

Supplementary figures:

Fig. 1. Perl Scripts for bioinformatics (seqone.pl, unigene.pl, blastresult.pl)—page 2.

Fig. 2. A list of significant genes based on SAM and Affymetric microarray—page 7.

Fig.3. A list of significant genes based on SAM and custom microarray—page 9.

Fig.4. Protein sequence comparison of mouse and human NDG2--page 12. Arrows indicate the four conserved cysteine residues that might be involved in metal binding.

Fig.5. Mitochondria localization of NDG2 protein—page 13.

Fig. 6. NDG1 but not NDG2 transgenic thymocytes from adult mice showed increased propensity to undergo apoptosis--page 14.

```

#!/usr/bin/perl -w

#Written by Astar Winoto This program will merge several files in the
folder given and put them
#into one file. For each file, we put > and the file name. This file will
#now be ready to be put into blastcl3 for batch blasting.
#further modified to test for the quality of the sequence, this is now
simply done
#by calculating the number of N in the sequence. If the fraction is
greater than
#0.50 or 50%, only the name but not the sequence will be printed.

print "what folder hold the sequence information?\n";
chomp ($where=<STDIN>);
print "what is the output file going to be named?\n";
chomp ($outfile=<STDIN>);
open (OUTFILE, ">$outfile");

chdir ("$where") || die "can not change to the folder mentioned\n";
#change to the directory and read the file names, sort first and then print
#with the name preceded by the fasta file sign >
foreach (sort by_mostlyname<*>) {
    print OUTFILE (">$_", "\n");
    open (DATAFILE, "$_") || die "can not open files";
    while (<DATAFILE>) {
#will only write if the sequence quality is good
        if (sequence_test($goodseq)=~/1/) {
            print OUTFILE ("$_", "\n");
        }
    }
}

#subroutine sorting
sub by_mostlyname {
($a cmp $b || $a <=> $b);
}

#subroutine to test if the sequence quality is good or not
sub sequence_test {
    @seqtest=split('',$_);
    $numberN=0;
    foreach $seq(@seqtest) {
        if ($seq=~ /N/) {
            $numberN++;
        }
    }
}

```

```
    }
    if ($numberN/($#septest+1)>0.50) {
#the sequence is not good
    return $goodseq=0;
    } else {
#the sequence is good
    return $goodseq=1;
    }
}
```



```
        if ($num=~ /ACC=(.*)/) {
            $genbank=$1;
            $genbank=~ s/\;//g;
        } elsif ($num=~ /CLONE\=(.*)/) {
            #still have to solve the problem that E is treated as power multiplier--
            #this is the problem of excel and not perl. Call the output file as a word
            #document...add .doc at the end.
            $clone=$1;
            $clone=~ s/\;//g;
        }
    }
    #the position of write is very important
    write UNIGENEOUT;
        } else {
        }
    }
}
```


