# Supporting Information

**Yu *et al*. 10.1073/pnas.0813002106**

```
LOCUS       GXP_551436    1101 bp    DNA
DEFINITION  loc=GXL_338887|sym=AQP2|geneid=486552|acc=GXP_551436|
            taxid=9615|spec=Canis familiaris|chr=27|ctg=NC_006609|
            str=(-)|start=7759584|end=7760684|len=1101|tss=1001|
            descr=aquaporin 2 (collecting duct)|
            comm=GXT_21955060/XM_543678/501/bronze
ACCESSION   GXP_551436
BASE COUNT    281 a  253 c  377 g  190 t
ORIGIN
        1 TGGTGGCTTA CTGTGGGCCA GGTGTGGCTC TGAGCACTTT ACATCATTTC ATTTCATTCA
       61 AACCCCGATC CCTTGTAAAT TGATTATTCC CTATTTACAG AGATGGACAA GAAAGGGTAA
      121 ATCTCAGAGA CAGAGCAGGG TGATGCTGCA ATGGAAGCCC GAAGCTTCCT CAGTCAACCA
      181 CAGGTTGATC CCATCACGGG GGGCCAGGGT CCCCCCGCCA CTGAAGCCAC CCCCTCCCTC
      241 CAGGCTATGG AAGCTGCCTT TTGGCTGGGG TTCCAGGGAG GAGGTGAAGC TTAGTAACTT
      301 ATGGGTAATT GGCTGCAAGA AAGGGAGCAC AGAAGAGACC TTACGCCTTG GAAAGTTCAT
      361 GTCCCAAGCC AACCAGTTGT TTTCCATCCC TTTCAAGGGG GGATAATGGA AGGTTAACAA
      421 CAGGAAGAGC TCCCCCATCA CCAGAAGGGG TAGCATTGAC ATCCTGTCTC ATGGTCAGTG
      481 TCGAGTGAGG TGCTGTCCCT GCCTTGACCA ATACTTTGAG CTGGAGGGAG AGGGGTTTGG
      541 CATCATCCCC TTGGAGAATC CCCATTCAAG GGACCAGTGG GGGTAGGGGA GGCCCTGGGG
      601 CAGCAGCCCT GGGGGAGCCA GGGAACAAAC ACGGAAAACC AGGGGACGTC AGTCCTTATC
      661 TGGAGCTCAT TAATGGGGAA CATTAGTCAG CTGTGAAGGC CAAGATGGGG TGATAGGCAT
      721 GCGGGTGGGC AGGGGAGGGG GGGGCAGGAG GCAGCCAGAG GAGAGACAAG GTAAGGGCCC
      781 AGACTCCACC TGCGCCTCAT CGTGTCCAAA TACAGAGAGA AGCGCCCTAT AAATGCCCGT
      841 AGCCCAGCCT CCCCAGAGAC CCTGAGAGAG AGAGAGAGAG AGAGAGAGAG GGAGGGAGGT
      901 AGAGAGGGAG GGAGAGAGGG AAGAAGGGAG AGAGGGAAGA AGAGAGAGAG GGAGGGAGAG
      961 AGGGAGGGAG GGAGGGAGGG AGGGAGGGAG GGAGGGAGCA CCCACGAGCT TCCATCGCCG
     1021 GCCTGGGCCA GCCCACAGAG GGCCCCACAG CATGTGGGAA CTGCGCTCCG TAGCCTTCTC
     1081 CAGGGCGGTG TTCGCCGAGT T
//
LOCUS       GXP_593991    1101 bp    DNA
DEFINITION  loc=GXL_358532|sym=aqp2|geneid=539870|acc=GXP_593991|
            taxid=9913|spec=Bos taurus|chr=5|ctg=NC_007303|str=(-)|
            start=30623056|end=30624156|len=1101|tss=1001|
            descr=aquaporin 2|comm=GXT_21986527/XM_589978/1373/bronze
ACCESSION   GXP_593991
BASE COUNT    278 a  290 c  305 g  228 t
ORIGIN
        1 CGGGACCCTA ACATGTGAGT TTGGGGAGTC ATGGTCCCTC CTGATGGTAC AAATAATGCT
       61 ATTAATAACA GATGATATCA ACTGACAGCT TGCTGTGGGC CAGGAGCAGC TCTGAGCACT
      121 TTCCATACAT TAGTTCATTT AATCTAAACC CCAAGCCCAC AGAGGTTTAT TATCCCCTTT
      181 TGAAATCACA GAGATGCACT ATGGAATGGA AAACCATGGC TTCCTCAGTC GACCTGAGCT
      241 TGACTCCATC ACCCAAGGCC AGGATCCCAG CGCCTCTTAA ATTGCCCCCT GCTTCCAGGC
      301 TATGAAAGCT GCTATTTGGC AGGGGTTCCA GGGAGCAAGT GAAGCTTAAT GATTTATGGA
      361 TGACTGAGTG CAAGAAAGCA AACACAGGAG AGACACGAAA CCTCTATGCC TTGGAAGGTT
      421 CGTGTCACAC CAACCGGTTG TTTTCCATCC CCTGCGAGAT GGAGCTCAGT CAGCAGGAAG
      481 GATGTCCCGA TCTCCAGAAG GGGTAGCACT GACACCCTGT TTCATGGTCA GTTCTGTAAG
      541 GTGTTGCCCC TGACTTGGCA AACAGTTTGA ATCAGAGAGA GGGGTTTGGC AGCATAAGCC
      601 GTACAGAATC CCCATAGGAC CGGTGGGGAC AGGGGAGGCC CTGGGGCAGC TGCCCTAGGA
      661 TAACCTAGGG AACAAACACG GAAAACCAGG GGACGTCAAT CCCTATCTGG AGCTCATTAA
      721 TGGGGAACTC CATTCAGCTG TGAAGTCCTA GATGGGGTGA TAGGCATGCT GCGCCCAGGG
      781 GTAGGGGGTA GGGGCAGGAG GCGGCCATAG GGGAGAAAGG GTATCGGCCA GACTCCACCT
      841 ACATCTCACG TGTCCAAATT CAGGATGGAG GGGCCCTATA AATGCCCAGG GCCAGCCTCC
      901 CCGGAGACCC TGAGAGAGAG ACAGACAGAC AGAGGGAGAG CAAGAGCTGC AAGCTCCCCA
      961 CCCCGAGACC CCTGGGCCAG CCCACGGAAG GCTCCGCAGC ATGTGGGAAC TCCGGTCCAT
     1021 AGCCTTCTCC AGGGCAGTGC TTGCAGAGTT CCTGGCCACA CTCCTCTTTG TCTTCTTTGG
     1081 CCTCGGCTCA GCCCTCAACT G
//
```

