of Huelva ''Atmospheric Pollution'', Center for Research in Sustainable Chemistry (CIQSO), University of Huelva, E21071 Huelva, Spain

Table of Contents:

Supplementary Table 1. Primers used for qPCR microarray confirmation.

Gene name	Human Gene Symbol	Entrez Gene ID	Primer Sequence
Clarin 3	*CLRN3*	XM_001340649.3	F:GTCGCTGATTTTCTACGTGTCG R:ACGGGATCAGGAGAAAGAAACC
Cadherin-related family member 2	*CDHR*	XM_009291246.1	F:CTACTGTGGGCTTCGTACTGTT R:AGTGGCCGCAACTATTGAAAAC
Chondroitin sulfate synthase 1	*CHSY1*	NM_212678.1	F:GGAGATGATTAACGCCAATGCC R:TGTCATAGTCTTCCCCTTGTGC
Rho guanine nucleotide exchange factor (GEF) 1a	*ARHGEF*	NM_001081631	F:ATCATTCAGCAGACAGTGAGCA R:GGGAACGGTAACATTCTGACCT
Tripartite motif containing 59	*TRIM59*	XM_009291407.1	F:ACGTACATGCACTTGAAACTGC R:GTGTGTCTTGTTTGTGCCCTAC
Transient receptor potential cation channel, subfamily M, member 3	*TRPM3*	XM_009301562.1	F:CAAGCCTGACCTTTCCAACATG R:TCTACTGCCATGTCAGTGCATT
Eukaryotic translation elongation factor 1 alpha 1, like 1	*ELF1α*	NM_131263.1	F:GCCAAGACCAAGTGAATTTCCC R:GTCCTTAAGTAGAGTGCCCAGG

Supplementary Table 2. Morphological analysis at 72 hpf.

Concentrations (ppm)	Body length (µm)[a]	Head length (µm)	Head width (µm)
0	3,266 ± 52	593 ± 4	546 ± 3
0.1	3,290 ± 13	591 ± 3	549 ± 4
1	3,275 ± 17	587 ± 4	553 ± 4
10	3,324 ± 18	593 ± 3	548 ± 3
100	3,303 ± 14	592 ± 3	547 ± 3

[a]Data are representative of three independent replicates of 20 subsamples each ± standard error (p>0.05).

Supplementary Table 3. Gene enrichment table of physiological system development and function for 72 hpf zebrafish larvae at 1 ppm aqueous coal dust treatment.

Physiological System Development and Function – 1 ppm	p-value[a]	Number of Genes[b]
SKELETAL AND MUSCULAR SYSTEM DEVELOPMENT AND FUNCTION	4.96E-03 – 6.51E-06	11
	6.51E-06	6
Formation of osteoclasts	1.27E-03	6
Proliferation of smooth muscle cells	3.91E-03	5
Size of bone	3.80E-05	4
Quantity of trabecula		
ORGANISMAL DEVELOPMENT	5.30E-03 – 6.51E-06	24
Abnormal development of body cavity	1.13E-03	14
Abnormal morphology of abdomen	2.18E-04	13
Morphology of head	5.62E-04	12
ORGAN DEVELOPMENT	4.96E-03 – 6.51E-06	16
Formation of osteoclasts	6.51E-06	6
Formation of rhombencephalon	3.21E-03	4
Function of skin	8.95E-04	3
Maturation of thymocytes	3.41E-03	2
CONNECTIVE TISSUE DEVELOPMENT AND FUNCTION	4.96E-03 – 6.51E-06	13
Development of connective tissue cells	3.80E-05	7
Formation of osteoclasts	6.51E-06	6
Differentiation of osteoclasts	1.07E-03	5
Quantity of bone	3.49	4
EMBRYONIC DEVELOPMENT	4.98E-03 – 6.51E-06	20
Development of body trunk	4.21E-03	12
Development of abdomen	4.98E-03	8
Formation of osteoclasts	6.51E-06	6
Hematopoiesis of liver	3.02E-05	3

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.
[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 4. Gene enrichment table of physiological system development and function for 72 hpf zebrafish larvae at 10 ppm aqueous coal dust treatment.

