Supplemental data

Table of Contents

Supplemental figure legends

Fig. S1 **The majority of genes on chromosome 4 were found in large zones that showed coordinately reduced expression in BP cells.** DIGMAP analysis was performed as described in Fig. 1. The entire chromosome showed a gradual reduction in expression in the course of HF1 transformation. (A: triplicate K samples, B: triplicate E samples, C: triplicate L samples, D: triplicate BP samples)

Fig. S2 **The majority of transcripts on chromosome 18 were located within large chromosomal zones with reduced expression in BP cells.** DIGMAP analysis was performed as described in Fig. 1. The entire chromosome showed a gradual reduction in expression in the course of HF1 transformation.

Fig. S3 **Chromosome 5p has large chromosomal zones with induced expression in BP cells.** DIGMAP analysis was performed as described in Fig. 1. Long regions with induced expression in BP cells were observed on chromosome 5.

Fig. S4 **Chromosome 3p has large chromosomal zones with reduced expression in BP cells.** DIGMAP analysis was performed as described in Fig. 1. Long regions with reduced expression in BP cells were observed on chromosome 3p.

Fig.S1 AAA BBB CCC DDD

Supplementary Material (ESI) for Molecular BioSystems
This journal is © The Royal Society of Chemistry 2011

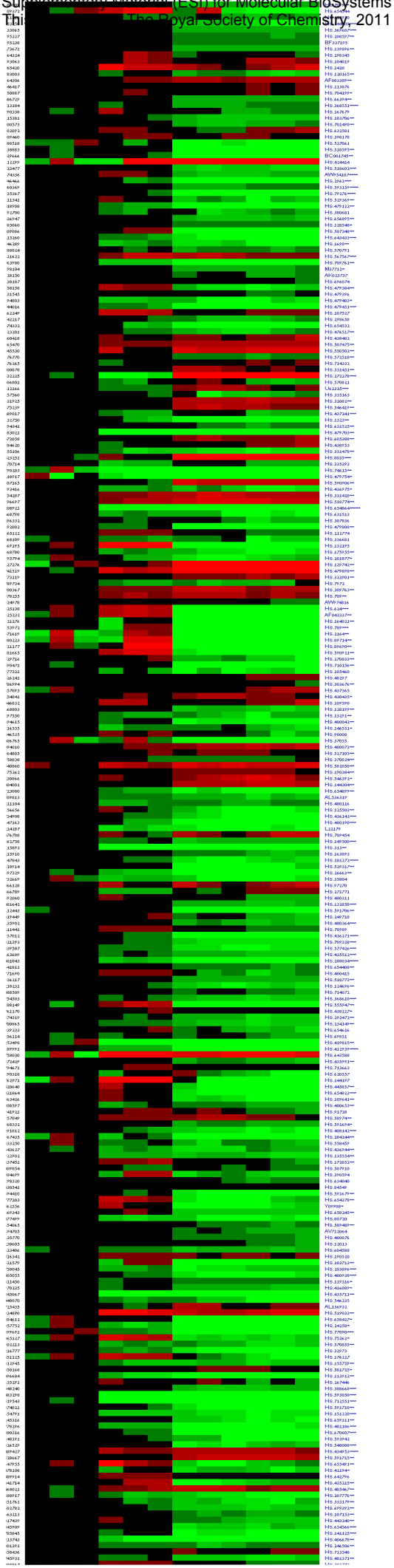


Fig.S2

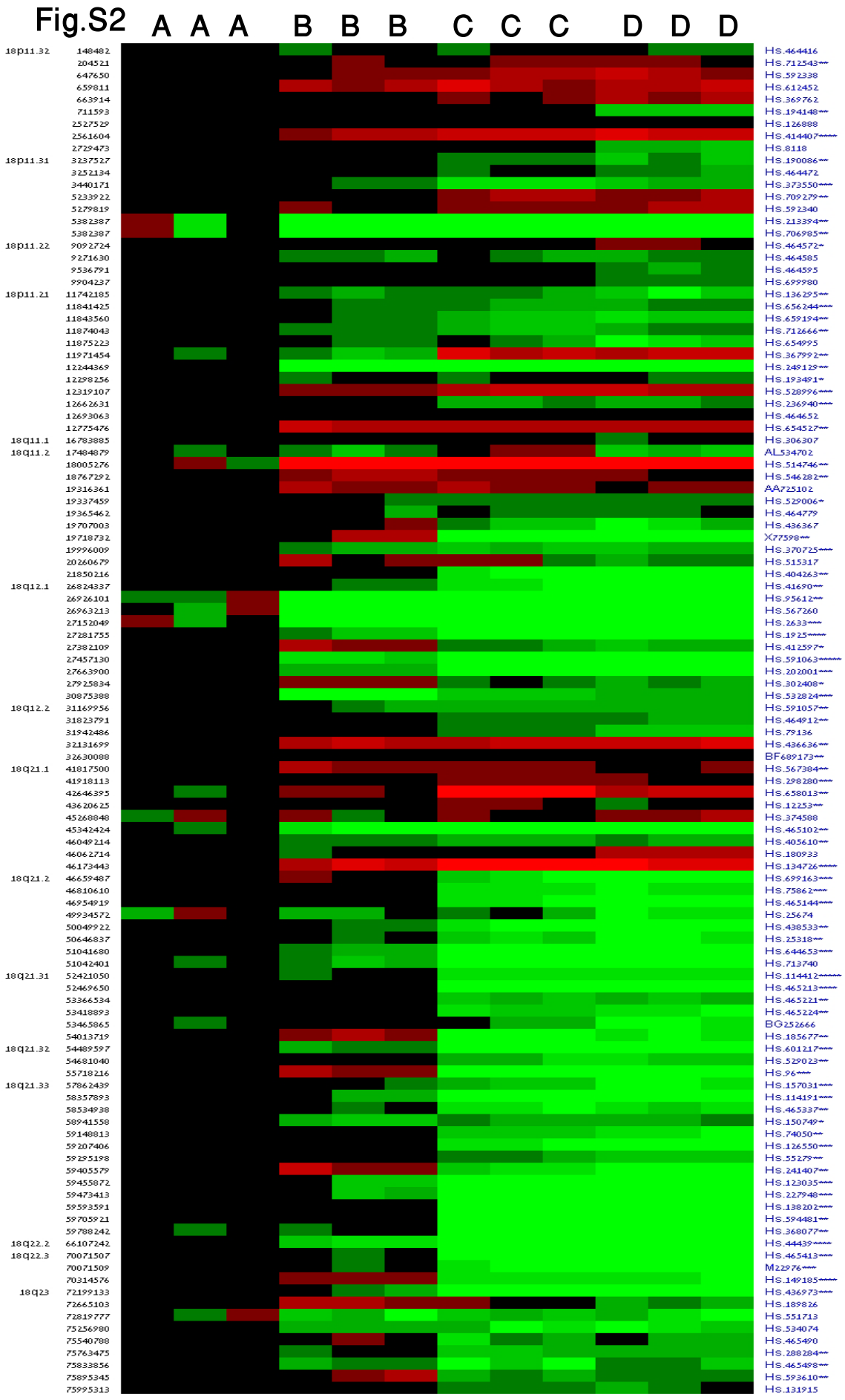
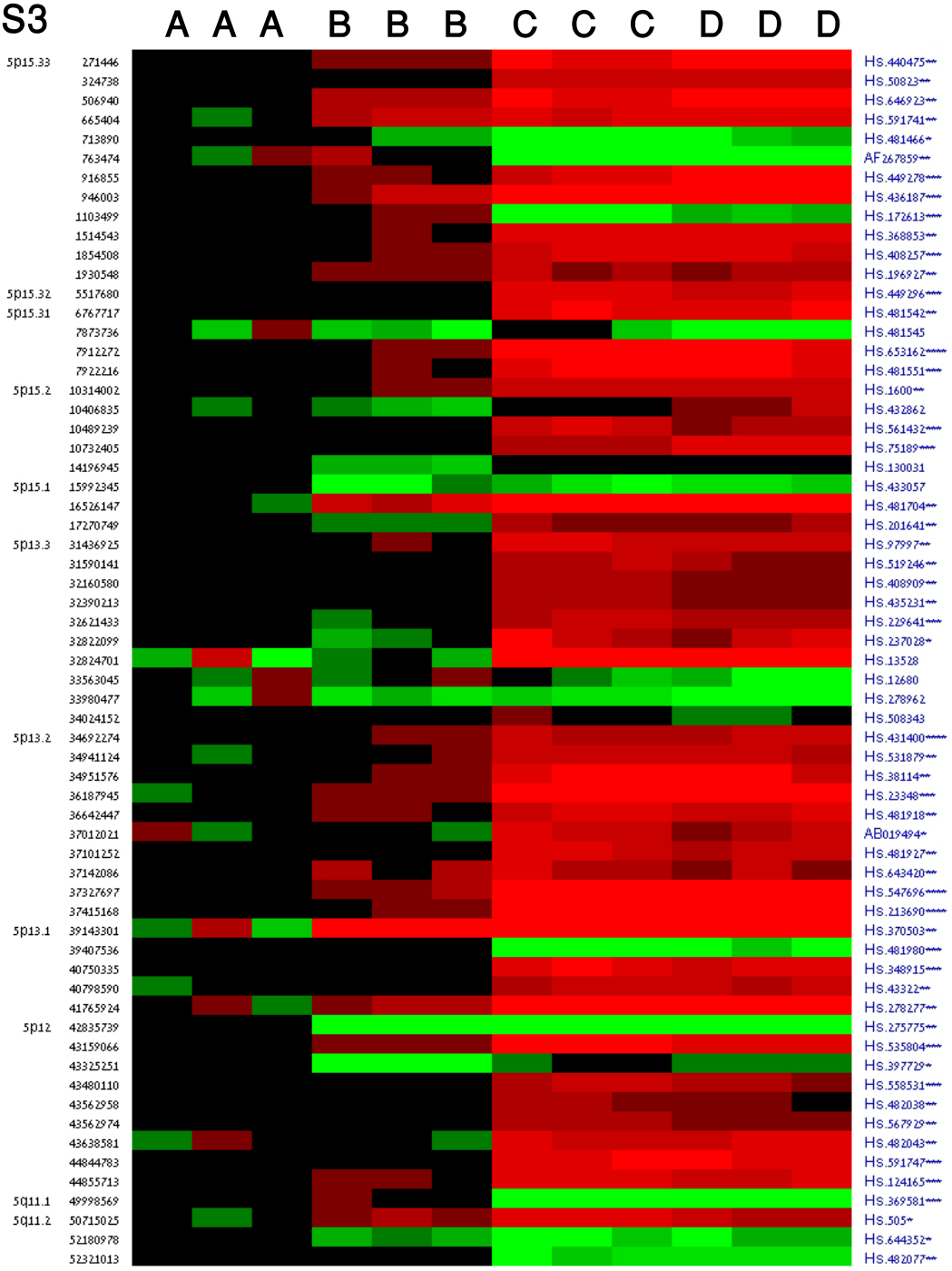

