**Do SNPs in glutathione *S*-transferase-omega allow predictions of the susceptibility of vertebrates to SARS-CoV-2?**

John T. Hancock1, David Veal1, Tim J. Craig1 and Ros C. Rouse2

1. Department of Applied Sciences, University of the West of England, Bristol, UK
2. Research, Business and Innovation(RBI), University of the West of England, Bristol, UK

\*Correspondence:

Prof. John T. Hancock

Faculty of Health and Applied Sciences,

University of the West of England, Bristol, BS16 1QY, UK.

[john.hancock@uwe.ac.uk](mailto:john.hancock@uwe.ac.uk)

Short title: COVID-19 and GST

**Abstract**

Infection with the SARS-Cov-2 virus causes COVID-19 in humans, and is the cause of the pandemic around the world in 2020 and on. However, some animals have been found to be susceptible to the virus too. This has included the non-human primates, dogs, cats and mustelids. Mink, and very recently hamsters and deer, have been shown to be able to contract the virus and pass it back to humans. However, which animals are susceptible to the virus has been very hard to predict. Many groups have looked at the sequence homology of the angiotensin converting enzyme 2 receptor (ACE2) across species, but this has had limited success. Similar work on other proteins such as Transmembrane Serine Protease 2 (TMPRSS2), neuropilin-1 and furin have also been unfruitful. Recently it has been suggested that single nucleotide polymorphisms (SNPs) in the glutathione *S*-transferase-omega (GSTO) genes of humans could alter viral susceptibility. Therefore, here, the presence of related sequences in vertebrates has been investigated. The SNPs in the GST-omega-1 (GSTO1) gene reported to increase COVID-19 in humans do not appear in the vertebrate species. However, the GST-omega-2 (GSTO2) SNP is represented in several vertebrate species known to have contracted the SAR-CoV-2 virus. Of course, animals may contain unknown SNPs at disruptive points in these genes too. In summary, GST-omega-1 genes are unlikely, at least at the moment, to be of value in predicting the susceptibility of an animal to the SARS-CoV-2 virus or disease progression, but a further study of the GST-omega-2 genes would be worthwhile. Therefore, more work on SARS-CoV-2 infections on vertebrates is recommended.

**Keywords**: ACE2; COVID-19; GST; polymorphisms; SARS-CoV-2; SNPs.

**Introduction**

The SARS-CoV-2 virus caused the COVID-19 pandemic which has spread across the world. At present (28th February 2022) there have been 435,626,514 confirmed cases and 5,952,215 deaths worldwide1. The virus is an RNA-containing enveloped virus, and has been shown to be transmitted through respiratory droplets and being deposited on surfaces2. The virus interacts and then enters the host cells through the interaction of the viral surface proteins with the cell surface. This involves spike proteins3 which protrude out of the virus and give it its image, and hence name, as being a coronavirus. The host cell protein of note that the virus interacts with is the angiotensin converting enzyme 2 receptor (ACE2)4, although other proteins are also involved, such as neurophilin5. Severe COVID-19 is underpinned by what is termed the cytokine storm6, and this can lead to death.

The SARS-Cov-2 is thought to have derived from a virus from a bat7, but exactly how it entered the human population is still contested8. However, regardless of its origin, it is known that several species of animals have been found to be SARS-CoV-2 positive. Susceptible animal species include non-human primates, dogs, cats and mustelids. This has been reviewed before9-13. However, two factors are important here. Firstly, contraction of the virus has been found to be fatal in some animals. For example, lions have died at zoos in Honolulu14 and Chennai, India15. Secondly, the virus can be passed from animal to animal, such has been found in cats16, but more concerning, it can be passed back to humans. This was first reported in mink17,18, and led to a mass euthanasia programme for mink in several mink farms, for example in Denmark19. More recently hamsters were thought to have given the virus back to humans in a pet shop in Hong Kong20, and were subsequently culled, whilst deer were a worry for the same reason in Canada21. The main concern for human health is the potential for the virus to mutate in the animal before re-transmission back to humans, and therefore the efficacy of vaccines may be compromised, and for animals’ health in relation to the impact on endangered populations which may be susceptible, such as black footed ferrets(*Mustela nigripes*), and the culling of domestic species during an outbreak.

