

**The genetic equidistance result of molecular evolution is independent of
mutation rates**

Shi Huang, Ph.D.

The Burnham Institute for Medical Research

10901 North Torrey Pines Roads

La Jolla, CA 92037

shuang@burnham.org

Tel: 1-858-646-3120

Fax: 1-858-646-3192

Key words: Genetic equidistance result, evolution, molecular clock,

Running title: The genetic equidistance result is independent of mutation rates

Abstract

The genetic equidistance result shows that sister species are approximately equidistant to an outgroup as measured by DNA or protein dissimilarity. The equidistance result is the most direct evidence, and remains the only evidence, for the constant mutation rate interpretation of this result, known as the molecular clock. However, data independent of the equidistance result have steadily accumulated in recent years that often violate a constant mutation rate. Many have automatically inferred non-equidistance whenever a non-constant mutation rate was observed, based on the unproven assumption that the equidistance result is an outcome of constant mutation rate. Here it is shown that the equidistance result remains valid even when different species can be independently shown to have different mutation rates. A random sampling of 50 proteins shows that nearly all proteins display the equidistance result despite the fact that many proteins have non-constant mutation rates. Therefore, the genetic equidistance result does not necessarily mean a constant mutation rate. Observations of different mutation rates do not invalidate the genetic equidistance result.

Introduction:

A remarkable result of molecular evolution is the apparent linear correlation between genetic distance as measured by DNA and protein sequence dissimilarity and time of species divergence as inferred from fossil records. The result has been commonly interpreted to mean a constant mutation rate, which in turn directly provoked the molecular clock hypothesis.

Two kinds of sequence alignment can be made using the same set of sequence data. The first aligns a recently evolved organism such as a mammal against those that evolved earlier such as amphibians and fishes. The second aligns an outgroup organisms such as fishes against those sister species that appeared later such as amphibians and mammals. The first alignment indicates a linear correlation between genetic distance and time of divergence, implying indirectly a constant mutation rate among different species. The second alignment shows the genetic equidistance result where sister species are approximately equidistant to the outgroup. This directly triggered the idea of constant mutation rate among different species. Since both alignments use the same sequence data set, either alone is sufficient to reveal any information on genetic distance. But the data that most directly and obviously support the interpretation of a constant mutation rate is the genetic equidistance result.

The molecular clock hypothesis was first informally proposed in 1962 based largely on data from the first alignment (Zuckerandl and Pauling 1962). Margoliash in 1963 performed both alignments and made a formal statement of the molecular clock after noticing the genetic equidistance result (Margoliash 1963). "It appears that the number of residue differences between cytochrome C of any two species is mostly conditioned by the time elapsed since the lines of evolution leading to these two species originally diverged. If this is correct, the cytochrome c of all mammals should be equally different from the cytochrome c of all birds. Since fish diverges from the main stem of vertebrate evolution earlier than ether birds or mammals, the cytochrome c of both mammals and birds should

be equally different from the cytochrome c of fish. Similarly, all vertebrate cytochrome c should be equally different from the yeast protein.”

The molecular clock hypothesis asserts that the rate of amino acid or nucleotide substitution is approximately constant per year over evolutionary time and among different species. Two different species are thought to gradually accumulate mutations over time since their most recent common ancestor. Their genetic distance in ancient times is thought to be smaller than their distance today.

The genetic equidistance result merely shows the outcome of evolution and says nothing about the mutation process during the past history of evolution. In contrast, the common interpretation or restatement of this result, i.e., the constant mutation rate or molecular clock, is all about the mutation process in the past history of evolution. So there is a clear distinction in meaning between the equidistance result and its common interpretation known as ‘the constant mutation rate’. The equidistance result does not necessarily entail a constant mutation rate or any other ideas about the mutation process, while the constant mutation rate idea covers the equidistance result and much more and represents an over-interpretation of the actual result.