A A A B B B C C C D D D
18p11.32 148482 Hs.464416
204521 Hs.712543**
647650 Hs.592338
659811 Hs.612452
663914 Hs.369762
711593 Hs.194148**
2527529 Hs.126888
2561604 Hs.414407*****
2729473 Hs.8118
18p11.31 3237527 Hs.190086**
3252134 Hs.464472
3440171 Hs.373550****
5233922 Hs.709279**
5279819 Hs.592340
5382387 Hs.213394**
5382387 Hs.706985**
18p11.22 9092724 Hs.464572*
9271630 Hs.464585
9536791 Hs.464595
9904237 Hs.699980
18p11.21 11742185 Hs.136295**
11841425 Hs.656244****
11843560 Hs.659194**
11874043 Hs.712666**
11875223 Hs.654995
11971454 Hs.367992**
12244369 Hs.249129**
12298256 Hs.193491*
12319107 Hs.528996****
12662631 Hs.236940****
12693063 Hs.464652
12775476 Hs.654527**
18q11.1 16783885 Hs.306307
18q11.2 17484879 AL534702
18005276 Hs.514746**
18767292 Hs.546282**
19316361 AA725102
19337459 Hs.529006*
19365462 Hs.464779
19707003 Hs.436367
19718732 X77598**
19996009 Hs.370725****
20260679 Hs.515317
21850216 Hs.404263**
18q12.1 26824337 Hs.41690**
26926101 Hs.95612**
26963213 Hs.567260
27152049 Hs.2633****
27281755 Hs.1925*****
27382109 Hs.412597*
27457130 Hs.591063******
27663900 Hs.202001****
27925834 Hs.302408*
30875388 Hs.532824****
18q12.2 31169956 Hs.591057**
31823791 Hs.464912**
31942486 Hs.79136
32131699 Hs.436636**
32630088 BF689173**
18q21.1 41817500 Hs.567384**
41918113 Hs.298280****
42646395 Hs.658013**
43620625 Hs.12253**
45268848 Hs.374588
45342424 Hs.465102**
46049214 Hs.405610**
46062714 Hs.180933
46173443 Hs.134726*****
18q21.2 46659487 Hs.699163****
46810610 Hs.75862****
46954919 Hs.465144****
49934572 Hs.25674
50049922 Hs.438533**
50646837 Hs.25318**
51041680 Hs.644653****
51042401 Hs.713740
18q21.31 52421050 Hs.114412******
52469650 Hs.465213*****
53366534 Hs.465221**
53418893 Hs.465224**
53465865 BG252666
54013719 Hs.185677**
18q21.32 54489597 Hs.601217****
54681040 Hs.529023**
55718216 Hs.96****
18q21.33 57862439 Hs.157031****
58357893 Hs.114191****
58534938 Hs.465337**
58941558 Hs.150749*
59148813 Hs.74050**
59207406 Hs.126550****
59295198 Hs.55279**
59405579 Hs.241407**
59455872 Hs.123035****
59473413 Hs.227948****
59593591 Hs.138202****
59705921 Hs.594481**
59788242 Hs.368077**
18q22.2 66107242 Hs.44439*****
18q22.3 70071507 Hs.465413****
70071509 M22976****
70314576 Hs.149185******
18q23 72199133 Hs.436973****
72665103 Hs.189826
72819777 Hs.551713
75256980 Hs.534074
75540788 Hs.465490
75763475 Hs.288284**
75833856 Hs.465498**
75895345 Hs.593610**
75995313 Hs.131915

