Supplementary Material

Methodological considerations

The results generated by the automatic Pic/Greg gel image analysis software package (Fraunhofer Gesellschaft, Sankt Augustin, http://www.fit.fraunhofer.de/projekte/greg/index-en.xml) were treated in the following way: Protein spots were included if spot volume was at least 1.e10 (in units used by Greg) and background intensity was at least four times less than spot intensity. All spots were checked manually using the Pic program and accepted or rejected depending on their actual appearance. In this way, artifacts were sorted out. For the differential analysis of, e.g., substance treated vs. control proteome, differences in spot intensities were computed. Using a t-test, the probabilities of beta error, generally known as P-values, are computed. If there is good evidence from the gel images that one of the spot intensities is an outlier, e.g., stripes or other distortions occur or the spot is not properly covered by the segmentation, then this one spot's intensity from the respective gel image may be discarded. If a spot is not detected on one image its intensity is set to zero for that image unless it can clearly be identified as an outlier, in which case it is discarded. In general, a spot is selected for further analysis if its p-value is ≤ 0.01, and the abundance ratio of the compared spot intensities is >1.5 or <0.667. At that point the image quality is checked to assure that the spot was consistently detected on all images. In this way, for each comparison of a toxicant vs. its appropriate solute control, a list of differential spots along with their corresponding abundance ratios is obtained. For the subsequent cluster analysis, the union of all protein spots/abundance ratios that had been found differential for any of the probed substances was used.

The question arises if the differential spots identified represent specific markers of embryotoxicity or are simply more abundant than other proteins. With the applied proteomic technique, typically some thousand protein spots per gel are detected covering spot volumes of more than three orders of magnitude. From these, normally the vast majority is found to be not differentially affected and only between 10 and 100 differentials are identified. Their absolute sizes may cover a wide range of spot

volumes, but rarely the most abundant protein spots are among these differentials. Thus, the applied technique definitely does not only detect the more abundant proteins. What can be said about specificity so far is that there is a marked coincidence of the observed volume changes in many of these spots for all substances from one and the same cluster. So, yes there is specificity of the current observations. Nonetheless, the authors are well aware of the fact that it is not entirely clear yet what the observed changes in biomarker abundance are actually specific for. More research is needed to identify a (small) set of biomarkers that can be used to reliably predict different types of embryotoxicity of an unknown potential toxicant.

In a recent publication, concerns have been raised regarding the predictive power of cluster analyses in gene expression studies (59). This article criticizes “a spurious claim of correlation between clusters and clinical outcome, made after clustering samples using a selection of outcome-related differentially expressed genes“. In the present study, all spots were used for clustering the volume of which was significantly changed by any of the assessed substances (which had carefully been selected by an expert committee to include representatives of all possible types and strengths of embryotoxic effects). On the other hand, the relevant “outcome” in this analysis is “membership in a class of substances with similar effects on protein abundances”. Obviously, these criteria for spot selection and outcome assignment are entirely unrelated to each other. In consequence, the selection of spots was not “outcome-related” as presupposed by the cited article and hence the article’s concerns do not apply. Had the outcome been related to toxic effects in general, the criticism might hold true. E.g., if the various effect vectors of toxicant treatments along with those of solute “treatments” had been used for clustering there is little doubt that in one of the principal clusters all toxicants and in another one all solutes (plus, maybe, some non-toxic substances) would have been grouped together.

Supplementary Table1S: Mass spectrometric identifications of relevant proteins. The table shows all accepted proteins in all spots in which one of the proteins from Tables 1 or 3 was identified, along with pertinent information relating to the detection process: Mascot PMF score, expectation value, sequence coverage, number of matching and not matching masses in the respective Mascot identification, molecular mass and pI of the identified protein, interpolated molecular mass and pI of the analyzed spot in the gel. For spots with an asterisk in the last column, MS/MS sequence data were obtained for one or more peptides (see supplementary Table 2S).

