**Title of Senior Honors Thesis**

**Tanscriptome comparison of ionizing radiation resistant strain of *Escherichia coli*, CB2000, and naturally occurring strain of *Escherichia coli*, Founder**


**Abstract**

CB2000 is a strain of *Escherichia coli (E. coli)* that is extremely resistant to ionizing radiation. It was derived from a colony, referred to as the founder strain, isolated from naturally occurring Escherichia coli K12 MG1655. Using microarray technology developed by Affymetrix, a transcriptome comparison was done on CB2000 and Founder to determine the gene(s) that are differentially expressed in CB2000. A large number of genes which are involved in metabolism, transport and stress response were identified to display changes in expression levels in CB2000 when compared to Founder.



Meng Kwang Marcus Tan /Biochemistry major.
Author Name/Major

Author Signature

05/10/07
Date

Michael M. Cox, Professor, Biochemistry Department
Mentor Name/Department

Mentor Signature

**COVER SHEET**

TITLE: **Tanscriptome comparison of ionizing radiation resistant strain of *Escherichia coli*, CB2000, and naturally occurring strain of *Escherichia coli*, Founder**

AUTHOR'S NAME: **Meng Kwang Marcus Tan**

MAJOR: **Biochemistry**

DEPARTMENT: **Biochemistry**

MENTOR: **Michael M. Cox**

DEPARTMENT: **Biochemistry**

YEAR: **2007**

(The following statement must be included if you want your paper included in the library's electronic repository.)

# Tanscriptome comparison of ionizing radiation resistant strain of *Escherichia coli*, CB2000, and naturally occurring strain of *Escherichia coli*, Founder

## Abstract

CB2000 is a strain of *Escherichia coli (E. coli)* that is extremely resistant to ionizing radiation. It was derived from a colony, referred to as the founder strain, isolated from naturally occurring Escherichia coli K12 MG1655. Using microarray technology developed by Affymetrix, a transcriptome comparison was done on CB2000 and Founder to determine the gene(s) that are differentially expressed in CB2000. A large number of genes which are involved in metabolism, transport and stress response were identified to display changes in expression levels in CB2000 when compared to Founder.

## Introduction

Previous studies have shown that exposure to ionizing radiation has many adverse effects on many cells, including death in *E.coli* at sufficiently high doses (1, 2). Radiation sensitivity has often been associated with DNA damage. DNA damage usually triggers DNA damage repair and the SOS response in *E. coli*, which involves proteins such as RecA and lex proteins (3). However, many studies have indicated that radiation sensitivity in prokaryotes depends on many cellular mechanisms other than DNA repair. Heat shock responses (4) and protein oxidation (5) have been implicated in conferring radiation resistance to prokaryotes. Hence, it can be seen that though DNA repair is important in preventing cell death when cells are exposed to radiation, other cellular responses also play important roles in the survival of these cells.

CB2000 was obtained by subjecting a culture obtained from a single colony, referred to as the Founder, isolated from naturally occurring *E. coli* K12 MG1655 to twenty cycles of exposure to ionizing radiation and growth. The length of each exposure was adjusted to kill greater than 99% of the population. The dose administered increased from 2040Gy for the first cycle to 10000Gy for the last cycle. Cells which survived after each exposure cells were harvested, and allowed to recover in rich media until reaching stationary phase. This population was then irradiated again. After 21 cycles of irradiation and growth, single colony isolates were tested for ionizing radiation and from those isolates which displayed between 1000- to 5000-fold increase in survival, one colony, CB2000 was selected for further analysis. This culture was grown for at least 100 generations without selection and expressed no considerable change in ionizing radiation resistance. CB2000 has about a 4500-fold increase in survival relative to the Founder after exposure to 3,000Gy.