The equidistance result could just as easily provoke many other ideas about the actual mutation process in the past. One obvious one is that God made the equidistance result. To distinguish between competing ideas that could all be equally deduced from an empirical fact, the only scientific way is to perform experimental tests of the predictions of these ideas. The God idea is obviously not testable and is hence not a scientific idea. Remarkably, the constant mutation rate idea, despite sounding scientific, is also not testable as no one can perform other independent tests of mutation rates on geological timescales (> 1 million years). All tests on genealogical timescales (< 1 million years) in fact falsify the constant mutation rate idea. It is easy to show that different species have different mutation rates when measured on genealogical timescales, not to mention their difference in generation times. Against all expectations, such differences are irrelevant to the evolution outcome of genetic equidistance

over geological timescales. As noted by Nei and Kumar, “the constant rate of evolution was unthinkable for classical evolutionists.” (Nei and Kumar 2000).

Therefore the equidistance result so far has not provoked a truly scientific interpretation that can be tested. The constant mutation rate idea has often been violated when it was given an independent meaning (from the equidistance result) that is testable (Ayala 1999; Ho and Larson 2006; Pulquerio and Nichols 2007). But it is non-testable or non-scientific when it has no independent meaning or merely means a restatement of the empirical result of equidistance. It is correct only in the trivial sense of tautology. It is true as a factual restatement of the equidistance result. But it has not been independently proven true as a scientific explanation of the equidistance result. It is not true as a scientific hypothesis of the mutation process that can be actually measured in real time using extant organisms.

The tautology fallacy of the constant mutation rate interpretation can be illustrated by a simple example. Two turtles and a rabbit are running a 1-mile race. No one watches the race and one is only informed of the race result by a video camera aimed at the finish line. The result of the race is that the turtles and rabbit arrive at the finish line at approximately the same time in 1 hour. To explain this fact, one can deduce the same speed hypothesis. One can also deduce many other hypotheses such as ‘God did it’. To determine which hypothesis is correct, we must perform independent tests of the predictions of each hypothesis. For it to be a true explanation and not a tautology, the same speed hypothesis or any other hypothesis must be backed up by independent evidence. Of course, any independent tests of running speed would reveal that the two turtles have similar speeds while the rabbit is much faster. After performing such independent tests, one can conclude that the same speed hypothesis is likely a true explanation for the two turtles but cannot be true for the rabbit. The hypothesis is a real explanation for the two turtles but is merely a tautology for the rabbit.

The molecular clock interpretation of the equidistance result is the equivalent of the same speed hypothesis for the turtle and rabbit race. If it explains the equidistance of

organisms that can be independently measured to have similar mutation rates and generation times, the same idea cannot also explain the equidistance of organisms that can be independently measured to have different mutation rates and different generation times.

The existence of large amount of data that contradict the predictions of the constant mutation rate idea shows that the idea is a tautology. What we have as fact is that human and frog are equidistant to fish. This fact is then restated by the molecular clock interpretation as that the human and frog lineage have the same mutation rate. For this restatement to not to be a tautology, the idea of 'having the same mutation rate' cannot be supported only by the equidistance fact that the clock hypothesis is trying to explain in the first place. The same mutation rate idea should be tested by independent methods that can actually measure the rate in real time, such as pedigree analysis and counting generation times. Of course, it is well known that when people actually did that, they found that mutation rates and generation times differ vastly in different organisms. So when the constant rate idea is testable and tested, it is falsified. When the idea is not independently testable, such as independent measurements of mutation rates covering the same time frame on geological time scales, it no longer qualifies as a scientific idea. So the equidistance result remains unexplained by a truly scientific hypothesis that can be independently tested to be true.