spot #	Identified protein	Accession number	PMF score	Expec-tation	Sequence coverage	Matched masses	No matches	Molec. mass	pI	Mol. mass on gel	pI on gel
288	Ran binding protein 5 [Mus musculus]	gi|12057236	199	2.8E-15	24.9 %	21	15	128298	4.5	139000	4.
288	RAN binding protein 5 [Mus musculus]	gi|29789199	199	2.8E-15	25.0 %	21	15	128115	4.6	139000	4.
288	RAN binding protein 5, isoform CRA_a [Mus musculus]	gi|148668272	187	4.4E-14	24.5 %	20	16	125104	4.8	139000	4.
288	Ranbp5 protein [Mus musculus]	gi|32451775	163	1.1E-11	24.2 %	18	18	121417	4.6	139000	4.
507	Heat shock protein 8 [Mus musculus]	gi|42542422	289	2.8E-24	45.2 %	28	30	71536	5	83000	5.5
507	heat shock protein 9 [Mus musculus]	gi|6754256	288	3.6E-24	48.2 %	30	25	74368	6	83000	5.5
507	stress-70 protein (PBP74/CSA) [Mus musculus domesticus]	gi|903309	286	5.7E-24	48.2 %	30	25	74301	5.9	83000	5.5
507	Hspa8 protein [Mus musculus]	gi|63101351	255	7E-21	43.1 %	26	32	69444	5.4	83000	5.5
522	Heat shock protein 8 [Mus musculus]	gi|42542422	297	4.6E-25	48.5 %	32	29	71536	5	79500	5.3
523	Heat shock protein 8 [Mus musculus]	gi|42542422	303	1.1E-25	49.4 %	31	27	71536	5	80000	5.4
524	Heat shock protein 8 [Mus musculus]	gi|42542422	281	1.7E-23	45.7 %	29	18	71536	5	77500	5.
524	Hspa8 protein [Mus musculus]	gi|63101351	261	1.7E-21	44.2 %	28	19	69444	5.4	77500	5.
540	ras-GTPase-activating protein SH3-domain binding protein [Mus musculus]	gi|7305075	248	3.5E-20	59.4 %	25	41	51974	5.3	71000	5.
540	mKIAA4115 protein [Mus musculus]	gi|60359872	233	1.1E-18	47.7 %	23	21	56298	6.4	71000	5.

541	ras-GTPase-activating protein SH3-domain binding protein [Mus musculus]	gi|7305075	168	3.5E-12	46.7 %	22	57	51974	5.3	77500	5.6
541	mKIAA4115 protein [Mus musculus]	gi|60359872	150	2.2E-10	40.6 %	21	58	56298	6.4	77500	5.6
544	calreticulin [Mus musculus]	gi|6680836	142	1.4E-09	38.5 %	11	13	48496	4.1	77000	4.
547	dihydropyrimidinase-like 2 [Mus musculus]	gi|40254595	124	8.8E-08	30.1 %	11	26	63478	6.3	80000	6.
547	Ulip2 protein [Mus musculus]	gi|1915913	107	4.4E-06	26.2 %	10	27	63371	6.3	80000	6.
550	ras-GTPase-activating protein SH3-domain binding protein [Mus musculus]	gi|7305075	241	1.7E-19	43.9 %	22	29	51974	5.3	72000	5.
550	mKIAA4115 protein [Mus musculus]	gi|60359872	217	4.4E-17	38.0 %	21	30	56298	6.4	72000	5.
558	dihydropyrimidinase-like 2 [Mus musculus]	gi|40254595	90	0.00018	28.0 %	9	19	63478	6.3	79000	6.
558	Ulip2 protein [Mus musculus]	gi|1915913	76	0.0054	24.1 %	8	20	63371	6.3	79000	6.
573	stress-induced phosphoprotein 1 [Mus musculus]	gi|14389431	95	0.000063	31.9 %	14	43	64490	6.8	78000	6.
573	Stress-induced phosphoprotein 1 [Mus musculus]	gi|13277819	93	0.000092	27.3 %	13	25	64476	6.8	78000	6.
605	stress-induced phosphoprotein 1 [Mus musculus]	gi|14389431	248	3.6E-20	46.6 %	27	33	64490	6.8	73000	6.6
605	Stress-induced phosphoprotein 1 [Mus musculus]	gi|13277819	232	1.4E-18	43.6 %	26	34	64476	6.8	73000	6.6
624	stress-induced phosphoprotein 1 [Mus musculus]	gi|14389431	195	7E-15	39.0 %	19	16	64490	6.8	73000	6.
624	Stress-induced phosphoprotein 1 [Mus musculus]	gi|13277819	179	2.8E-13	36.1 %	18	17	64476	6.8	73000	6.
624	stress-induced phosphoprotein 1 [Rattus norvegicus]	gi|20302113	149	2.8E-10	35.5 %	16	19	64478	6.8	73000	6.
684	calreticulin [Mus musculus]	gi|6680836	114	9.1E-07	45.0 %	15	71	48496	4.1	61000	4.
685	calreticulin [Mus musculus]	gi|6680836	138	3.5E-09	45.7 %	15	51	48496	4.1	59000	4.4
715	CCT (chaperonin containing TCP-1) beta subunit [Mus musculus]	gi|468546	179	2.8E-13	44.1 %	20	29	58474	6.4	60500	6.
715	chaperonin subunit 2 (beta) [Mus musculus]	gi|126521835	179	2.8E-13	44.1 %	20	29	58504	6.4	60500	6.
715	chaperonin subunit 2 (beta), isoform CRA_b [Mus musculus]	gi|148689878	85	0.0007	33.2 %	9	24	45028	6.1	60500	6.
717	Fscn1 protein [Mus musculus]	gi|144719132	317	4.4E-27	63.2 %	30	46	53597	6.6	60500	6.4
717	Unknown (protein for IMAGE:3599038) [Mus musculus]	gi|116283253	316	5.5E-27	62.8 %	30	46	53868	6.9	60500	6.4
717	fascin homolog 1, actin bundling protein [Mus musculus]	gi|113680348	311	1.7E-26	59.6 %	30	46	56776	6.9	60500	6.4