Physiological System Development and Function – 10 ppm	p-value[a]	Number of Genes[b]
BEHAVIOR	4.85E-03 – 7.97E-06	14
Behavior	7.97E-06	14
Emotional behavior	1.23E-04	6
Maternal behavior	8.26E-05	3
Learning	3.99E-04	7
HEMATOLOGICAL SYSTEM DEVELOPMENT AND FUNCTION	5.47E-03 – 9.64E-06	13
Quantity of leukocytes	3.27E-04	11
Quantity of T lymphocytes	1.25E-05	10
Differentiation of blood cells	1.67E-03	9
T cell development	1.63E-03	7
HEMATOPOESIS	5.47E-03 – 9.64E-06	11
Development of leukocytes	1.81E-04	9
Maturation of leukocytes	9.36E-04	5
Quantity of thymocytes	8.56E-05	6
T cell development	1.63E-03	7
TISSUE MORPHOLOGY	5.47E-03 – 1.24E-05	15
Quantity of leukocytes	3.27E-04	11
Quantity of B lymphocytes	7.72E-04	6
Permeability of blood vessel	1.23E-03	3
Quantity of neuroglia	3.23E-03	3
CONNECTIVE TISSUE DEVELOPMENT AND FUNCTION	5.47E-03 – 2.42E-05	11
Differentiation of osteoclast	2.49E-05	6
Formation of osteoclasts	2.42E-05	5
Proliferation of fibroblasts	5.41E-03	5
Angiogenesis of stromal tissue	2.74E-03	1

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.

[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 5. Gene enrichment table of physiological system development and function for 72 hpf zebrafish larvae at 100 ppm aqueous coal dust treatment.

Physiological System Development and Function – 100 ppm	p-value[a]	Number of Genes[b]
ORGANISMAL SURVIVAL	8.56E-04 – 3.35E-19	331
Organismal death	3.35E-19	331
Perinatal death	9.81E-12	100
Neonatal death	1.56E-08	71
Death of embryo	8.56E-04	22
BEHAVIOR	9.02E-04 – 7.88E-13	164
Behavior	7.88E-13	151
Cognition	3.37E-05	60
Learning	1.22E-04	54
Locomotion	7.22E-04	43
NERVOUS SYSTEM DEVELOPMENT AND FUNCTION	1.14E-03 – 3.04E-12	313
Development of neurons	3.04E-12	120
Morphology of nervous system	1.02E-08	114
Neuritogenesis	2.49E-08	86
Neurotransmission	1.35E-08	67
Long-term potentiation	2.61E-04	34
TISSUE DEVELOPMENT	1.13E-03 – 3.04E-12	332
Development of neurons	3.04E-12	120
Proliferation of neuronal cells	2.22E-07	84
Axonogenesis	8.41E-07	38
Growth of neurites	5.26E-06	66
EMBRYONIC DEVELOPMENT	1.13E-03 – 5.68E-11	329
Development of body axis	5.68E-11	144
Development of head	1.11E-09	130
Formation of brain	5.00E-06	70
Growth of embryo	6.64E-05	60
Morphology of bone	8.95E-04	57

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.

[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 6. Gene enrichment table of diseases and disorders in 72 hpf zebrafish larvae in the 1 ppm aqueous coal treatments.

Diseases and Disorders – 1 ppm	p-value[a]	Number of Genes[b]
NEUROLOGICAL DISEASE	5.61E-03 – 9.24E-06	26
Movement disorders	9.24E-06	17
Seizures	2.04E-03	7
Ataxia	5.01E-04	6
Damage of nervous tissue	5.16E-03	3
HYPERSENSITIVITY RESPONSE	4.96E-03 – 1.19E-05	7
Degranulation of mast cells	1.19E-05	6
Morphology of mast precursor cells	3.63E-03	1
Proliferation of mast cells	4.96E-03	2
INFLAMMATORY RESPONSE	5.17E-03 – 1.19E-05	16
Inflammatory response	3.02E-03	10
Immune response of cells	4.01E-03	8
Activation of T lymphocytes	9.33E-04	7
Cell movement of neutrophils	1.71E-03	6
CONNECTIVE TISSUE DISORDERS	5.30E-03 – 1.71E-05	16
Rheumatic disease	1.43E-03	13
Arthritis	1.69E-03	12
Damage of cartilage tissue	3.55E-03	3
IMMUNOLOGICAL DISEASE	5.07E-03 – 1.71E-05	14
Rheumatoid arthritis	5.07E-03	9
Splenomegaly	3.37E-04	4
Onset of autoimmune disease	1.33E-03	2

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.
[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 7. Gene enrichment table of diseases and disorders in 72 hpf zebrafish larvae in the 10 ppm aqueous coal treatments.