With this knowledge, we set out to elucidate the genes that are differentially expressed in CB2000, as part of an effort to determine which genes are critical to radiation resistance. cDNA obtained from reverse transcription of RNA isolated from CB2000 and Founder were compared to identify changes in gene expressions. Three sets of comparisons between CB2000 and Founder were performed to ensure the accuracy and consistency of the results of this study. Microarray technology developed by Affymetrix was employed in our study. Changes in expression (increased or decreased) were analyzed. Changes in gene expression levels which were significant (having a p-value >1.0 for increased gene expression and <-1.0 for decreased gene expressions) were

collated and used for analysis. The p-value used here is specific to Affymetrix's algorithm and is different from the p-value used in statistical hypothesis testing.

Genes with significant changes in expression levels in CB2000 compared to Founder were first obtained for each of the 3 sets of comparisons. This was followed by the identifying genes with significant changes in expression levels which were common in all three comparisons.

## Materials and Methods

*Organisms and culture methods*

Radiation resistant strain of *Escherichia coli*, CB2000 and naturally occurring Escherichia coli isolate, Founder, were obtained from Elizabeth Wood, University of Wisconsin- Madison. These *E. coli* strains were cultured in Luria broth (LB) and incubated overnight at 37 ºC.

*Total RNA extraction*

Total RNA from both overnight cultures were extracted using the MasterPure™ RNA purification kit developed by Epicentre Biotechnologies. Concentrations of the RNA obtained from both strains were determined using the NanoDrop® ND-1000 UV-Vis Spectrophotometer.

*Obtaining cDNA from RNA*

Total RNA extracted from both strains was used to create complementary DNA (cDNA) via reverse transcription. This was done according to the *E. coli total RNA labeling protocol for Affymetrix GeneChips.* The $pd(N)_6$ sodium salt (hexamer) was

obtained from Amersham. A 12 uL mixture containing 10 ug of total RNA, 0.5 ug/uL hexamer and Diethylpyrocarbonate (DEPC) water (RNase-free) was prepared for each strain. This mixture was incubated at 70 ºC for 5 minutes. The RNA/hexamer mix for each strain was chilled on ice for 2 minutes. This 12 uL RNA/hexamer was incubated at 42 ºC in a 60 uL solution containing 10mM dNTP mix, 5x First Strand Buffer, 0.1M DTT, DEPC water and 200U/ul SuperScriptII. The dNTP mix was obtained from Amersham; the 5x First Strand Buffer and 200U/ul SuperScriptII from Invitrogen. Note: Incubations at 70 ºC and 42 ºC were carried out in a thermal cycler.

*Removing RNA remnants*

The total cDNA from each strain was incubated at 37 ºC for 10 minutes with 2U/uL RNase H and 5ug/uL RNase A.

*Purifying cDNA*

The total cDNA obtained for each strain was purified using Qiagen's Minelute PCR purification kit. The concentrations of the total cDNA obtained from each strain were determined using the NanoDrop® ND-1000 UV-Vis Spectrophotometer.

*Digestion of cDNA*

DNase I was used to fragment the obtained cDNA. 10uL of purified cDNA was added to make a 40uL solution containing 10x DNase I buffer, water and 0.1U/uL DNase I. This mixture was incubated at 37 ºC for 8 _ minutes and followed by 99 ºC for 10 minutes to inactivate the DNase I. Fragmentation of cDNA was checked on a 2%

agarose 1x TAE gel using 2 uL of each sample. cDNA fragments must have sizes between 50 and 100 base pairs. The DNase I and buffer were obtained from Epicentre Biotechnologies.

*cDNA end-terminus labeling*

38 uL of the fragmented cDNA from each strain was added to a final volume of 100uL solution containing 10x Terminal deoxynucleotide transferase (TdT) buffer, 2.5 mM cobalt chloride, 1 mM Biotin-$N^6$-ddATP, water and 20U/uL TdT. The mixture was incubated in a thermal cycler for 2 hours at 37 ºC. The Biotin-$N^6$-ddATP was obtained from Enzo, the TdT from New England Biolabs and RNase H from Invitrogen.