**Fig. S1.** Sequences of AQP2 5′-flanking regions from 5 species.

```
LOCUS       GXP_89332    1117 bp    DNA
DEFINITION  loc=GXL_74454|sym=Aqp2|geneid=11827|acc=GXP_89332|
            taxid=10090|spec=Mus musculus|chr=15|ctg=NC_000081|str=(+)|
            start=99406090|end=99407206|len=1117|tss=1001,1017|
            descr=aquaporin 2|
            comm=GXT_12978455/AK002689/1251/gold;
            GXT_13074328/NM_009699/1267/bronze;
            GXT_21805231/AK142567/1251/gold
ACCESSION   GXP_89332
BASE COUNT    286 a  278 c  320 g  233 t
ORIGIN
        1 CACCAAGACT AGAGGTTACT AGCAGGCACC ATGGCCTCCT TGAGCCTGGA GAGCTAGCTG
       61 CAACTCCAGA AACATCTTCC TCCTGTGGTA AGGGTGGCTC TGAGCTCGAG ATGGTATCTA
      121 GAGCCTGTTG TAGCCTTGGA GACTCTGAGT TTGGGGGAGA CCTGTTTGCC TCAAAATGCA
      181 AATAGCACTG AGCAATAGCT GACAAAACCT GAGTTCTTCC TGAGAACCAG GAGCCATATT
      241 AATGCAAATC TCTGTCCTAT GAGATAGTTT TATTATCTCA TGTACACAGG CAGAGCAGAA
      301 CTGTCACATC TCAACGACAT GAGTGTGTGC TGCAGAACAG AAACCCTCAT GTTCCTAAGT
      361 TGACCTTGGT GACTCTATCA CGCAGGGCTA GGGTCCCAGG GCCACTGAAG CCACCTTTGG
      421 CTCCAGGCAA TGGGAGTTCC TCTCTGGCTG AGATTCCTGG GTGAAGGTGG AACATAATGA
      481 CTTATGGGTG ATTAACTGCA AGAAAGCAAG CCCAAGACCT TTTGCCTTGG AAATTTGTCC
      541 CAGGCCAGCC AGTTGTTTTC CATCCTCTTC CAGAAGAAAG ACCATCCAGT CCGCGGAAAG
      601 GCAGCGCCCA CATTTCCTCA CAGTTTGGGG TAAGGCATTG CCCCTTCCTT GGCCAGCCTT
      661 TTAGTCAAAG AGAAGGTCAC TGGACACAGC CTCCTCTGCA GGAACTGGTG CTGGTGGTCA
      721 TTGTGGGGGC TGGGGCAGCC CTGAGGCAGC TCCATGGGGT AACTGAGGAA AAAACGAGGA
      781 AAACAGAGAC GTCAATCCTT ATCTGGAGTC CATTAATGAG GAGAACATTA GTCAGCTGTG
      841 AAAGCTAAGA TGGGGTGATA GGCCTTCGGG TGGGCAGGAG CAGGGATGGG GGCAAGTCCG
      901 CCATGGGGGA GCACAGGGTT GGCAGGAACG CCTCCTCCAC CCCACGTGCC CAGGTCTACG
      961 ATAGGAAGGC CCTATAAGTG CCCACAGTCT AGCCTCTCCG GAGGCCCAGA GGAAGAGAGA
     1021 AGAGAAAGAG AGAGGGAGGG AGGAAGAGCC ACCCCCGTGG CCCAGACCAC TGGCCAGCGC
     1081 TCAGAAGGCC GAGCAGCATG TGGGAACTCC GGTCCAT
//
LOCUS       GXP_241970    1101 bp    DNA
DEFINITION  loc=GXL_201711|sym=Aqp2|geneid=25386|acc=GXP_241970|
            taxid=10116|spec=Rattus norvegicus|chr=7|ctg=NC_005106|
            str=(+)|start=138324855|end=138325955|len=1101|tss=1001|
            descr=aquaporin 2|comm=GXT_21770862/NM_012909/1261/bronze
ACCESSION   GXP_241970
BASE COUNT    284 a  269 c  323 g  225 t
ORIGIN
        1 AGGTAGGTCA GTCTTCTGAG GCTCCACAGC CACAGGAGCT GTGAGGGCAG GTGACCTTGT
       61 AGAACACTGA CCAAGCTGAG TGGTCCCGCC ACCAGAAACA TCTTCCTCCT CTGAGCTCAA
      121 AGATGGTACC TAGAGTCTGT TGTAGCCTCA GAGACTCTGA GTTTGGGGGA CAGGACCTAT
      181 TTACCTCAAA ATGCAAATAG CACTGAGCAA TGACTGACAT AGTAACGCAG ATCTCTATCC
      241 TACGAGATAG TTTTATTACC TCATCTATAG AGACAGAGCG GGGATGTCAC ATCTGGAGTG
      301 CGTGCTGCAG AACAGAAGCC CTTGTGTTTC TTGTTGGCTC TGGATGACTC TACCACACAG
      361 GGCTAGGGTC CCCGGGGGCC ACTGAAGCCA CCCTTGGCTC CGGGCAATGG GAGTTACTCT
      421 CTGGCTGGGG TTCCTGGGTG AAGGTGAAAC ATAATAACTT CTGGGTGATT AACTGCAAGA
      481 AAGCAAGCAC AAGACCTTTG CCTTGGAAAT CTGTCCCAGG CCAATCAGTT GTTTTCCATC
      541 CTCTTCCAGA AGAAAGACCG ACCATCCAGT CAGCGGAAAG GTAGTTCCCA CATTTTCTCA
      601 CAGTTTGGGC TGAGGTGGTG CCCCTTCCTT GGCCAGCCTT TTAGTCGAAG AGAAGGTCAC
      661 TGGACACGGT CTCCAATGCA GGAACTGGAA CTGGTGGTCA CTCTGGGGGC TGGGGAAGCC
      721 CTGGGGCAGC TCCCTGGAGT AACTGAGGAG AAAAAAAATG AGGAAAACAG AGGACGTCAA
      781 TCCTTATCTG GAGTTCATTA ACGAGGAGAA CATTTAGTCA GCTGTGGAAG CTAAGATGGG
      841 ACGATAGGGC CTTCAGGTGG GCAGGAGCAT GGGTTGGGGC AAGGCCACCA CGGGAGAGCA
      901 TGGTGTGGGC AGGAACGCCT CCTCCATACC ACGTGCCCAA GCCTACGATA GGAAGGTCCT
      961 ATAAGTGCCC ACGGTCTAGC CTCCCCAGAG GCCCAGAGGA AGAGAGAAGA GAAAGAGAGA
     1021 GGGAGGGAGG AAGAGCCACC CCCGTGGCCC AGACCCCTGG CCAGCGCGCA GAAGTCGGAG
```