Fig.S3


A A A B B B C C C D D D
5p15.33 271446 Hs.440475***
324738 Hs.50823***
506940 Hs.646923***
665404 Hs.591741***
713890 Hs.481466*
763474 AF267859***
916855 Hs.449278***
946003 Hs.436187***
1103499 Hs.172613***
1514543 Hs.368853***
1854508 Hs.408257***
1930548 Hs.196927***
5p15.32 5517680 Hs.449296***
5p15.31 6767717 Hs.481542***
7873736 Hs.481545
7912272 Hs.653162***
7922216 Hs.481551***
5p15.2 10314002 Hs.1600***
10406835 Hs.432862
10489239 Hs.561432***
10732405 Hs.75189***
14196945 Hs.130031
5p15.1 15992345 Hs.433057
16526147 Hs.481704***
17270749 Hs.201641***
5p13.3 31436925 Hs.97997***
31590141 Hs.519246***
32160580 Hs.408909***
32390213 Hs.435231***
32621433 Hs.229641***
32822099 Hs.237028*
32824701 Hs.13528
33563045 Hs.12680
33980477 Hs.278962
34024152 Hs.508343
5p13.2 34692274 Hs.431400***
34941124 Hs.531879***
34951576 Hs.38114***
36187945 Hs.23348***
36642447 Hs.481918***
37012021 AB019494*
37101252 Hs.481927***
37142086 Hs.643420***
37327697 Hs.547696***
37415168 Hs.213690***
5p13.1 39143301 Hs.370503***
39407536 Hs.481980***
40750335 Hs.348915***
40798590 Hs.43322***
41765924 Hs.278277***
5p12 42835739 Hs.275775***
43159066 Hs.535804***
43325251 Hs.397729**
43480110 Hs.558531***
43562958 Hs.482038***
43562974 Hs.567929***
43638581 Hs.482043***
44844783 Hs.591747***
44855713 Hs.124165***
5q11.1 49998569 Hs.369581***
5q11.2 50715025 Hs.505*
52180978 Hs.644352*
52321013 Hs.482077***