38 positive

Row	Gene Name	Gene ID	Score(d)	Numerator(r)	Fold Change
913	102371_AT X16995 X16995:Hormone receptor /cds=(2,1807) /gb=X	X16995	10.17571	973.125	15.91379
1241	96900_AT AW125480 AW125480:UI-M-BH2.2-aqm-c-05-0-UI.s1	Mus AW125480	5.78844	550.5	7.68285
3928	101412_AT U58887 U58887:SH3 domain protein 2C /cds=(159,1202) U58887		5.35732	126.25	6.02488
102	103992_F_AT AA636966 AA636966:vn05b05.r1	Mus musculus cDNA, AA636966	9.177354	1905.75	5.66524
5963	92676_AT D14636 D14636:Core binding factor alpha 1 /cds=(1015,2	D14636	2.314641	28.875	4.20833
4972	98423_AT M81445 M81445:Gap junction membrane channel protein l	M81445	3.720701	102.75	3.75839
4624	97113_AT U06948 U06948:Fas antigen ligand /cds=(124,963) /gb=U	U06948	3.384206	93.4375	2.9749
1789	96344_AT X61600 X61600:M.musculus gene for beta-enolase /cds=(1	X61600	5.397943	374.25	2.79389
3065	104445_AT AW046694 AW046694:UI-M-BH1-alm-d-07-0-UI.s1	Mus r AW046694	2.927295	102.625	2.7431
3599	98836_AT X67914 X67914:M.musculus PD-1 mRNA /cds=(63,929) /g	X67914	3.669062	147.375	2.62621
2148	96095_I_AT AI648018 AI648018:uk39c03.x1	Mus musculus cDNA, 3' AI648018	3.56967	198.625	2.52495
1448	96096_F_AT AI648018 AI648018:uk39c03.x1	Mus musculus cDNA, 3' AI648018	3.097433	248.125	2.34759
3946	103715_AT U04354 U04354:Mus musculus ADSEVERIN mRNA, compl	U04354	2.568527	206.75	2.33603
1102	99030_AT M29697 M29697:Interleukin 7 receptor /cds=(47,1426) /g	M29697	4.614432	435	2.32319
1569	96666_AT U57692 U57692:N-terminal Asn amidase /cds=(33,965) /g	U57692	2.798845	244.5	2.28094
3078	97551_AT AI846182 AI846182:UI-M-AM1-agb-c-09-0-UI.s1	Mus mus AI846182	2.494474	118	2.17559
2426	94369_AT AW123026 AW123026:UI-M-BH2.1-apb-f-07-0-UI.s1	Mus r AW123026	4.153287	173.75	2.16513
5172	95664_AT AW048159 AW048159:UI-M-BH1-amc-a-01-0-UI.s1	Mus r AW048159	2.660687	45.125	2.13522
290	102695_AT AF021335 AF021335:Mus musculus T cell receptor gamm	AF021335	3.571438	1063.75	2.07626
4993	102955_AT U83148 U83148:Mus musculus NFIL3/E4BP4 transcription	U83148	2.354288	61.625	2.02495
1119	103367_AT U18975 U18975:Ganglioside expression 2 /cds=(3,1604)	U18975	2.594561	289.125	1.9957
3382	98554_AT AB033922 AB033922:Mus musculus mRNA for Ndr1 relatec	AB033922	3.03691	86.875	1.97067
3234	102299_AT M25811 M25811:Protein kinase C, alpha /cds=(255,2273	M25811	3.036649	137.375	1.94012
2597	95665_AT AI852087 AI852087:UI-M-BH0-aja-g-04-0-UI.s1	Mus musc AI852087	2.301292	150.5	1.9012
2403	92558_AT M84487 M84487:Vascular cell adhesion molecule 1 /cds=(1	M84487	2.766707	134.875	1.87016
2377	97977_AT AA645293 AA645293:vs81a07.r1	Mus musculus cDNA, 5' e AA645293	2.292333	119	1.85689
4666	102676_AT U12570 U12570:Von Hippel-Lindau syndrome homolog /c	U12570	2.540234	48.5	1.776
1941	93319_AT U20238 U20238:Mus musculus GTPase-activating protein (U20238	2.664754	159	1.75
1570	97718_AT X05719 X05719:Mouse mRNA for immunoglobulin superfar	X05719	2.97096	240.75	1.73737
3827	103739_AT AW230977 AW230977:uo70h08.y1	Mus musculus cDNA, !AW230977	2.304105	61.25	1.71118
1832	101014_AT Y09864 Y09864:M.musculus mRNA for type I interferon r	e Y09864	2.565719	173.375	1.584
685	97553_AT AI841884 AI841884:UI-M-AO0-acd-c-10-0-UI.s1	Mus mus AI841884	3.007479	462	1.56949
1740	97819_AT AI843119 AI843119:UI-M-AK1-aes-f-05-0-UI.s1	Mus musc AI843119	2.421513	209.25	1.55837

1176	95721_AT	AW120722	AW120722:UI-M-BH2.3-any-a-09-0-UI.s1	Mus	AW120722	2.284876	227.5	1.51398	
754	104115_AT	AA867340	AA867340:vx97a11.r1	Mus musculus	cDNA, 5'	AA867340	2.509063	322.125	1.44339
451	103574_AT	AI841606	AI841606:UI-M-AP0-abn-h-07-0-UI.s1	Mus musculus	cDNA, 3' end /clone=UI-M	AI841606	2.725387738	405.25	1.39273
1795	97405_AT	M28489	M28489:Mus musculus (clone Mu6a)	ribosomal	pr	M28489	2.518512	146	1.31389
18	101906_AT	AA032310	AA032310:MTA.G11.085.A	Mus musculus	cDN	AA032310	3.953102	819.25	1.30728