**Fig. S1.** (continued)

```
     1081 CAGCATGTGG GAACTCAGAT C
//
LOCUS       GXP_71921    1106 bp    DNA
DEFINITION  loc=GXL_59683|sym=AQP2|geneid=359|acc=GXP_71921|taxid=9606|
            spec=Homo sapiens|chr=12|ctg=NC_000012|str=(+)|
            start=48629791|end=48630896|len=1106|tss=1001,1006|
            descr=aquaporin 2 (collecting duct)|
            comm=GXT_2814311/NM_000486/1281/gold;
            GXT_22215291/AK225940/1276/gold
ACCESSION   GXP_71921
BASE COUNT    279 a  290 c  308 g  229 t
ORIGIN
        1 CAGAAATGTC TTTCTCCTGT GGCCAGAGCG GCCCTGAGCT CCTCAGGGTT TCTAGGACCT
       61 GTCACAGCTT CTGGGACCCC AACAAATGAG TGTGCAAAGA CCTGTTCTCC TCAAAGTGCA
      121 AATAATACTG TAAGTAAGGG CTGATATTCA TTGACCACCT ACTGTGGACG AGGCGCTGCT
      181 CTGAGCACTT GACATACATT GTTTCATTCA ATCCAAACCC CAATCCACTA GGTCAGTTTT
      241 ATTCTCCCCA TTTTACAGAG AAAGACAAGA AAGGTTAAGT CACATAGATG GAACAGGCTG
      301 ATGCCAGAAT GGACACCCTC AGCTTCCTCA GCCGCCCCAG GTTGACTCCA TCACCAAGGG
      361 CCAGGGTCCC AGTGCCCCCA GGCTCTAGGA GCTGCTGTTT GGCTGGGGCT CCCAGGAGAA
      421 GGTGAAGCTT AATGATTTAT GGGTGATTAG CTGCAAGAAT GCAAGCACAG AAGACACAAA
      481 CCTTTATGCC TTGGAAATTT GTCACAAGCC AACCAGTTGT TTTCCATCCC TGTGAAGCAG
      541 GAATAATTGG AGGTCTGACA GAAGAACTTC CCAGTTACCA GAAGGAGCAG CACTCATGTT
      601 CTCTCTTAGT TTTGTGTGAG GTGTTGCCCC TGCCTTGGTC AACAGTTTGA GTCAGAGAGA
      661 TGGGGGCCGG GCACAATCCC CACCAGACGT TCCCATTCAC AAGACCCTGT GGGGGCTGGG
      721 AGACGCCCTG GGGCAGCAGC CCTGGGGTAA CCAAGGGAAC AAACACGGAA AACCAGGGAC
      781 GTCAGTCCTT ATCTGGAGCT CATTAAATGG GGAACATTAG TCAGCTGTGA AGGCCAAGAT
      841 AGGGTGATAG GCCTGTGGGT GGGCTGGGAT GGGGCATGGG GGCAGAGGCC GCCATGGAGG
      901 AGAAGAGGTA TTGGCCTCAA CGACTCCACC TCCCCGCCAC GTGCCCAGAT CCGGGATGGA
      961 AGGACCCTAT AAATGCCCAC AACCCAGCCT CCCCAGAGGC CTTGAGAAAG AGAGCGATAG
     1021 AGTGCGAGAG CGAGTGCCCG GAGCATCCTG GCCCTGAGAC AGCTGGGCCA GCCCCGCAGG
     1081 GCTCTGCAGC ATGTGGGAGC TCCGCT
//
```