The automatic rephrasing of the equidistance result as 'the constant mutation rate' has hindered a direct understanding of the equidistance result. All past efforts on this empirical observation have focused instead on explaining the constant mutation rate as if it were an empirical fact of the past mutation process. Various selectionist ideas as well as non-selectionist ideas have been proposed in the past to account for the constant mutation rate (Clarke 1970; Kimura 1968; Kimura and Ohta 1971; King and Jukes 1964; Richmond 1970; Van Valen 1974). The 'Neutral Theory' has come out as the favorite. But this theory is now widely acknowledged to be an incomplete explanation. For example, Ayala noted: "The theoretical foundation originally proposed for the clock, namely the neutrality theory of molecular evolution,

is untenable. The vagaries of molecular rates of evolution have contributed much to invalidating the theory.”(Ayala 1999). Pulquerio and Nichols noted: “The ‘Neutral Theory’ is not a complete explanation, however. For example, it predicts a constant substitution rate per generation, whereas empirical evidence suggests something closer to a constant rate per year.” (Pulquerio and Nichols 2007). Thus, despite numerous efforts in the past 45 years, the constant mutation rate remains unexplained by any fundamental principle of biology. However, no one has even attempted to explain the real original empirical fact, the genetic equidistance result, without presupposing a constant mutation rate.

The constant mutation rate idea has often been mistakenly treated as the same thing as the equidistance result. What has become popular known is the constant mutation rate interpretation rather than the original result itself. Even the original discoverer of this result, E. Margoliash, has subsequently avoided highlighting the result. In a 1967 paper, Fitch and Margoliash compared the cytochrome c of 20 species. Table 3 of the paper clearly showed the genetic equidistance result, for example, the yeast *Saccharomyces* has 57 mutational differences from the yeast *Neurospora*, 57 from monkey, 56 from human, and 58 from kangaroo. But Fitch and Margoliash did not comment on the obvious equidistance and instead concluded the opposite. “Indeed, from any phylogenetic ancestor, today’s descendants are equidistant with respect to time but not, as computations show, equidistant genetically. Thus the method indicates those lines in which the gene has undergone the more rapid changes. For example, from the point at which the primates separate from the other mammals, there are, on the average, 7.5 mutations in the descent of the former and 5.8 in that of the latter, indicating that the change in the cytochrome c gene has been much more rapid in the descent of the primates than in that of the other mammals.”

But the genetic equidistance result shows minor sampling variations around a mean. The eagerness to interpret insignificant variations of the equidistance result as significant

differences in mutation rates probably reflects a compromise to accommodate the mindset of classical evolution biologists who view the idea of a constant mutation rate “unthinkable”.

The common but unjustified practice of interpreting minor variations from exact equidistance as significant has caused many biologists to be unaware of the equidistance result. Whenever the constant mutation rate idea is violated, many would automatically infer that there would be no equidistance. It is commonly thought that if there is no constant mutation rate, there is no equidistance result. Here I show that the equidistance result remains valid regardless of independent results showing violations of the constant mutation rate.

Results and Discussion:

The genetic equidistance result is independent of variation in mutation rates in different species

It can be easily shown that different species have different mutation rates. A typical violation of the constant mutation rate can be illustrated by the Lsd1 protein. The time of divergence for two different bony fishes such as pufferfish (*T. nigroviridis*) and zebra fish (*D. rerio*) is ~ 140-200 MyBP (million years before present) as inferred from fossil records (Powers 1991), or from slow evolving proteins such as cytochrome c (unpublished observation). However, the genetic distance between the two fishes (13% dissimilarity in protein sequence) in Lsd1 is greater than that between chickens and mice (6% dissimilarity) which diverged ~ 310 MyBP, much earlier than the two fishes. This indicates that the mutation rate in Lsd1 is higher in fishes than in birds and mammals. However, this protein shows the equidistance result where sea urchins are approximately equidistant to all vertebrates (31% dissimilarity to fishes, 30% to frogs, 27% to chickens, 28% to mice). So violation of a constant mutation rate does not mean violation of the genetic equidistance result. For a protein such as cytochrome c, the fishes have comparable mutation rate as birds and mammals and it is well known that most

vertebrates are equidistant to an outgroup in this protein. The equidistance result therefore holds for both types of proteins that either has a constant mutation rate or has not. It is independent of mutation rate variations.