Supplementary Material (ESI) for Molecular BioSystems

Fig.S4

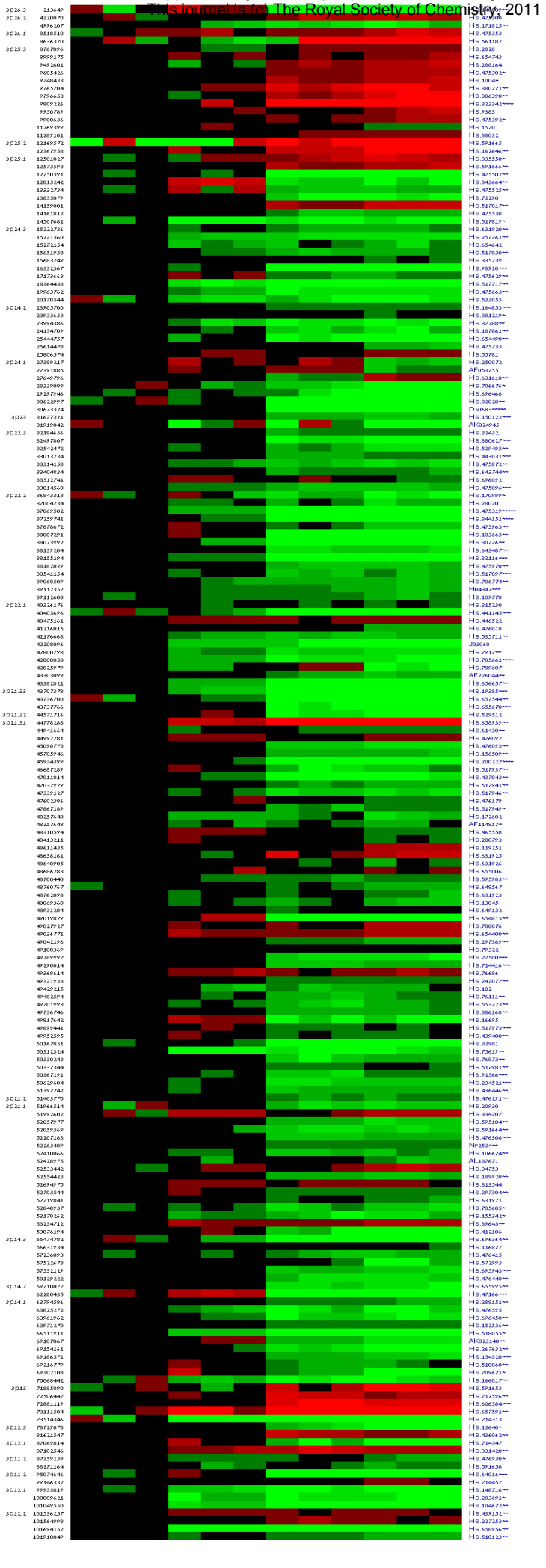
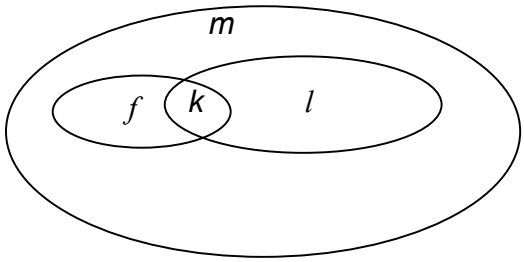


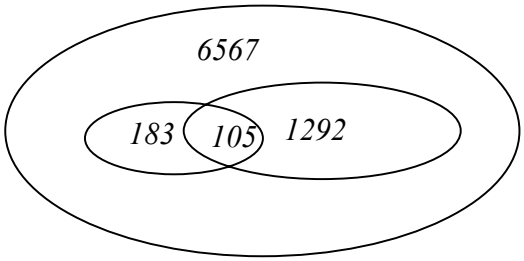
Figure S5

The probability of obtaining an overlap of size k randomly between two sub-populations of size l and f within a population of size m can be calculated as follows:

$$P(n=k) = \frac{\binom{m}{l}\binom{l}{k}\binom{m-l}{f-k}}{\binom{m}{f}\binom{m}{l}}$$