717	fascin [Mus musculus]	gi|497775	195	7E-15	44.4 %	20	33	56673	6.6	60500	6.4
737	Fscn1 protein [Mus musculus]	gi|144719132	346	5.5E-30	55.1 %	26	15	53597	6.6	56000	6.
737	Unknown (protein for IMAGE:3599038) [Mus musculus]	gi|116283253	345	7E-30	54.7 %	26	15	53868	6.9	56000	6.
737	fascin homolog 1, actin bundling protein [Mus musculus]	gi|113680348	340	2.2E-29	51.9 %	26	15	56776	6.9	56000	6.
738	Unknown (protein for IMAGE:3599038) [Mus musculus]	gi|116283253	104	8.8E-06	32.3 %	12	68	53868	6.9	56000	6.
738	Fscn1 protein [Mus musculus]	gi|144719132	104	8.8E-06	32.5 %	12	68	53597	6.6	56000	6.
738	fascin homolog 1, actin bundling protein [Mus musculus]	gi|113680348	101	0.000017	30.6 %	12	68	56776	6.9	56000	6.
739	Heat shock protein 8 [Mus musculus]	gi|42542422	90	0.00018	25.9 %	13	32	71536	5	56000	6.
749	retinoblastoma binding protein 7 [Mus musculus]	gi|123223058	110	2.2E-06	40.2 %	11	31	44808	5.2	54000	4.
749	retinoblastoma binding protein 7 [Mus musculus]	gi|123223060	102	0.000014	35.1 %	9	36	47941	4.9	54000	4.
753	Heat shock protein 8 [Mus musculus]	gi|42542422	118	3.5E-07	24.9 %	13	18	71536	5	56333	6.3
992	PREDICTED: similar to Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38) [Mus musculus]	gi|149251776	132	1.4E-08	50.7 %	14	49	33177	4.6	40000	4.
992	nucleophosmin 1 [Mus musculus]	gi|6679108	115	7E-07	42.5 %	13	50	33071	4.4	40000	4.
992	Nucleophosmin 1 [Mus musculus]	gi|55153941	115	7E-07	42.5 %	13	50	33099	4.4	40000	4.
992	nucleophosmin 1 [Mus musculus]	gi|56206422	81	0.0015	38.3 %	11	56	30037	4.2	40000	4.
993	PREDICTED: similar to Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38) [Mus musculus]	gi|149251776	105	0.000007	46.9 %	12	55	33177	4.6	39000	4.
993	nucleophosmin 1 [Mus musculus]	gi|56206422	99	0.000028	34.5 %	8	21	30037	4.2	39000	4.
993	nucleophosmin 1 [Mus musculus]	gi|6679108	96	0.000045	31.2 %	8	21	33071	4.4	39000	4.
993	Nucleophosmin 1 [Mus musculus]	gi|55153941	96	0.000045	31.2 %	8	21	33099	4.4	39000	4.
1314	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	gi|13435747	128	3.5E-08	41.7 %	8	33	23554	4.8	28500	5.0
1314	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	gi|31982030	128	3.5E-08	45.6 %	9	32	23570	4.8	28500	5.0
1330	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	gi|13435747	86	0.00046	38.7 %	9	29	23554	4.8	28000	5.