There is, therefore, a need to be able to predict which animals are susceptible to the SARS-Cov-2 virus. This has in the past concentrated on looking at the ACE2 receptor, which is the primary protein which interacts with the viral spike proteins4. Similarity analysis between the human ACE sequences and those found in animal genomes has been carried out extensively by others9,22,23. However, the predictions do not appear to match reality. Animals which have been reported to be SARS-CoV-2 susceptible do not show sequence similarities in their ACE2 genes. Therefore, other proteins were examined too, including Transmembrane Serine Protease 2 (TMPRSS2), neuropilin-1 and furin10. Again, limited usefulness was found. Not surprisingly, the animals most similar to humans, such as non-human primates, were predicted to be susceptible to the virus, and this has been borne out in reality24. At the lower end of the vertebrate evolutionary scale, reptiles, fish, amphibians and birds all seem to be predicted to not be susceptible to SARS-CoV-2, and there have been to date no reports of such animals being infected with the SARS-CoV-2 virus. However, in between this range of animals, viral susceptibility seems hard to predict. For example, the mustelids such as mink were not predicted to be highly susceptible and yet mink is one of the species most affected. Similarly, the susceptibility to the virus of dogs and cats does not jump out of the sequence alignment data reported, and yet again, these animals seem to be being tested positive for the virus more than predicted. Being companion animals, this may be due to their close contact with infected people. Therefore, there is a need for more analysis of the proteins involved in the SARS-Cov-2 infection process. Recently, single nucleotide polymorphisms (SNPs) within the genes of two isoforms of glutathione *S*-transferase- Omega (GST-omega; GSTO) have been found to correlate with COVID-2 severity25. Therefore, these gene sequences were aligned in animal species to see if there was any evidence that the SNPs were represented and therefore could lead to a better understanding as to why certain animals are SARS-CoV-2 susceptible.

**Methods**

Representative vertebrate species were chosen and the sequences for GSTO1 and GSTO2 were obtained from the National Center for Biotechnology Information (NCBI) database. For some animal species there were several isoforms or variants available and these were extracted too (see Tables 1 & 2). Alignments of the sequences were undertaken using Clustal Omega26. The multiple sequence alignment tool, Clustal Omega, is based on seeded guide trees and Hidden Markov Model (HMM) profile-profile techniques. Further analysis of functional and control sequences were undertaken using Prosite27.

**Results**

**Previous analysis of the GSTO SNPs**

Glutathione-*S* transferase (GST) is a superfamily of proteins28. GST proteins have a range of roles. They are involved in antioxidant metabolism, and hence reactive oxygen species (ROS) metabolism, and are instrumental in many cell signalling pathways and in the detoxification of certain molecules29. In mammals there are eight classes of cytosolic GST enzymes: alpha ((α)-*GSTA*); mu ((μ)-*GSTM*); pi ((π)-*GSTP*); omega ((ω)-*GSTO*); theta ((θ)-*GSTT*); sigma ((σ)-*GSTS*); kappa ((κ)-*GSTK*); zeta ((ζ)-*GSTZ*). Of relevance here GSTO1 has a role in chronic obstructive pulmonary disease (COPD), through its action as an antioxidant and its control of glutathione (GSH) homeostasis. GSTO1 activity also impacts in the inflammatory response through its action on the lipopolysaccharide (LPS)/Toll-like receptor (TLR-4) and the subsequent downstream activation of nuclear factor-kappa B (NF-ĸB), and hence gene expression. There is also a physiological inhibitor of GSTO1 which can impact the inflammatory response25. Therefore, it appears that this enzyme has a significant role in inflammation: COVID-19 severity in humans involves a cytokine storm, and hence loss of control of the inflammatory response. However, even though some animals are reported to be virus positive, this does not necessarily result in a cytokine storm, with some animals appearing to have less severe symptoms. However, GSTO1 polymorphisms were the focus of a recent study25. GSTO2, meanwhile, is also involved in redox homeostasis and control of vitamin C levels. Vitamin C has been shown to be beneficial for patients with severe COVID-1930. As discussed below, again a SNP in this enzyme’s gene sequence was analysed with respect to COVID-19 severity too25.