97 Positive Significant Genes

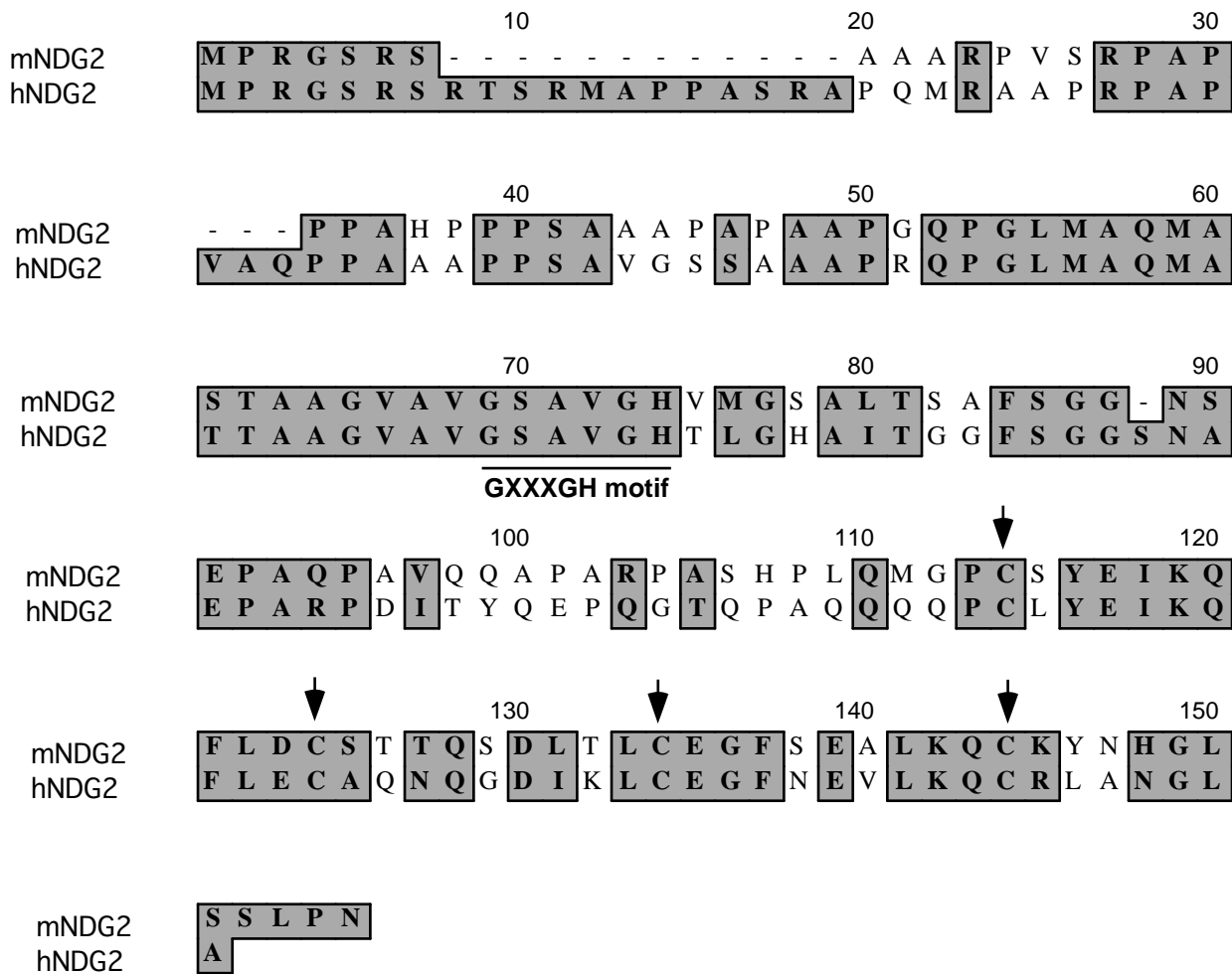
Row	Gene Name	Gene ID	Score(d)	Fold Change	q-value (%)
291	Nur77	Mm.119	5.659865494	9.22256	2.43902439
988	ESTs Highly similar t	BE633821	3.148979026	3.59294	7.52688172
634	TGF-beta3	Mm.1291	3.512046456	3.34810	5.633802817
150	TCF-4	Mm.4269	7.450914529	3.28991	2.43902439
950	schlafen 3	AA175847	4.238699586	3.01359	2.43902439
1270	ESTs Moderately sim	BE377985	14.08474612	26.05181	2.43902439
307	Interleukin 7 recepto	AI505954	5.534961302	2.99176	2.43902439
1314	COUP-TFII	Mm.16519	4.475303678	2.91209	2.43902439
1333	M.musculus mRNA f	X03802	5.497355906	2.78962	2.43902439
1908	Wasp	Mm.4735	4.000653404	2.71211	2.43902439
1071	CTLA-4	Mm.390	3.239997192	2.62767	7.52688172
1296	ESTs Weakly similar	AI450144	6.845136874	2.41965	2.43902439
1811	thyroid hormone rece	AI173152	3.681532386	2.25341	3.636363636
712	DNA segment Chr 7	AA216881	3.472860072	2.24826	5.633802817
978	Toll like receptor 5	Mm.116894	3.295911214	2.23911	7.52688172
1261	TCRg-V4	Mm.1289	4.297669707	2.22211	2.43902439
1754	Mouse mitochondriol	EMPTY	3.029457995	2.20650	7.920792079
1282	Hexokinase 2	AI021011	5.265096018	2.20447	2.43902439
1274	T-cell receptor gamr	NM_011558	4.566807197	2.18235	2.43902439
1850	fibroblast growth fact	BE136182	3.334068281	2.15016	7.52688172
506	growth hormone	AW742222	3.793327061	2.13427	3.636363636
1255	histocompatibility 2 C	AA690320	5.400901763	2.12830	2.43902439
1817	glycoprotein 49 A	AA175494	4.639243706	2.12209	2.43902439
1248	Mus musculus mRN	AI561460	5.097472277	2.10681	2.43902439
1293	MEKK5	Mm.6595	3.461059395	2.08579	5.633802817
1325	MMP-4	Mm.20935	3.622259287	2.02950	3.636363636
705	Ksr	Mm.4745	3.223322612	2.02227	7.52688172
1848	p38-2G4 proliferatio	Mm.4742	3.635888306	2.01202	3.636363636
1232	E2F transcription fac	AA396123	3.242063462	1.98553	7.52688172
1077	integrin alpha V (Cd	BE309701	3.58490815	1.97691	5.633802817
623	?small inducible cyto	AA217348	5.86583995	1.92913	2.43902439
1379	Mus musculus TNF-r	AA163835	3.178532416	1.90625	7.52688172
655	Zinc finger protein 2	BF682561	3.038433543	1.89688	7.920792079
121	Mus musculus genor	AA760194	3.686366669	1.89465	3.636363636
1288	Histocompatibility 2	AA200325	3.421689918	1.89359	5.633802817
2181	?Annexin VI p6	AA178221	3.023854642	1.89338	7.920792079
82	TBX1	Mm.156370	4.208473926	1.88222	2.43902439
1728	?integrin alpha 4 (Cd	AI322934	3.2975375	1.88054	7.52688172
147	Mad4	Mm.38795	3.133472206	1.86663	7.52688172
559	Mus musculus resisti	AF323080	3.354891937	1.86308	5.633802817
554	RAD50	Mm.4888	3.986083394	1.85002	2.43902439
187	wingless-related MM	W29961	3.023748035	1.84921	7.920792079
1867	sulfotransferase hydr	AA267313	4.05690327	1.84615	2.43902439
134	melanoma antigen 8	AA726940	4.209002552	1.84300	2.43902439
1227	IRAK	Mm.38241	3.853977065	1.84122	3.636363636
1254	ESTs Weakly similar	AI847419	6.124521471	1.82438	2.43902439