Hence, the probability of obtaining an overlap of 105 or more genes randomly between the genes participating in the apoptotic network with at least 2 fold reduced expression in the BP cells (183 genes) and all the genes located in large reduced regions (1292 out of 6567 significant genes passed t-test analysis) can be calculated as shown below:



$$P(n > k) = \sum_{k=105}^{183} \frac{\binom{6567}{1292}\binom{1292}{k}\binom{6567-1292}{183-k}}{\binom{6567}{183}\binom{6567}{1292}} = 3.0 \times 10^{-30}$$

Fig S5. **Probability of random overlap between genes participating in the apoptotic network with at least 2 fold reduced expression in the BP cells and all the BP genes located in large reduced regions.**

Calculation of contribution of global chromosomal alterations to the changes in the general enzymatic activity of the transformed BP cells

BP cells proliferate much more rapidly than K and E cells but have similar ATP levels and similar rates of O_2 and glucose consumption [1]. The overall reduction in gene expression in BP cells apparently led to a reduction in protein concentration, thereby modifying the general enzymatic activity or "flux" inside the cell. Enzymatic flux is defined as the rate of flow into and out of the pool of metabolites in the cell and is a function of all the enzymes participating in a particular network. We hypothesize that the energy conserved by the shrinkage of many processes, such as cap-dependent translation and apoptosis, was channeled toward rapid cell proliferation [1, 2]. In this section we examine this hypothesis, by calculating the changes in the enzymatic flux during transformation. Since enzymatic activity is ATP dependent, a reduction in flux represents a saving in ATP.

In order to estimate the changes in general enzymatic flux inside the BP cells that result from the overall decrease in gene expression (where "general flux" is an ensemble of fluxes inside the cell), we used the equation derived by Rasnick et al. [3]. First we demonstrated that this equation is correct for networks of interlocking pathways as well as for linear pathways. We then estimated the influence of the global chromosomal aberrations in BP cells on the general enzymatic flux in these cells.

Derivation of flux equation for networks: The equation of Rasnick et al. [3] calculates changes in general flux inside a cell during the transformation process. To derive the equation, Rasnick et al. used a theory that was developed by Kacser and Burns and Heinrich and Rapoport for the analysis of phenotypes generated by complex assembly lines for genes [4-7].

Every organism can be viewed as a complex network, consisting of a large array of specific and saturable catalysts organized into diverging and converging pathways, all transforming molecular species and resulting in a flow of metabolites [7]. Therefore we performed a revised analysis of enzymatic flux, which avoids the assumption of linearity of the system. Using the theory of "The control of flux" [8], we provide below a new derivation of Rasnick's equation [3]:

$$\frac{1}{Fa} = 1 - \phi + \frac{\phi}{\pi} \quad (1)$$

assessing the flux changes during transformation (see section "New derivation of the equation $\frac{1}{Fa} = 1 - \phi + \frac{\phi}{\pi}$ "). Fa is the enzymatic flux of eukaryotic cell resulting from aneuploidy. A diploid cell has enzymatic flux, F=1. The $1-\phi$ term represents the fraction of unaffected gene products. The composite term $\frac{\phi}{\pi}$ is the fraction, ϕ, of gene products undergoing a π-fold change in expression.

New derivation of the equation $\frac{1}{Fa} = 1 - \phi + \frac{\phi}{\pi}$

A change in the rate of an enzymatic reaction is equivalent to a change in the concentration of the enzyme [8]. The symbol Δ Ei refers to a small change and is applied to the concentration of any enzyme, Ei; and ΔFj is the net effect on any flux Fj, anywhere in the system.

The ratio $\frac{\Delta F_j / F_j}{\Delta E_i / E_i}$ represents the effectiveness of Ei in controlling the flux Fj [8].

$$\frac{dF_j / F_j}{dE_i / E_i} = C_i$$

C_i can be described as the flux control coefficient of the system flux with respect to enzyme concentration. According to the flux summation theorem [8] , the sum of all the flux control coefficients is equal to unity. This means that for almost all enzymes a reduction in the quantity or activity of an individual enzyme does not have an appreciable effect on the overall cellular flux [8] . Indeed, all the enzymes are "in excess", so a reduction in the activity of an enzyme hardly affects the output. Moreover, the heterozygote usually differs very little from the wild type homozygote when pool sizes and fluxes are measured [7]. Therefore the coefficients of the system enzymes are all smaller than unity and the sum of their coefficients approaches 1 [8]:

$$\sum_{1 \to n} C_i = 1$$

Thus, for a system at least as complex as a cell, each enzyme makes only a small individual contribution to the general flux.