The idea of GST SNPs being important in the onset and severity of COVID-19 has been reported before. For example, it was found that patients with the GSTT1−/− genotype (i.e., lacking the theta 1 form) had high mortality31. However, recently a paper focused on SNPs in the GST-omega genes and how their presence correlates with the susceptibility and severity of SARS-Cov-2 infection in humans25. Interestingly, it was found that polymorphisms in the ACE2 receptor (rs4646116) were not useful as a predictor of susceptibility, similar to that reported previously10. However, SNPs in GSTO1 (rs4925) and GSTO2 (rs156697) correlated with COVID-19 infection. 255 COVID-19 patients and 236 matched healthy individuals were assessed, and the correlation of disease severity and presence of polymorphisms analysed25. Haplotype analysis was used to confirm the findings. The authors say that the presence of these SNPs in GSTO1 or GSTO2 reached a “statistically significant association” with the development of the disease. The presence of GSTO1\*AA (as opposed to GSTO1\*CC) showed significantly increased COVID-19 development (2.45 fold), whilst presence of the GSTO2\*GG SNP (as opposed to GSTO2\*AA) led to a 3.7 increase in disease development.

In GSTO1 the rs4925 SNP leads to the amino acid change of A140D, whilst the GSTO2 SNP rs156697 leads to a N142D (or N142Y) change in the amino acid sequence of that polypeptide. Therefore, if, as stated, these changes can increase the development and disease severity of COVID-19 in humans, are these “new” amino acid sequences present in the GSTO genes of animals and can their presence be used as a predictor of COVID-19 in those animal species?

**Analysis of GSTO genes, SNPs and animals**

As stated above, the GSTO1 SNP (rs4925) leads to a nucleotide change of C to A, and this translates to an amino change of A140D (Ala to Asp). Here, three GSTO1 isoforms from humans were aligned with homologous sequences from twenty vertebrate species (Table 1).

Table1: GSTO1 sequences used for alignments

|  |  |  |
| --- | --- | --- |
| **Species** | **Latin name** | **Sequence used** |
| Human isoform 1 | *Homo sapiens* | NP\_004823 |
| Human isoform 3 | *Homo sapiens* | NM\_001191003 |
| Human isoform 2 | *Homo sapiens* | NM\_001191002 |
| Chimpanzee | *Pan troglodytes* | NP\_001233486 |
| Gorilla (western lowland) isoform X1 | *Gorilla gorilla gorilla* | XP\_004050100 |
| Gorilla (western lowland) isoform X2 | *Gorilla gorilla gorilla* | XP\_018890528 |
| Rhesus monkey | *Macaca mulatta* | NP\_001247484 |
| Dog | *Canis lupus familiaris* | XP\_535007 |
| Cow (cattle) | *Bos taurus* | NP\_001068682 |
| Mouse (house) | *Mus musculus* | NP\_034492 |
| Rat (Norway) | *Rattus norvegicus* | NP\_001007603 |
| Chicken | *Gallus gallus* | NP\_001264304 |
| Zebrafish | *Danio rerio* | NP\_001002621 |
| Bat (large flying fox) | *Pteropus vampyrus* | XP\_011362082 |
| Bat (Myotis) | *Myotis myotis* | XP\_036191958 |
| Bat (Kuhl's pipistrelle) | *Pipistrellus kuhlii* | XP\_036294846 |
| Ferret (domestic) | *Mustela putorius furo* | XP\_004749360 |
| Mink (American) | *Neogale vison* | CCP84655 |
| Whale (killer) | *Orcinus orca* | XP\_033276967 |
| Dolphin (common bottlenose) isoform X1 | *Tursiops truncatus* | XP\_033697527 |
| Hamster (Chinese) isoform X1 | *Cricetulus griseus* | XP\_003515091 |
| Hamster (Chinese) isoform X2 | *Cricetulus griseus* | XP\_007653253 |
| Hamster (Chinese) isoform X3 | *Cricetulus griseus* | XP\_035294030 |
| Hamster (Chinese) isoform X4 | *Cricetulus griseus* | XP\_007653254 |
| Hamster (Chinese) isoform X5 | *Cricetulus griseus* | XP\_027262854 |
| Hamster (Chinese) isoform X6 | *Cricetulus griseus* | XP\_027262855 |
| Hamster (Chinese) isoform X7 | *Cricetulus griseus* | XP\_027262856 |
| Frog (tropical clawed) isoform X1 | *Xenopus tropicalis* | XP\_012821965 |
| Crocodile (Australian saltwater) | *Crocodylus porosus* | XP\_019411533 |
| Squirrel (grey) | *Neosciurus carolinensis* | MBZ3873070 |
| Snake | Not found |  |