I also found that flowering plants have higher mutation rates than mammals and yet flowering plants and mammals are still equidistant to the outgroup protists. Biology textbooks commonly teach that flowering plants and animals coevolved. Based on the fossil record, the first flowering plants evolved at about the same time as the earliest mammals during the early Cretaceous period, about 125 MyBP. I randomly selected 5 proteins from the apple tree (*M. domestica*) and determined the sequence identity in these 5 proteins between the apple tree and the flowering plant *A. thaliana* (Table 1). The time of divergence between these two flowering plants is not precisely known but must be less than 125 MyBP. I also determined the sequence identity in these 5 proteins between two highly diverged mammals (human and cattle or *B. taurus*), between human and bird (*G. gallus*), between human and amphibians (*X. tropicalis*), and between human and fish (*D. rerio*). As shown in Table 1, the sequence identity between the two flowering plants is much less than that between the two mammals and is equivalent to that between human and fish. So, the flowering plants have reached a genetic distance that is much higher than that reached by the mammals after about the same amount of time of evolution. The genetic distance of flowering plants after less than 125 million years of evolution is about equivalent to that reached by vertebrates after 450 million years of evolution.

Yet, despite the faster rate of genetic divergence in flowering plants, they and mammals are equidistant to their common outgroup protists. For example, for the EF1a gene, the alveolata protist (*S. lemnae*) is 74% identical to humans and 73% identical to *A. thaliana*. Again, violation of a constant mutation rate does not mean violation of the genetic equidistance result.

It is commonly argued that the molecular clock is a stochastic clock. It does not tick at a constant rate like a real clock. It is sometimes slow and sometimes fast. It is erratic and unpredictable. But the *average* rate over long time is constant and predictable. Thus to explain

the equidistance to sea urchin of zebra fish and mouse, when zebra fishes can be shown to have faster mutation rates than mouse in the last 140-200 million years as discussed above, it is argued that the ancestor of zebra fish must have had slower mutation rate than the ancestor of mouse. Similarly, to explain the equidistance to protests of flowering plants and mammals, when flowering plants can be shown to have faster mutation rates than mammals in the last ~125 million years as discussed above, it is argued that the ancestor of flowering plants must have had slower mutation rate than the ancestor of mammals.

Such argument has several fatal flaws. First, it is not testable and hence not scientific. It cannot be expected to have independent factual support and is merely a tautology. Second, it does not have a biological reason or mechanism. It is not a deduction of a fundamental biological principle. It has no independent merit and cannot exist independent of the equidistance result. Finally, it is not logical. The constant mutation rate idea is obviously a sensible explanation for the equidistance to sea urchins for a million different individuals of zebra fishes that can be independently confirmed to have the same mutation rate. By logical inference, if the constant mutation rate idea is a true explanation for the equidistance of different organisms that can be independently confirmed to have the same mutation rate, then it already means that different organisms that can be independently confirmed to have different mutation rates would not be equidistant to an outgroup. The same idea therefore cannot also be the reason for the equidistance of different organisms that can be independently confirmed to have different mutation rates.

If millions of descendant species of bacteria are all equidistant to bacteria, regardless of their great variations in mutation rates that can be independently measured, the most straightforward logical conclusion would be that the equidistance result is not related to mutation rates. If we insist on restating such a result as constant mutation rate, we still have not explained the biological reason for the constant mutation rate, since no theories so far proposed

to explain the constant mutation rate are complete explanations. From such restatement, we have learned nothing about the biology behind the genetic equidistance result.