**Fig. S1.** (continued)

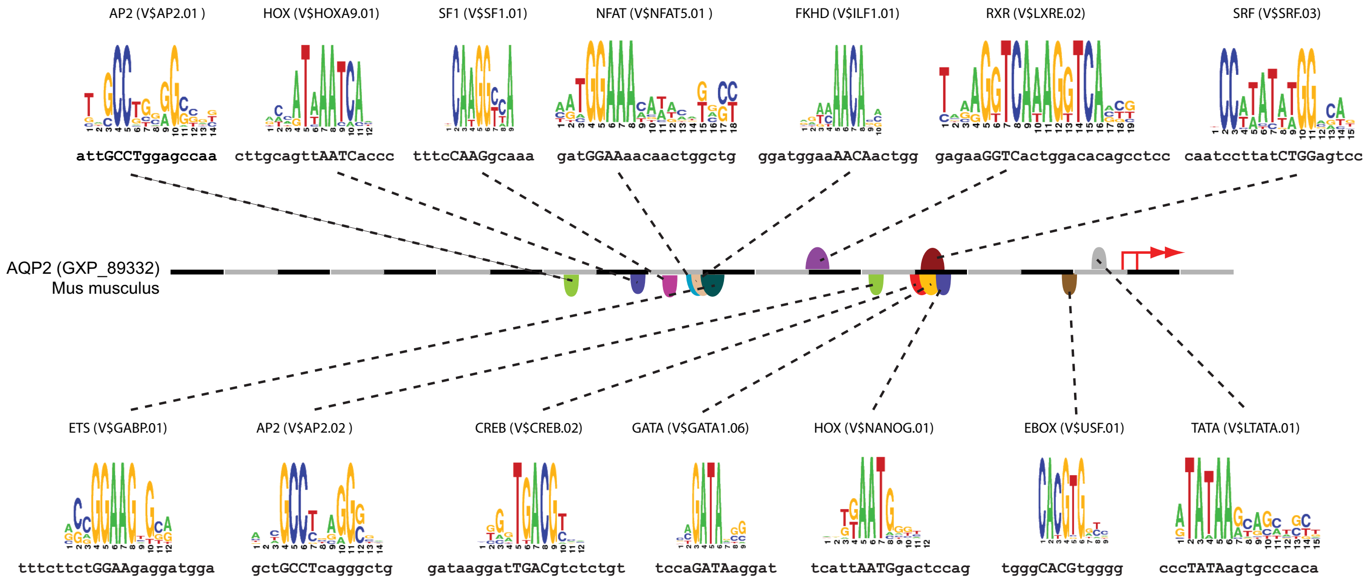

**Fig. S2.** Position-weighted transcription regulator binding elements (TRBEs) in the mouse AQP2 5′-flanking region. Transcription starting site is indicated by red arrows. TRBE family names and specific position-weighted binding element are indicated. Mouse sequences are shown.