Some animal species had more than one isoform in the NCBI database, so these were also included. For example, as hamsters are thought to be a species able to cause animal-to-human viral transmission, seven isoforms were analysed here. The human SNP of interest is A140D, and is highlighted as yellow in Figure 1. Although the Ala was conserved in humans and non-human primates there was little conservation of this amino acid across species, many have a Ser in this position. None analysed had the Asp which would have predicted greater disease progression. Therefore, looking for the presence of this amino acid change in GSTO1 appears to have little use in prediction of SARS-Cov-2 infection in vertebrates.

To further investigate the protein, the human isoform 1 sequence of GSTO1 was analysed using Prosite, to find domains and sequences of interest in functionality and possible control of the polypeptide. The reported SNP at residue 140 is well within the C-terminal functional domain of the protein, as shown in purple (Figure 1). Several amino acids are predicted to be phosphorylated, suggesting possible control, along with amidation and *N*-myristoylation sites. However, none of these seem to be impinged upon by the SNP, at least as predicted by this rather limited analysis (no structure was analysed here).

Similar analysis was carried out with GSTO2. The SNP here (rs156697) was at amino acid 142. This led to changes of N142D or N142Y. Interestingly, when all the human GST sequences were aligned the two SNPs are in the same region of the protein (adjacent amino acids), as can be seen in Figure 2 (highlighted in yellow and purple). This region is missing from three of the isoforms, and although there are some conserved amino acids in this section of the protein (shown in blue), the area is not overall heavily conserved.

To analyse the possible use of the GSTO2 SNP (rs156697) in SARS-Cov-2 infections in animals the four human isoform sequences were aligned with those of fifteen animal species (Table2). The SNP is shown in yellow in Figure 3.

Table 2: GSTO2 sequences used for alignments

|  |  |  |
| --- | --- | --- |
| **Species** | **Latin name** | **Sequence used** |
| Human isoform 1 | *Homo sapiens* | NP\_899062 |
| Human isoform 2 | *Homo sapiens* | NP\_001177942 |
| Human isoform 3 | *Homo sapiens* | NP\_001177943 |
| Human isoform 4 | *Homo sapiens* | NP\_001177944 |
| Gorilla | *Gorilla gorilla gorilla* | XP\_004050103 |
| Rhesus monkey | *Macaca mulatta* | NP\_001276891 |
| Mouse (house) | *Mus musculus* | NP\_084327 |
| Rat (Norway) | *Rattus norvegicus* | AAH79295 |
| Whale (killer) | *Orcinus orca* | XP\_033276969 |
| Ferret (domestic) | *Mustela putorius furo* | AER99471 |
| Dolphin (common bottlenose) | *Tursiops truncatus* | XP\_019795507 |
| Hamster (golden) isoform X1 | *Mesocricetus auratus* | XP\_040599162 |
| Hamster (golden) isoform X2 | *Mesocricetus auratus* | XP\_040599171 |
| Hamster (golden) isoform X3 | *Mesocricetus auratus* | XP\_040599174 |
| Frog (tropical clawed) | *Xenopus tropicalis* | NP\_001005086 |
| Cow (cattle) | *Bos taurus* | NP\_001193084 |
| Dog isoform X1 | *Canis lupus familiaris* | XP\_038296425 |
| Dog isoform X2 | *Canis lupus familiaris* | XP\_038296426 |
| Cat (domestic) isoform X1 | *Felis catus* | XP\_023096662 |
| Cat (domestic) isoform X2 | *Felis catus* | XP\_044896585 |
| Cat (domestic) isoform X3 | *Felis catus* | XP\_044896586 |
| Bat (greater spear-nosed) | *Phyllostomus hastatus* | XP\_045710275 |
| Bat (common vampire) | *Desmodus rotundus* | XP\_045047595 |
| Bat (big brown) | *Eptesicus fuscus* | XP\_008142501 |
| Reptile | Not found |  |
| Mink | Not found |  |