Most proteins show the genetic equidistance result

Many proteins are found to violate the molecular clock in experiments examining the genetic distance between similar species such as two different fishes. For example, pufferfish (*T. rubripes*) and zebrafish (*D. rerio*) are believed to have diverged not more than 140-200 MyBP (million years before present) based on the first fossil evidence of Teleostei in the early Cretaceous period (Powers 1991). One would expect most genes to show more identity between the fishes than between human and bird since the time of divergence for human and bird is much earlier (~ 310 MyBP). In a survey of 40 randomly picked proteins, I found only 19 (48%) with more identity between the two fishes than between human and bird. So about half of all genes in fishes have faster mutation rate than the molecular clock deduced from macroevolution of vertebrates. It is now common practice to exclude these genes in calculating divergence time for microevolution (Kumar and Hedges 1998).

The fact that about half of all genes have different mutation rates in different species offers another way to resolve whether the genetic equidistance result is independent of the measurable variations in mutation rates in different species. If most genes can be shown to display the equidistance result despite the independent fact that half of them have different mutation rates in different species, then we can conclude that the equidistance result is independent of rate variations.

I randomly selected 50 proteins from frogs (*X. laevis*) and compared each to chickens (*G. gallus*) and humans (*H. sapiens*). Among these proteins, 11 (22%) showed exact equal distance (to frogs) of chickens and humans, 28 (56%) showed greater distance between humans and frogs than between chicken and frogs, and 11 (22%) showed less. For most of these proteins (46/50 or 92%), the difference between chicken and human in their percent

identities to frogs is less than 4% (Table 2), indicating approximate equidistance. For 4 other proteins (4/50 or 8%), the difference between chicken and human in their percent identities to frogs is 7% to 8%. However, all 4 proteins showed approximate equidistance when sea urchins were used as the outgroup (Table 2). Thus, the seeming non-equidistance to frogs in these 4 proteins may not represent a significant violation of the equidistance result. Since all of the 50 randomly selected proteins showed the equidistance result, whereas one expects only half of them since at least half is known to have non-constant mutation rates, the data suggest that the equidistance result is independent of the constancy of mutation rates ($P < 0.0001$). It also suggests that nearly all vertebrate proteins show the genetic equidistance result.

The proper way to establish that small variations in distance are not significant is to sample multiple individuals of each sister species. A single individual of species A may be either more or less distant to an outgroup than a single individual from sister species B. However, if large number of individuals were analyzed, the mean distance to the outgroup should not be significantly different between the two sister species. Also, the number of comparisons that show A to be more distant to the outgroup than B should be similar to the number of comparisons that show A to be less distant to the outgroup than B. This kind of analysis has shown that humans and chimpanzees are equidistant to gorillas (Huang, 2008). In a study using mitochondrial DNA from 30 randomly selected human individuals and 30 chimpanzee individuals, the number of comparisons that showed greater distance between humans and gorillas than between chimpanzees and gorillas (13) was similar to the number of comparisons that showed greater distance between chimpanzees and gorillas than between humans and gorillas (11), while 6 showed that human and chimpanzees are exactly equal distant to gorillas (Huang, 2008).

At this point in time, for most species, we do not yet have sequence information for multiple individuals of a species. Thus it is not yet possible to statistically establish that the small variations in equidistance in many cases are indeed non-significant. However, given the

overwhelming data of approximate equidistance, when expectation based on non-constant mutation rates would be much greater variations in distance, it is easy to infer that the real result here is equidistance (with minor coincidental variations from the mean) rather than non-equidistance with equidistance being coincidental. It can therefore be reasonably concluded that the minor variations from exact equidistance are not significant in most cases. Indeed, if the equidistance result were not real, the molecular clock idea would not have been invented in the first place.

Some common practice such as the relative rate test has often interpreted small variations from exact equidistance to be significant. Many evolution biologists who perform such tests mistakenly consider the real phenomenon to be non-equidistance with equidistance being coincidental. But the relative rate test is flawed because it does not consider sampling variations. It also presupposes the truth of the gradual mutation model of speciation when it remains an open question whether genetic distance had always increased with time in the past history of life on Earth. The recent analysis of fossil organisms in fact shows that genetic distance had not always increased with time in the past history of life on Earth (Huang, 2008).