1330	Rho GDP dissociation inhibitor (GDI) alpha [Mus musculus]	gi|31982030	86	0.00046	38.7 %	9	29	23570	4.8	28000	5.
1424	Heat shock protein beta-1 (HspB1) (Heat shock 27 kDa protein) (HSP 27) (Growth-related 25 kDa protein) (P25) (HSP25)	gi|547679	110	2.2E-06	49.3 %	8	24	23177	6.5	26000	6.
1424	heat shock protein 1 [Mus musculus]	gi|7305173	93	0.00009	45.7 %	7	25	23064	7	26000	6.

Supplementary Table2S: MS/MS sequence identifications of individual peptides from the proteins of Table 1S marked by an asterisk in the rightmost column. The table shows the peptide sequences along with pertinent information relating to the detection process: precursor mass, ion charge z, Mascot ion score, expectation value, MS/MS technique (MALDI postsource decay PSD or QTOF collision inducted dissociation CID. Note the insignificant expectation value for spot 540 "Phosphor" which is only meant as an additional confirmation of the phosphorylation shown for spot 541.

spot #	Precurser mass	z	Ion score	Expec-tation	Sequence	PSD or CID
507	833.413	2	83	2.70E-06	^{57}NQVAMNPTNTVFDAK^{71} +Oxidation (M)	CID
522	596.676	3	92	9.40E-08	^{172}IINEPTAAAIAYGLDKK ^{188}	CID
523	744.369	2	86	1.30E-06	^{37}TTPSYVAFTDTER^{49}	CID
524	600.350	2	52	0.0026	^{160}DAGTIAGLNVLR^{171}	CID
540	902.914	2	27	1.5	^{229}STSPAPADVAPAQEDLR^{245} +(Phosphor)	CID
540	862,927	2	48	0.014	^{229}STSPAPADVAPAQEDLR^{245}	CID
540	578.270	2	58	0.00053	^{319}EAGEPGDVEPR^{329}	CID
541	902.914	2	43	0.0016	^{229}STSPAPADVAPAQEDLR^{245} +(Phosphor)	CID
550	578.280	2	58	0.00073	^{319}EAGEPGDVEPR^{329}	CID
605	558.834	2	65	7.30E-05	^{534}LMDVGLIAIR^{543}	CID
685	2976.468	1	100	4.60E-08	^{163}CKDDEFTHLYTLIVRPDNTYEVK^{185} +BHPECAM ©	PSD
704	665.839	2	79	7.90E-06	^{377}GATQQILDEAER^{388}	CID
715	654.338	2	80	5.20E-06	^{237}ILIANTGMDTDK^{248} +Oxidation (M)	CID
717	537.810	2	73	2.20E-05	^{399}KVTGTLDANR^{408}	CID
737	564.796	2	90	5.30E-07	^{91}FLVVAHDDGR^{100}	CID
749	1909.894	1	73	5.70E-05	^{147}HPAKPDPSGECNPDLR^{162} +BHPECAM ©	PSD
992	784.882	2	111	4.20E-09	^{33}VDNDENEHQLSLR^{45}	CID
992	565.816	2	65	0.0002	^{238}GPSSVEDIKAK^{248}	CID
993	565.817	2	48	0.0097	^{238}GPSSVEDIKAK^{248}	CID
1314	1650.909	1	83	2.80E-06	^{59}VAVSADPNVPNVIVTR^{74}	PSD
1424	538.293	2	67	0.00013	^{84}QLSSGVSEIR^{93}	CID