*Microarray analysis*

The *E. coli* Antisense Genome Arrays were purchased from Affymetrix.

## Results

### Table 1 : Genes with increased expression in CB2000 compared to Founder

| Gene locus | Gene | Gene product | Function |
|---|---|---|---|
| b0411 | tsx | nucleoside channel, receptor of phage T6 and colicin K | ion transport<br>detection of virus<br>entry of virus into host cell |
| b1288 | fabI | enoyl-(acyl carrier protein) reductase | fatty acid biosynthesis<br>metabolism<br>lipid biosynthesis<br>response to antibiotic |
| b1356 | racR | Rac prophage; predicted DNA-binding transcriptional regulator | regulation of transcription, DNA-dependent |
| b1513 | lsrA | fused AI2 transporter subunits of ABC superfamily: ATP-binding components | transport |
| b1518 | lsrG | autoinducer-2 (AI-2) modifying protein LsrG | antibiotic biosynthesis |
| b1975 | serU | tRNA | |
| b2143 | cdd | cytidine/deoxycytidine deaminase | cytidine metabolism |
| b2181 | yejG | Hypothetical protein yejG | |
| b2393 | nupC | Nucleoside permease nupC<br>nucleoside (except guanosine) transporter | transport |
| b2504 | yfgG | hypothetical protein | |
| b2598 | pheL | pheA gene leader peptide | |
| b2872 | ygeY | hypothetical protein | proteolysis |
| b2964 | nupG | nucleoside transporter | nucleoside transport |
| b2980 | glcC | Glc operon transcriptional activator<br>DNA-binding transcriptional dual regulator, glycolate-binding | regulation of transcription, DNA-dependent |
| b3748 | rbsD | High affinity ribose transport protein rbsD<br>predicted cytoplasmic sugar-binding protein | carbohydrate transport |
| b3749 | rbsA | Ribose transport ATP-binding protein rbsA<br>fused D-ribose transporter subunits of ABC superfamily: ATP-binding components | transport |
| b3750 | rbsC | ribose high-affinity ABC transporter permease component<br>ribose ABC transporter permease protein | transport |
| b3751 | rbsB | D-ribose-binding periplasmic protein precursor<br>D-ribose transporter subunit | chemotaxis |
| b3752 | rbsK | Ribokinase | carbohydrate metabolism<br>D-ribose metabolism |
| b3753 | rbsR | Ribose operon repressor<br>DNA-binding transcriptional repressor of ribose metabolism | regulation of transcription, DNA-dependent |
| b3754 | hsrA | predicted multidrug or homocysteine efflux system | dicarboxylic acid transport<br>tetracycline transport |
| b3761 | trpT | tRNA | |
| b3831 | udp | Uridine phosphorylase | nucleoside,nucleotide catabolism |
| b4060 | yicB | predicted inner membrane protein | |
| b4062 | soxS | Regulatory protein soxS<br>DNA-binding transcriptional dual regulator | regulation of transcription, DNA-dependent<br>regulation of transcription |
| b4315 | fimI | Fimbrin-like protein fimI precursor, fimbrial protein involved in type 1 pilus biosynthesis | cell adhesion |
| b4381 | deoC | deoxyribose-phosphate aldolase | deoxyribonucleotide catabolism |
| b4382 | deoA | thymidine phosphorylase | pyrimidine base metabolism |
| b4383 | deoB | phosphopentomutase | nucleotide catabolism<br>deoxyribonucleotide catabolism<br>metabolic compound salvage |
| b4384 | deoD | purine nucleoside phosphorylase | nucleobase, nucleoside, nucleotide and nucleic acid metabolism, purine |