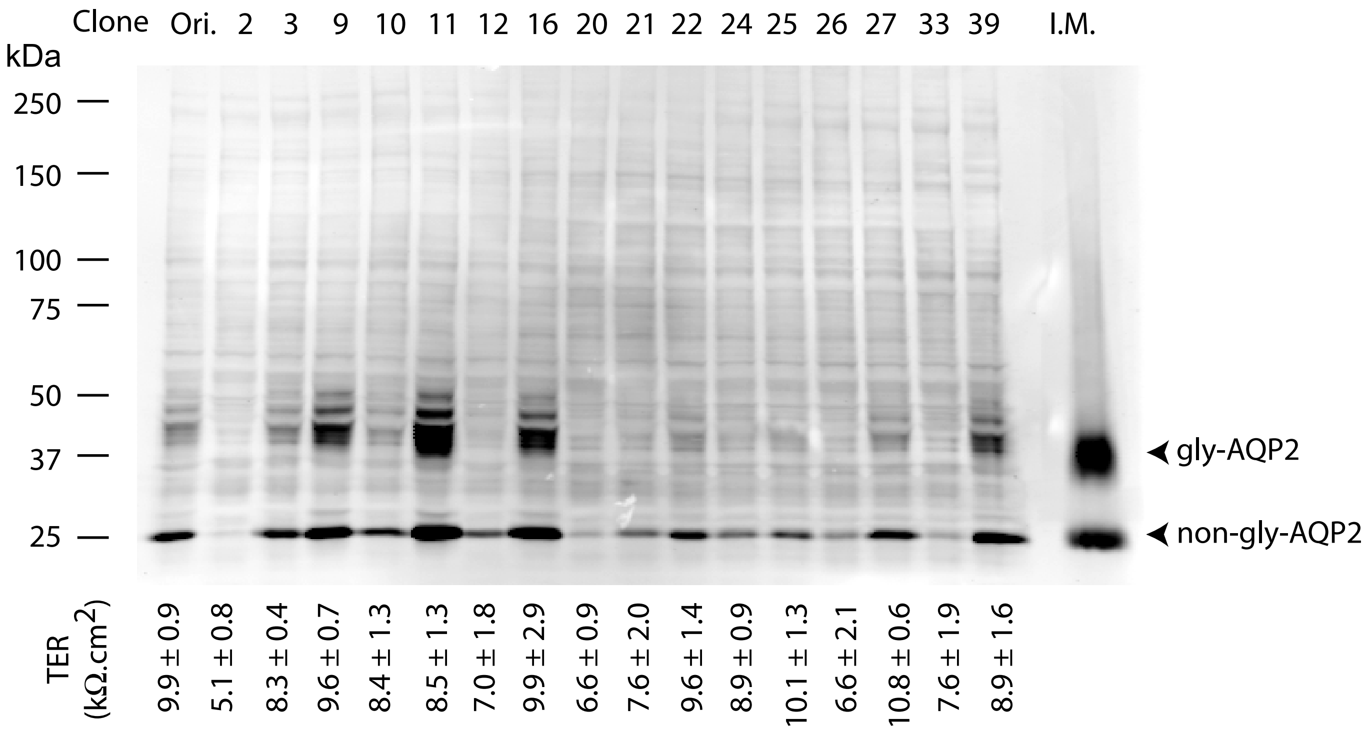

**Fig. S3.** AQP2 immunoblot illustrating *mpkCCD* cells with various AQP2 expression levels. The *mpkCCD*$_{c14}$ cell clones (numbered plus the original cell [Ori.]) were grown on membrane support until confluence (transepithelial resistance [TER] $> 5$ kΩ.cm$^2$). Cells were then exposed to 0.1 nM vasopressin analogy dDAVP added to the basolateral serum-free medium for 5 days. 10 $\mu$g of protein was tested for AQP2 expression using immunoblotting (N20 antibody from Santa Cruz Biotechnology). The average TER during the