Diseases and Disorders – 10 ppm	p-value[a]	Number of Genes[b]
CONNECTIVE TISSUE DISORDERS	2.74E-03 – 1.46E-06	11
Rheumatic disease	1.22E-03	11
Arthritis	1.86E-03	10
Juvenile rheumatoid arthritis	1.46E-03	6
IMMUNOLOGICAL DISEASE	5.47E-03 – 1.46E-06	11
Splenomegaly	4.96E-06	5
Thymic lymphoma	6.24E-04	3
Hyperplasia of lymphoid organ	2.46E-03	3
INFLAMMATORY DISEASE	3.67E-03 – 1.46E-06	12
Rheumatic disease	1.22E-03	11
Arthritis	1.86E-03	10
Juvenile rheumatoid arthritis	1.46E-03	6
Multiple sclerosis	2.52E-03	4
SKELETAL AND MUSCULAR DISORDERS	5.47E-03 – 1.46E-06	19
Rheumatic disease	1.22E-03	11
Neuromuscular disease	9.14E-04	10
Parkinson's disease	1.25E-04	6
Fibromyalgia	5.38E-04	3
CANCER	5.47E-03 – 1.29E-05	24
Hyperplasia	1.29E-05	10
Skin tumor	2.67E-03	9
Skin cancer	5.44E-03	8
Development of tumor	8.03E-04	6

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.

[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 8. Gene enrichment table of diseases and disorders in 72 hpf zebrafish larvae in the 100 ppm aqueous coal treatments.

Diseases and Disorders – 100 ppm	p-value[a]	Number of Genes[b]
CANCER	1.20E-03 – 1.19E-18	1133
Cancer	1.32E-17	1106
Solid tumor	2.20E-18	1078
Epithelial cancer	1.30E-15	924
Abdominal cancer	5.81E-13	906
Melanoma	2.24E-09	647
NEUROLOGICAL DISEASE	1.20E-03 – 1.95E-18	355
Movement disorders	1.95E-18	177
Neuromuscular disease	8.37E-15	142
Seizures	1.94E-09	68
Degeneration of neurons	1.34E-05	37
PSYCHOLOGICAL DISORDERS	8.38E-06 – 1.20E-15	190
Disorder of basal ganglia	1.20E-15	132
Huntington's disease	4.08E-14	104
Schizophrenia	3.19E-07	75
Hyperactive behavior	8.38E-06	25
SKELETAL AND MUSCULAR DISORDERS	1.16E-03 – 8.37E-15	222
Neuromuscular disease	8.37E-15	142
Huntington's disease	4.08E-14	104
Articular rigidity	8.05E-05	11
HEREDITARY DISORDER	1.16E-03 – 4.08E-14	211
Huntington's disease	4.08E-14	104
Autosomal dominant cerebellar ataxia	9.97E-07	15
X-linked hereditary disease	8.80E-06	39
Autosomal dominant disease	4.57E-04	63

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.

[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 9. Gene enrichment table of molecular and cellular function in 72 hpf zebrafish larvae in the 1 ppm aqueous coal treatments.

Molecular and Cellular Function – 1 ppm	p-value[a]	Number of Genes[b]
GENE EXPRESSION	3.63E-03 – 3.70E-06	15
Binding of DNA	3.70E-06	12
Transactivation	7.22E-05	11
Transactivation of RNA	1.97E-04	10
Bending of DNA	2.14E-03	2
CELLULAR DEVELOPMENT	4.96E-03 – 6.51E-06	27
Differentiation of cells	1.52E-03	21
Proliferation of blood cells	1.36E-04	13
Proliferation of immune cells	2.85E-04	12
Proliferation of lymphocytes	5.00E-04	11
CELLULAR GROWTH AND PROLIFERATION	4.96E-03 – 6.51E-06	21
Formation of cells	2.18E-04	14
Proliferation of blood cells	1.36E-04	13
Formation of osteoclasts	6.51E-06	6
Stimulation of cells	1.89E-03	6
CELLULAR COMPROMISE	4.63E-03 – 1.19E-05	10
Degranulation of cells	3.71E-05	7
Degranulation of mast cells	1.19E-05	6
Degranulation of natural killer cells	3.13E-03	2
Dysfunction of macrophage	3.63E-03	2
CELL TO CELL SIGNALING	5.07E-03 – 1.28E-05	19
Activation of blood cells	2.44E-03	10
Activation of T lymphocytes	9.33E-04	7
Stimulation of cells	1.89E-03	6
Long term potentiation	3.47E-03	5

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.
[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 10. Gene enrichment table of molecular and cellular function in 72 hpf zebrafish larvae in the 10 ppm aqueous coal treatments.