1598 fibroblast growth fact	AA689962	3.160501878	1.82246	7.52688172
1305 Tyk-2	Mm.20249	3.942205263	1.81663	2.43902439
452 retinoic acid early tra	BE852370	3.028632275	1.81058	7.920792079
1801 H2-D	Mm.33263	3.968683669	1.81056	2.43902439
1846 ESTs	BE853713	3.125273768	1.79668	7.52688172
1291 AU042671	Mm.184589	3.549517761	1.78226	5.633802817
2210 m-IAP3	Mm.23345	3.253440245	1.76787	7.52688172
758 shp-2	Mm.14844	3.130869052	1.76488	7.52688172
1220 lymphocyte antigen ε	D13695	3.665166242	1.75668	3.636363636
1250 HMG-I(Y)	Mm.4438	5.507523424	1.74583	2.43902439
925 seC2-1 PLCg1	EMPTY	3.074976414	1.73566	7.920792079
1691 Mus musculus 18 da	AK003687	3.371975245	1.70552	5.633802817
1233 Caspase	BE847020	3.430307789	1.69925	5.633802817
1830 Protein phosphatase	AA200189	3.107741901	1.69762	7.52688172
1360 Survivin	Mm.8552	3.787374297	1.69144	3.636363636
519 BM88 antigen	W13012	4.252826151	1.67784	2.43902439
1230 ESTs Highly similar t	AI837865	3.027714233	1.66225	7.920792079
1322 tumor necrosis factor	M59378	3.876698502	1.65784	3.636363636
1357 p80HT ?	BF320592	3.251280588	1.64219	7.52688172
764 RIKEN cDNA 50314f	Mm.16769	5.762156968	1.64099	2.43902439
1226 p19Ink4d	Mm.29020	3.169336037	1.64011	7.52688172
1324 Lyn	Mm.1834	3.671067355	1.63504	3.636363636
1268 Mus musculus dual ε	NM_011932	4.12335128	1.60908	2.43902439
1800 CHK1	Mm.16753	3.344817035	1.60439	5.633802817
60 Mus musculus Simile	BC004778	4.275273402	1.59832	2.43902439
1019 Lymphocyte antigen	AA109204	3.154222166	1.58385	7.52688172
1856 protein-tyrosine phos	AA217443	3.563689054	1.57813	5.633802817
1257 golgi autoantigen gol	AA546770	3.695903903	1.57614	3.636363636
570 MCBP-4	Mm.20890	9.325961973	1.57405	2.43902439
1114 TNFIP3	Mm.56971	4.643234811	1.55055	2.43902439
1223 Interleukin 6 signal tr	AA543149	3.831451007	1.52388	3.636363636
74 cyclin B2	AA189311	3.148440899	1.52216	7.52688172
1343 Prss16	Mm.32043	4.281550711	1.51593	2.43902439
525 caspase 2	Mm.3921	3.508093031	1.47364	5.633802817
104 large tumor suppressc	AI594372	3.472709406	1.46436	5.633802817
1218 ESTs Moderately sin	AA795608	4.242889546	1.44619	2.43902439
43 RIKEN cDNA 26100f	BG145132	3.393833521	1.43372	5.633802817
1356 ezrin	Mm.4551	3.701353763	1.41826	3.636363636
500 Interleukin 11	AA104491	7.110094227	1.41783	2.43902439
157 MEK1	Mm.1059	3.127941654	1.40878	7.52688172
156 p38beta	Mm.4437	3.763951801	1.39811	3.636363636
1240 HIF1alpha	Mm.3879	3.549022077	1.39110	5.633802817
1909 91% to unnamed hur	Mm.196371	3.099872487	1.38956	7.52688172
55 cyclinE	Mm.16110	4.070335607	1.37238	2.43902439
25 identical to EST BM2	BM248519	5.546103022	1.36021	2.43902439
1869 hsp60	Mm.1777	4.090765372	1.35928	2.43902439
2225 ESTs Moderately sin	AI464132	4.510307352	1.28410	2.43902439
2026 lymphoid blast crisis-	BE631635	3.135179807	1.27514	7.52688172
1301 lymphocyte antigen γ	BE634230	3.442676261	1.25857	5.633802817