**Table 2 : Genes with decreased expression in CB2000 compared to Founder**

| Gene locus | Gene | Gene product | Function |
|---|---|---|---|
| b0220 | ivy | inhibitor of vertebrate C-lysozyme | |
| b0240 | crl | DNA-binding transcriptional regulator | regulation of transcription, DNA-dependent<br>positive regulation of transcription |
| b0329 | yahO | hypothetical protein | |
| b0382 | yaiB | Hypothetical protein yaiB | |
| b0384 | psiF | hypothetical protein | |
| b0442 | ybaV | Hypothetical protein ybaV precursor | |
| b0485 | ybaS | predicted glutaminase | glutamine metabolism |
| b0598 | cstA | Carbon starvation protein A | cellular response to starvation |
| b0643 | ybeL | Hypothetical protein ybeL | |
| b0753 | ybgS | hypothetical protein | DNA binding |
| b0812 | dps | DNA protection during starvation conditions | iron ion homeostasis<br>response to stress<br>mitotic chromosome condensation |
| b0863 | artI | arginine transporter subunit | amino acid transport |
| b0880 | cspD | Cold shock-like protein cspD | regulation of transcription, DNA-dependent<br>response to stress<br>negative regulation of DNA replication |
| b0897 | ycaC | Protein ycaC<br>predicted hydrolase | metabolism |
| b1003 | yccJ | Hypothetical protein yccJ | |
| b1004 | wrbA | TrpR binding protein WrbA | negative regulation of transcription |
| b1137 | ymfD | e14 prophage; predicted SAM-dependent methyltransferase | |
| b1138 | ymfE | e14 prophage; predicted inner membrane protein | |
| b1143 | ymfI | e14 prophage; predicted protein | |
| b1145 | ymfK | e14 prophage; repressor protein phage e14 | transcription<br>regulation of transcription, DNA-dependent<br>proteolysis |
| b1188 | ycgB | Hypothetical protein ycgB | |
| b1304 | pspA | regulatory protein for phage-shock-protein operon | |
| b1305 | pspB | phage shock protein B | |
| b1306 | pspC | Phage shock protein C<br>DNA-binding transcriptional activator | regulation of transcription, DNA-dependent |
| b1307 | pspD | peripheral inner membrane phage-shock protein | |
| b1376 | uspF | stress-induced protein, ATP-binding protein | response to stress |
| b1417 | gapC | Glyceraldehyde 3-phosphate dehydrogenase A pseudo | glucose metabolism<br>glycolysis |
| b1443 | ydcV | predicted spermidine/putrescine transporter subunit | transport |
| b1492 | gadC | Amino acid antiporter<br>predicted glutamate:gamma-aminobutyric acid antiporter | amino acid transport<br>pathogenesis |
| b1493 | gadB | Glutamate decarboxylase alpha<br>glutamate decarboxylase A, PLP-dependent | glutamate metabolism<br>carboxylic acid metabolism |
| b1732 | katE | hydroperoxidase HPII(III) (catalase) | electron transport<br>response to oxidative stress<br>hydrogen peroxide catabolism |
| b1783 | yeaG | Hypothetical protein yeaG, conserved protein with nucleoside triphosphate hydrolase domain | |