Molecular and Cellular Function – 10 ppm	p-value[a]	Number of Genes[b]
GENE EXPRESSION	2.74E-03 – 6.08E-07	18
Transcription	9.51E-04	16
Transcription of RNA	2.32E-03	15
Transactivation	6.08E-07	12
Transactivation of RNA	2.31E-06	11
CELL DEATH AND SURVIVAL	5.47E-03 – 4.54E-06	27
Cell death	1.28E-03	23
Necrosis	1.93E-04	21
Apoptosis	2.96E-04	21
Cell viability	1.35E-04	14
CELL TO CELL SIGNALING	5.47E-03 – 5.93E-06	12
Stimulation of cells	5.39E-06	8
Neurotransmission	2.50E-03	6
Activation of neurons	1.03E-04	4
CELLULAR GROWTH AND PROLIFERATION	5.47E-03 – 5.93E-06	22
Formation of cells	7.69E-06	14
Proliferation of T lymphocytes	3.45E-05	10
Stimulation of cells	5.39E-06	8
Formation of osteoclasts	2.42E-05	5
CELLULAR DEVELOPMENT	5.47E-03 – 9.64E-06	24
Differentiation of cells	2.01E-04	19
Proliferation of T lymphocytes	3.45E-05	10
Differentiation of connective tissue	1.68E-04	10
Proliferation of liver cells	2.65E-03	4

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.
[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.

Supplementary Table 11. Gene enrichment table of molecular and cellular function in 72 hpf zebrafish larvae in the 100 ppm aqueous coal treatments.

Molecular and Cellular Function – 100 ppm	p-value[a]	Number of Genes[b]
CELLULAR DEVELOPMENT	1.17E-03 – 3.04E-12	414
Differentiation of cells	1.66E-05	244
Development of neurons	3.04E-12	120
Neuritogenesis	2.19E-08	86
Axonogenesis	8.41E-07	38
CELLULAR ASSEMBLY AND ORGANIZATION	1.19E-03 – 1.14E-10	271
Organization of cytoplasm	5.94E-09	192
Organization of cytoskeleton	5.48E-10	183
Microtubule dynamics	1.75E-10	162
CELL MORPHOLOGY	1.09E-03 – 1.14E-10	332
Formation of cellular protrusions	1.14E-10	131
Outgrowth of neurites	3.27E-06	60
Plasticity of synapse	3.66E-05	21
Neuritogenesis	2.19E-08	86
CELL DEATH AND SURVIVAL	1.18E-03 – 2.02E-11	434
Apoptosis	1.42E-10	319
Necrosis	1.81E-07	294
Cell viability	4.01E-05	149
Degeneration of neurons	1.34E-05	37
GENE EXPRESSION	4.68E-04 – 2.43E-13	296
Expression of RNA	1.25E-11	271
Transcription	1.12E-11	246
Transcription of RNA	5.54E-12	244
Expression of DNA	3.75E-12	220

[a]Derived from the likelihood of observing the degree of enrichment in a gene set of a given size by chance alone.