To consider small differences as being significant also makes it impossible to reconcile it with other contradicting facts. For example, the albumin protein of a specific bird individual is 47% identical to that of a specific human and 44% identical to that of a specific rat. Some evolution biologists have viewed such small differences to be statistically significant after performing the relative rate test (Nei and Kumar 2000). This however contradicts the fact that a frog (*X. tropicalis*) albumin gene is 38% identical to human and 40% to rat. It is impossible for the rat lineage to have a faster mutation rate than humans when birds are the outgroup but a slower mutation rate than humans when frogs are the outgroup. If the faster mutation rate than humans with birds as the outgroup is real, the rate with frogs as the outgroup can only be faster and cannot possibly be slower or equal, since rats and humans do not have separate ancestors prior to the frog to bird transition. Therefore, the facts can only be explained by considering such

small differences as insignificant variations of the equidistance result. Rats and humans are equidistant to birds as well as to frogs. All different mammals are equidistant to birds in the range of 43-47% identity in the albumin gene.

Conclusions:

The genetic equidistance result is arguably the most remarkable result of molecular evolution. The biology behind this result has remained obscure despite the past 45 years of research. The equidistance result could trigger many interpretations but the idea of constant mutation rate has become the most popular. However, this idea is not testable and there is no independent evidence for it other than the equidistance result that originally provoked it. It is merely a tautology. The observation of frequent violations of the constant mutation rate has misled many to automatically assume that there is no equidistance result. The study here establishes the fact that the equidistance result is an empirical observation that is independent of variation in mutation rates. The equidistance result shows the outcome of evolution but does not directly reveal any information about the mutation process in the past history of life on Earth. The constant mutation rate interpretation of this result represents an over-interpretation. New ideas are needed to explain the equidistance result that must be independently testable.

Acknowledgments:

This work was supported by NIH (RO1 CA 105347).

References:

- Ayala, F. J. 1999. Molecular clock mirages. *Bioessays* 21:71-75.
- Clarke, B. 1970. Darwinian evolution of proteins. *Science* 168:1009-1011.
- Ho, S. Y. W., and G. Larson. 2006. Molecular clocks: when times are a-changin'. *Trends Genet.* 22:79-83.
- Huang, S. 2008, Ancient fossil specimens of extinct species are genetically more distant to an outgroup than extant sister species are. In press, *Riv. Biol.* Available from Nature Precedings at <http://precedings.nature.com/documents/1676/version/1>
- Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624-626.
- Kimura, M., and T. Ohta. 1971. On the rate of molecular evolution. *J. Mol. Evol.* 1:1-17.
- King, J. L., and T. H. Jukes. 1964. Non-Darwinian evolution. *Science* 164:788-798.
- Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917-920.
- Margoliash, E. 1963. Primary structure and evolution of cytochrome c. *Proc. Natl. Acad. Sci.* 50:672-679.
- Nei, M., and S. Kumar. 2000, *Molecular evolution and phylogenetics*. New York, Oxford University Press.
- Powers, D. A. 1991. Evolutionary genetics of fish. *Advances in Genetics* 29:119-228.
- Pulquerio, M. J., and R. A. Nichols. 2007. Dates from the molecular clock: how wrong can we be? *Trends Ecol. Evol.* 22:180-184.
- Richmond, R. C. 1970. Non-Darwinian evolution: a critique. *Nature* 225:1025-1028.
- Van Valen, L. 1974. Molecular evolution as predicted by natural selection. *J. Mol. Evol.* 3:89-101.

Zuckerkandl, E., and L. Pauling. 1962, Molecular disease, evolution, and genetic heterogeneity: In *Horizons in Biochemistry* (Kasha, M. and B. Pullman, eds), New York, Academic Press.