Eqn. 1 can be rearranged to eqn. 2:

$$F = \frac{dF * Ei}{Ci * dEi} \qquad (2)$$

For n enzymes participating in a specific network we arrive at eqn.3:

$$\sum_{1\to n} F = n * F = \sum_{1\to n} \frac{dF * Ei}{Ci * dEi} \quad (3)$$

The rearrangement of eqn. 3 gives eqn. 4:

$$F = \frac{1}{n} \sum_{1\to n} \frac{dF * Ei}{Ci * dEi} \quad (4)$$

Since each enzyme makes only a small individual contribution to the general flux, all Ei terms can be approximated by replacing them with E, the mean of all the Ei:

$$F = \frac{E}{n} \sum_{1\to n} \frac{dF}{C_i * dE_i} \quad (5)$$

For each enzyme, Ei, the expression $\frac{dF}{C_i * dE_i}$ remains constant under normal conditions and can be replaced with Ai. Therefore eqn. 5 can be rewritten as follows:

$$F = \frac{(E \sum_{1\to n} A_i)}{n} \quad (6)$$

We can represent the sum of all the Ai (where Ai represents the effect of a change in the concentration of the enzyme on general flux) by a term W to give:

$$F = \frac{E * W}{n}$$ or

$$\frac{W}{F} = \frac{n}{E} \qquad (7)$$

Since the production of gene products is approximately proportional to gene dose [3] the fluxes in eqn. 7 can be partitioned (similarly to the concept of Rasnick et al. [3]) into those that are affected by aneuploidy and those that are not:

$$\frac{W}{Fa} = \frac{n-m}{E} + \frac{m}{\pi E} \qquad (8)$$

Fa is the phenotype of a eukaryotic cell resulting from aneuploidy. The number of genes experiencing a change in expression levels due to aneuploidy is m. The variable π is the segmental ploidy factor, reflecting the change in the number of gene copies or expression levels for m. The difference n-m is the number of genes not experiencing aneuploidy or a change in gene expression [3] . Eqn. (8) can be divided by eqn. (7) to arrive at eqn. (9)

$$\frac{F}{Fa} = \frac{n-m+\frac{m}{\pi}}{n} = 1 - \frac{m}{n} + \frac{m}{n\pi} \qquad (9)$$

Setting the normal phenotype F=1 and replacing the expression m/n with ϕ, we get eqn. 10 [3]:

$$\frac{1}{Fa} = 1 - \phi + \frac{\phi}{\pi} \qquad (10)$$

Calculation of altered flux in BP cells: 2199 significant genes out of the 8842 valid transcripts that passed t-test analysis were expressed at least twofold higher and 2902 genes were expressed at least twofold lower in BP cells in comparison with keratinocytes[2]. Thus, the calculated relative flux for the BP cells is:

$$\frac{1}{Fa} = 1 - \frac{2199}{8843} - \frac{2902}{8843} + \frac{2199}{8843*2} + \frac{2902}{8843*0.5} = 1.23$$
$$Fa = 0.82$$

Thus BP cells had a 18% decrease in general enzymatic flux compared to keratinocytes.

Flux change due to global chromosomal alterations: In order to estimate the contribution of global chromosomal changes to the change in flux in BP cells, we determined how many transcripts with significantly altered expression were located in chromosomal regions with coordinately altered expression. 2902 transcripts were significantly reduced in BP cells versus keratinocytes [2]. Of these, 1292 (44%) mapped to large chromosomal regions with coordinately altered expression. 2199 transcripts were significantly induced in BP cells [2], and of these 901 (41%) were located in large chromosomal regions with coordinately altered expression. The theoretical contribution of these transcripts to the change in the enzymatic flux of BP cells is :

$$\frac{1}{Fa} = 1 - \frac{1292}{8843} - \frac{901}{8843} + \frac{1292}{8843*0.5} + \frac{901}{8843*2} = 1.094$$
$$Fa = 0.91$$

Thus, a loss of flux in BP cells of 9% can be attributed to global chromosomal changes, as defined by the DIGMAP software. This minimal estimate, based on our stringent criteria for defining a chromosomal region with coordinately changed expression, indicates that at least half of the change in enzymatic flux in BP cells in comparison with keratinocytes (9% out of 18%) is due to global chromosomal alterations.