**Table 2 (continued)**

| | | | |
|---|---|---|---|
| b1784 | yeaH | hypothetical protein | |
| b1795 | yeaQ | Hypothetical protein yeaQ<br>conserved inner membrane protein | |
| b1967 | hchA | chaperone protein HchA | protein folding |
| b2080 | yegP | hypothetical protein | |
| b2097 | fbaB | fructose-bisphosphate aldolase | glycolysis<br>metabolism |
| b2266 | elaB | ElaB protein<br>hypothetical protein | protein binding |
| b2414 | cysK | cysteine synthase A, O-acetylserine sulfhydrolase A subunit | cysteine biosynthesis from serine<br>metabolism<br>amino acid biosynthesis<br>cysteine biosynthesis |
| b2750 | cysC | adenylylsulfate kinase | sulfate assimilation<br>amino acid biosynthesis<br>cysteine biosynthesis |
| b2978 | glcF | pseudo<br>glycolate oxidase iron-sulfur subunit | electron transport |
| b3098 | yqjD | Hypothetical protein yqjD | |
| b3100 | yqjK | hypothetical protein | |
| b3207 | yrbL | Hypothetical protein yrbL | |
| b3506 | slp | Outer membrane protein slp precursor<br>outer membrane lipoprotein | |
| b3509 | hdeB | Protein hdeB precursor<br>acid-resistance protein | |
| b3510 | hdeA | Protein hdeA precursor<br>stress response protein acid-resistance protein | |
| b3511 | hdeD | acid-resistance membrane protein | |
| b3512 | yhiE<br>gadE | Hypothetical protein yhiE<br>DNA-binding transcriptional activator | glycolysis<br>transcription<br>regulation of transcription, DNA-dependent |
| b3517 | gadA | Glutamate decarboxylase alpha<br>glutamate decarboxylase A, PLP-dependent | glutamate metabolism<br>carboxylic acid metabolism |
| b3555 | c4375<br>yiaG | Hypothetical protein yiaG<br>predicted transcriptional regulator | regulation of transcription, DNA-dependent |
| b3604 | lldR | DNA-binding transcriptional repressor | regulation of transcription, DNA-dependent |
| b3708 | tnaA | Tryptophanase<br>tryptophanase/L-cysteine desulfhydrase, PLP-dependent | tryptophan metabolism<br>tryptophan catabolism<br>aromatic amino acid family metabolism |
| b3709 | tnaB | tryptophan transporter of low affinity | tryptophan catabolism<br>transport<br>amino acid transport |
| b3917 | sbp | Sulfate-binding protein precursor<br>sulfate transporter subunit | sulfate transport |
| b3934 | cytR | Transcriptional repressor cytR<br>DNA-binding transcriptional dual regulator | transcription<br>regulation of transcription, DNA-dependent |
| b4045 | yjbJ | predicted stress response protein | |
| b4050 | yjbO | phage shock protein G | |
| b4139 | aspA | Aspartate ammonia-lyase | aspartate metabolism |
| b4244 | pyrI | aspartate carbamoyltransferase regulatory subunit | pyrimidine nucleotide biosynthesis<br>amino acid metabolism |
| b4245 | pyrB | aspartate carbamoyltransferase catalytic subunit | pyrimidine nucleotide biosynthesis<br>amino acid metabolism |
| b4369 | leuP | tRNA | |

## Discussion

CB2000 possesses an extraordinary ability to withstand the lethal effects of ionizing radiation. Using microarray technology, genes which may play a part in radiation resistance in *E. coli* were identified.

Contrary to previous reports which indicated that DNA repair proteins, such as RecA (6), are required for prokaryotic cell survival when DNA damage had occurred, the results obtained from our study did not show changes in expression levels in genes responsible for DNA repair such as RecA, LexA, helicases, etc. Since only changes in gene expression which were significant were selected, the genes responsible for DNA repair may have changed their expression levels, but were not significant, so they have been discounted in our analysis. Another reason for this is that DNA repair may play an integral role in cell survival under radiation, but it alone cannot guarantee the survival of the cell.

The majority of the genes with increased gene expressions in Table 1 are involved in nucleotide metabolism. One of the damaging effects of ionizing radiation is degrading DNA. Cells overcome this problem by activating DNA repair mechanisms (7, 8, 9). All of these mechanisms involve adding/replacing nucleotide bases in the damaged DNA. Thus, it is apparent that having increased levels of nucleotide metabolism may hasten and promote DNA repair. Fragmented DNA can be removed by nucleotide catabolism and the nucleotides obtained are recycled and become available for repair. Nucleotide anabolism creates more nucleotides needed to the DNA repair.

Upon close observation of the genes in Table 2, two groups of genes stand out among the genes identified to have decreased levels of expression. They are genes

responsible for the metabolism of various essential compounds such as glucose, proteins and lipids, and genes responsible for stress responses.