Table 1. Genetic distance within flowering plants is greater than that within mammals after similar amount of time of evolution. Five proteins from the apple tree (*M. domestica*) were randomly selected for determining the genetic distance between the apple tree and the flowering plant *A. thaliana*, between human and cattle (*B. taurus*), between human and bird (*G. gallus*), between human and amphibians (*X. tropicalis*), and between human and fish (*D. rerio*).

	<u>Percent identity</u>					<u>MyBP</u>
	<u>ADH</u>	<u>EF1a</u>	<u>eIF2b</u>	<u>Pin1</u>	<u>PP1</u>	
<i>M. domestica</i> v.s. <i>A. thaliana</i>	83	95	73	79	92	125
<i>H. sapiens</i> v.s. <i>B. Taurus</i>	94	100	96	98	100	125
<i>H. sapiens</i> v.s. <i>G. gallus</i>	87	98	94	N.A.	100	310
<i>H. sapiens</i> v.s. <i>X. tropicalis</i>	86	96	86	86	99	360
<i>H. sapiens</i> v.s. <i>D. rerio</i>	80	94	77	79	97	450

Table 2. The genetic equidistance result for 50 randomly selected proteins. Fifty proteins were randomly selected from frogs/*X. laevis* (F) and compared with chicken/*G. gallus* (C) and human (H). Percent identities in protein sequence are shown. For 4 of these proteins that showed greater variations from exact equidistance, a comparison with sea urchin/*S. purpuratus* (SE) was made to confirm approximate equidistance. Protein names and accession numbers are shown.

	<u>Percent identity</u>			
	<u>F-C</u>	<u>F-H</u>	<u>F-SE</u>	<u>H-SE</u>
HGF/AAB34354	70	67		
mN-cadherin/ AAB37685	74	73		
TrkB/ NP_001079579	95	94		
APP747/ AAB24853	86	86		
Rb1/ AAB23173	61	60		
ADPRT/ P31669	79	77		
B-catenin/ NP_001084045	97	97		
Chordin/ NP_001081778	59	51	34	34
c-Mos/ NP_001081563	57	53		
PR/ NP_001079100	64	68		
Chk1/ NP_001082039	75	76		
P53/ NP_001081567	52	51		
Hsp70/ P02827	95	97		
Prion/ NP_001082180	40	36		
Noggin/ NP_001079113	86	78	28	27
Ralb/ NP_001084154	94	94		
Desmin/ NP_001080177	81	78		
AAI60739	39	42		
AAI60733	39	36		
AAI60732	72	69		
AAI60718	64	63		
AAI60762	64	62		
AAI60704	34	36		
P85PI3K/ AAL68953	81	80		
DNA ligase IV/ AAL56554	71	72		
Stat5/ AAK94906	86	86		
P2X4/ AAG45104	70	70		
HIRA/ CAC41093	77	79		
DNase1/BAB20384	42	44		
Lis1/ AAK52334	96	96		
P450/ BAA37080	63	62		
Ptc1/AAK15463	81	79		

Paxillin/AAG40874	75	75		
Cdc21/NP_001081448	86	85		
ND4/AAA66467	60	59		
ND1/AAA66458	71	64	62	59
Rest/ AAF06720	64	61		
PTEN/AAD46165	89	88		
EphB2/ AAB94603	87	87		
RpS6/ AAD01647	95	96		
Cdc46/BAA09949	87	86		
Mcm2/ BAA09948	87	85		
DNMT1/ BAA11458	76	77		
CCND1/ CAA61664	84	77	45	45
EIF4E/ NP_001084017	92	91		
AAI60783	45	44		
Sdad1/ AAH70991	80	80		
COUP-TF2/ NP_001095229	77	75		
Src/ NP_001079114	87	87		
Serpin/NP_001079103	38	38		