References for Supplemental data:

1. S. Mizrachy-Schwartz, N. Kravchenko-Balasha, H. Ben-Bassat, S. Klein and A. Levitzki, *PLoS One*, 2007, **2**, e628.
2. N. Kravchenko-Balasha, S. Mizrachy-Schwartz, S. Klein and A. Levitzki, *J Biol Chem*, 2009, **284**, 11717-11727.
3. D. Rasnick and P. H. Duesberg, *Biochem J*, 1999, **340 (Pt 3)**, 621-630.
4. H. Kacser and J. A. Burns, *Symp Soc Exp Biol*, 1973, **27**, 65-104.
5. R. Heinrich and T. A. Rapoport, *Acta Biol Med Ger*, 1973, **31**, 479-494.
6. R. Heinrich and T. A. Rapoport, *Eur J Biochem*, 1974, **42**, 89-95.
7. H. Kacser and J. A. Burns, *Genetics*, 1981, **97**, 639-666.
8. H. Kacser and J. A. Burns, *Biochem Soc Trans*, 1995, **23**, 341-366.

Table SI

The amplified locus	Primers
E6 gene	E6-5': 5' tgtactgcaagcaacagttac 3' E6-3': 5' gcaacaagacatacatcgac 3'
E7 gene	E7-5': 5' ggagatacacctacattg 3' E7 -3': 5' gtgtgcccattaacaggtct 3'
E1 gene	E1-5': 5' ggatgctatatcagatgacg 3' E1-3': 5' ctacctgtaacatctgctg 3'
E2 gene	E2-5': 5' gtacagacctacgtgaccat 3' E2-3': 5' catatgtaagtgtaacaattgc 3'
E5 gene (nested PCR)	outer primers: E5-5': 5' gatagtgaatggcaacg 3' and E5 -3': 5' ggtggacatgtacctgcctg 3'; inner primers: E5in-5': 5'ctattacagtgtctactgg 3', E5in -3': 5' cgcagaggctgctgttatcc 3'
L1 gene	L1-1-5': 5' acggatgaatatgttgc 3' and L1-1-3': 5' tgtaatccatagatatac 3'; L1-2-5': 5' ggtgattgtccaccattagag 3' and L1-2-3': 5' gcagttgtagaggtagatgagg 3'
L2 gene (nested PCR)	L2-5': 5' gctggtgcaccaacatc 3' and L2-3': 5' gcaacatattcatccgt3' L2in-5': 5' catctgtattgcagcctcc 3' and L2in-3': 5' gacaggaggcaagtagacag 3'
Boundary of the viral insertion within the L2 gene	From E5 gene: E5L2 5': 5' caggcaggtacatgtccacc 3'; From L2 gene: E5L2A 3': 5' ctaaatcctgatacatctg 3'; E5L2B 3': 5' cctagaggttaatgctggcctatg 3'; and E5L2C 3': 5' ggaatattgtatgcaccacc 3'
Boundary of the viral insertion within the L1 gene	From E6 gene: L1-E6 rev: 5' gtaactgttgcttgcagtaca 3';

	From L1 gene: : L1-E6A: 5’ tgcacatgggtgtgtgc 3’; L1-E6-B: 5’ gtagcgccagcggccatt 3’; and L1-E6C: 5’ ctgttgttgatactacacgc 3’
Joints between the viral and human DNA	Hinc2-out: 5’ ctgctaaacacagatgtaggac 3’ Pst1+out: 5’ ctgataccacacctgctatattag 3’ Hinc2-in: 5’ acctgaccacccgcatgaac 3’ Pst1+in: 5’ catctgtattgcagcctcca 3’ L1g-forout: 5’ ggaggaatatgatttacag 3’ L1g-revout: 5’ gatatggcagcacataatgac 3’ L1g-forin: 5’ ccttaactgcagacgttatg 3’ L1g-1revin: 5’ gcgtgtagtatcaacaacag 3’

Table SI **Primers used in mapping of the viral HPV16 insertion in HF1 cells.**

Primers used in PCR walking along the incorporated viral genome are shown.