The decrease in expression levels of genes responsible for metabolism may be attributed to the evolution of more efficient metabolic mechanisms in CB2000 compared to the Founder. When cells are made to survive in harsh environments, they have to make use of whatever carbon, nitrogen, sulphur, mineral and energy sources available to sustain themselves. Therefore, CB2000, having to survive under conditions with high doses of ionizing radiation, may have developed metabolic pathways which are more efficient than its naturally occurring counterpart, the Founder strain.

Genes responsible for stress response also had decreased expression levels in CB2000 compared to Founder. Cellular stress responses help cells to counter and adapt to environmental stresses such as heat and shock. These responses are essential for cells to survive when changes in environmental conditions change. The highly radiation resistant CB2000 has adapted to harsh living conditions such as exposure to high doses of ionizing radiation, by developing a higher tolerance or threshold for environmental extremities. Under normal and ambient living conditions, stress response genes expression levels in CB2000 are lower than wild type levels because the increased threshold for harsh conditions in CB2000 has not been reached to trigger the expression of these genes.

Another striking observation is that radiation resistance is not controlled by a single gene or a few genes. It can be attributed to the coordinated, universal cellular responses which protect the cells and allow these cells to survive and thrive under harsh

conditions. Changing a single or a few phenotypes is insufficient in the preservation of cells living in harsh conditions.

More work will be needed to determine the extent to which certain groups of proteins are important in radiation resistance. Some of the genes identified may not even play any role in radiation resistance, while others may be essential for conferring radiation resistance. A way to investigate this is by using gene knock-outs or to prevent the translation of the mRNA of the genes of interest using siRNA. Strains of CB2000 with certain genes removed from their genome can be created. The viability and survivability of these knock-out strains under similarly high doses of ionizing radiation can be determined and compared to the original CB2000 strain. The importance of any given gene or groups of genes is directly proportional to how well the cells which have this gene or genes removed survive compared to the original CB2000 strain.

Another method of identifying important radiation resistance conferring genes is by comparing the gene expression levels of CB2000 with other radiation resistant strains of *E. coli*. By doing so, genes with common significant changes in expression levels among the different radiation resistant strains can be deduced to have important roles in radiation resistance in *E. coli*.

**Acknowledgements**

I will like to thank Professor Michael M. Cox for his guidance, advice and patience. I am grateful to Professor Cox for having faith in me and giving me this golden opportunity to develop my research skills during my undergraduate years. I will also like to thank Liz Woods for her help and instructions. Another person who I want to thank is Michael Hobbs. Michael Hobbs guided me during the early stages of my research in the Cox lab

and was always willing to lend a helping hand. Last but not least, I will like to thank everyone in the Cox lab for making my stay in the Cox lab both meaningful and enjoyable.

## References


1. Ulmer, K. M., Gomez, R. F. and Sinskey, A. J. (1979) *Journal of Bacteriology* **138(2)**, 486-491.
2. Demple, B. and Halbrook, J. (1983) *Nature* **304**, 466-468.
3. Mount, D. W., Low, K. B. and Edmiston, S. J. (1972) *Journal of Bacteriology* **112(2)**, 886-893.
4. Mitchel, R. E. J. and Morrison, D. P. (1982) *Radiation Research* **90**, 284-291.
5. Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Zhai, M., Leapman, R. D., Lai, B., Ravel, B., Li, S-M., W., Kemner, K. M. and Fredrickson, J. K. (2007) *Public Library of Science Biology* **5(4)**, 92
6. Daly, M. J., Ouyang, L., Fuchs, P. and Minton, K. W. (1994) *Journal of Bacteriology* **176(12)**, 3508-3517.
7. Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P. and Tainer, J. A. (1999) *Annual Review of Biophysics and Bimolecular Structure* **28**, 101-128.
8. Braithwaite, E., Wu, X. and Wang, Z. (1999) *Mutation Research* **424(1-2)**, 207-219.
9. Zhou, B-B. S., Elledge, S. J. (2000) *Nature* **408**, 433-439.