dDAVP exposure was shown as mean ± SE. Gly-AQP2 and non-gly-AQP2 indicate glycosylated and non-glycosylated AQP2 protein bands. I.M. is 10 $\mu$g protein sample from rat kidney inner medulla.

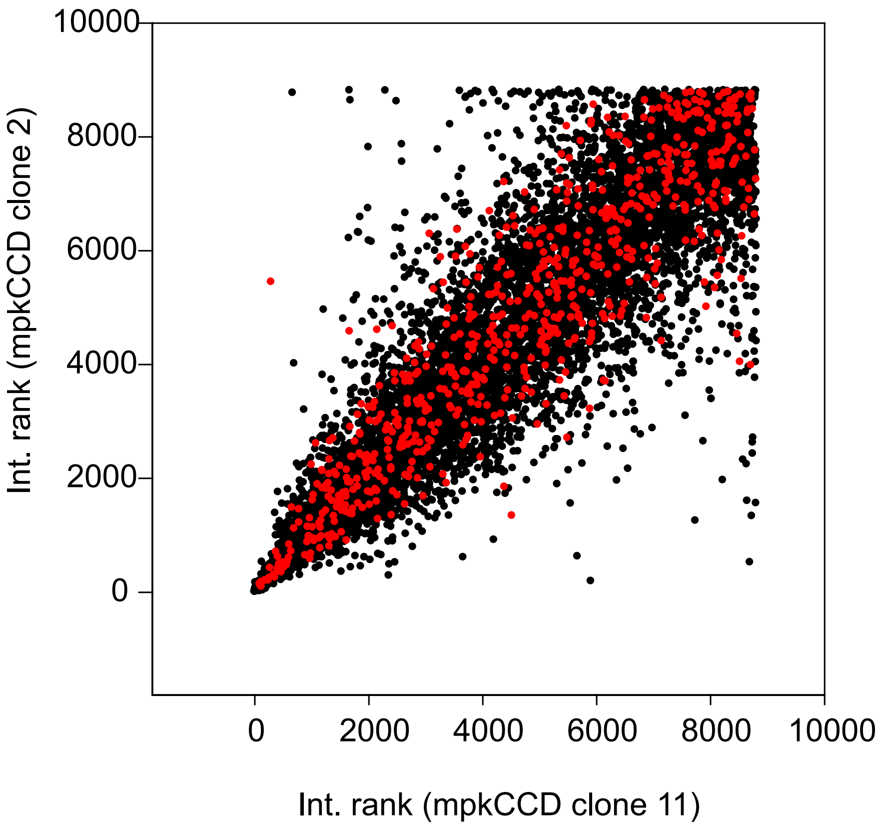

**Fig. S4.** A Scattergram comparing transcripts of *mpkCCD* clone 11 cells vs. clone 2 cells. Comparison was based on the rank of transcript intensity. Red dots indicate transcription factors and cofactors.

## Other Supporting Information Files

Table S1 (PDF)

Table S2 (PDF)

Table S3 (PDF)

Table S4 (PDF)

Table S5 (PDF)

Table S6 (XLS)