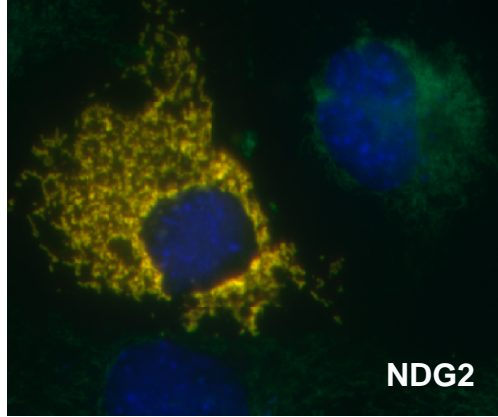
2237 Noxa	Mm.143798	3.203304666	1.13025	7.52688172
57 ESTs Weakly similar	AI616314	3.075522021	1.08053	7.920792079

4 Negative Significant Genes

Row	Gene Name	Gene ID	Score(d)	Fold Change	q-value (%)
766	tcr a 8	Mm.57186	-5.832362348	0.59147	2.43902439
2268	macrophage activati	BE633577	-5.205881244	0.79628	2.43902439
1133	?T-cell receptor alph	AA939602	-5.086439257	0.55680	2.43902439
2182	dual specificity phos	Mm.1791	-4.84494335	0.60012	2.43902439

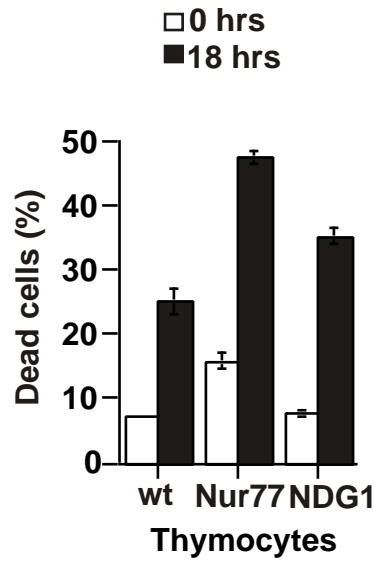


mitochondria localization of NDG2 protein was predicted by the MITOPROP program located in <http://mips.gsf.de/proj/medgen/mitop/>



Blue: DAPI staining
Red: NDG2
Green: HSP60

A



B