[b]Classified as being differentially expressed that relate to the specified function category; a gene may be present in more than one category.
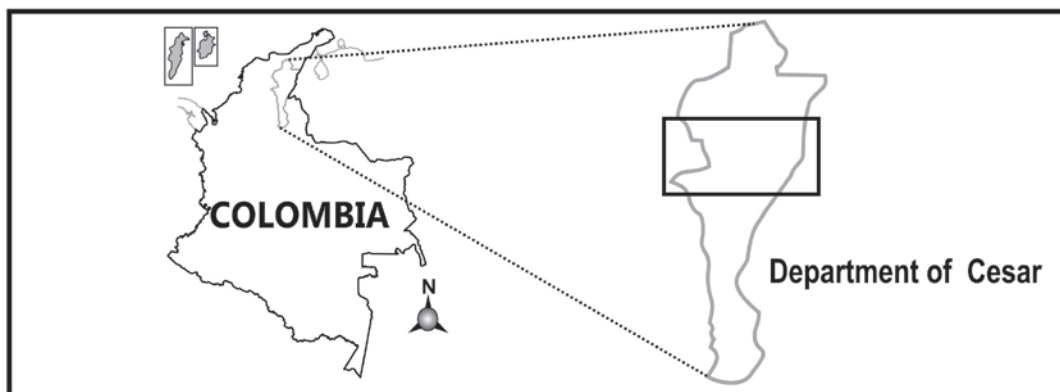
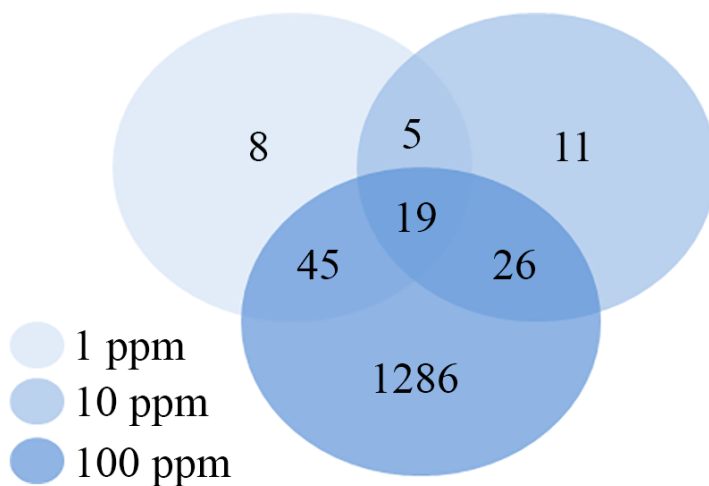**Figure S1.** Map of Colombia showing the location of sample collection.

Figure S2. Venn diagram of number of altered genes by microarray analysis. Following an embryonic exposure (1-72 hpf) to 1, 10, or 100 ppm aqueous coal dust extract. A total of 77, 61, and 1,372 genes were indicated to have altered expression, respectively. 19 of these genes had altered expression in all three treatments.
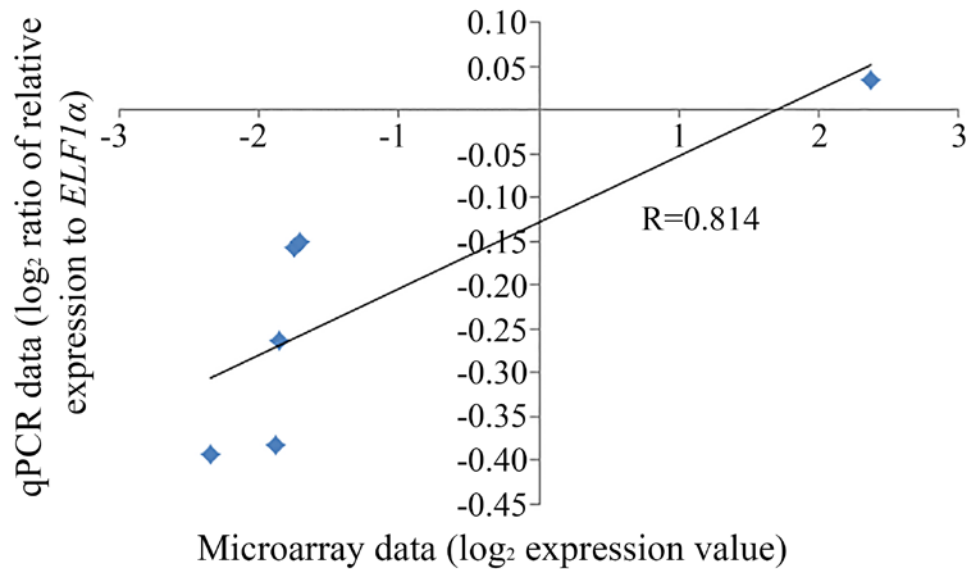
Figure S3. qPCR comparative analysis of microarray data. qPCR analysis was performed to compare gene expression changes detected on microarray analysis in the 100 ppm treatment group. Six target genes were analyzed (*ARHGEF, CDHR, CHSY1, CLRN3, TRIM59*, and *TRPM3*). All targeted genes were significantly altered on the microarray. A linear correlation was completed and statistical significance of the correlation was determined using a Pearson's correlation. The data showed a positive and significant correlation (R=0.814, p=0.0448). The microarray data is represented as $\log_2$ expression values and qPCR values are the ratios of the relative expression against *ELF1α* represented as $\log_2$.