As mentioned above, not all the human isoforms of GSTO2 have the regions in which the SNP falls, but it does occur in isoforms 1 & 3. The SNP manifests as changes from N to D or Y. Interestingly, although Y at residue 142 is not represented at all in the vertebrate analysis here, many animals have a D at this position. Interestingly, gorilla and Rhesus monkey have identical sequences around the SNP but have the N142D alteration. As this increases COVID-19 development and disease severity in humans25 it is tempting to suggest this would have the same effect in these primate species and that non-human primates are more susceptible to the SARS-CoV-2 virus. The D at position 142 is also found in cats, dogs, and ferrets, all of which are known to contract the virus. It is also represented in two of the hamster sequences (X1 & X2, but not X3), again, a species known to become infected. It is also found in rats and mice, suggesting that rodents may be more susceptible, although the region of the protein is not completely conserved. The susceptibility of mice to the SARS-CoV-2 virus appears to be strain specific32. Interestingly, two of the bat species analysed retained the N at position 142, but one also had the D substitution (big brown bat). Does this suggest that some bat species are more susceptible than others? Also of note, perhaps, is that marine mammals have the D at position 142, and it is suggested here that these species ought to be monitored more closely in the future. The frog sequence appears to be quite different in this region, and amphibians are known, at least to date, to be free from SARS-CoV-2 infections.

Again, using the human isoform 1 as a model, the GSTO2 protein was analysed by Prosite. As with GSTO1, the SNP in GSTO2 was not in a region which appears to conflict with predicted phosphorylation and other modification sites (Figure 3). However, further analysis of the gene sequence of GSTO2 may be of use in the prediction of whether vertebrates are susceptible to the SARS-CoV-2 virus.

**Discussion and Conclusions**

The factors which lead to severe COVID-19 in humans has been looked at by the analysis of of genomes. 7,491 critically ill patients were compared to 48,400 controls, and sixteen independent factors which can lead to severe outcome were identified. Many of these were involved in the immune response, for example in interferon signalling, but GST was not listed33. However, GST-omega SNPs were highlighted by the work of Djukic *et al*.25 and these are the focus here.

One of the manifestations of a SARS-CoV-2 infection is the onset of a cytokine storm, where there is hyperactivation of immune cells and elevated levels of circulating cytokines34. There is also a rise in the accumulation of ROS and accompanied oxidative stress, and it has been suggested that there is a vicious circle of the rise of ROS and cytokines, which will lead to Acute Respiratory Distress Syndrome (ARDS) and potentially death35. GSTO has a role in dampening the oxidative stress response, which can have profound effects on disease36. Oxidative stress during COVID-19 can lead to the oxidation of cellular components, such as lipids, which can be part of the disease mechanism37. GST homologues, including GSTO, are known to be involved in reducing oxidative stress and to catalyse the reduction of oxidative factors in a variety of animal species38,39. Because of their structures, and having a Cys in the active site, GSTOs catalyse reactions which cannot be carried out by other GST enzymes, and GSTOs are often involved in thiol transferase and reduction reactions40. Therefore, they are instrumental in redox signaling and ROS metabolism. GSTOs are also involved in mediating aspects of apoptosis and autophagy41. Downstream effects of GSTO1-1 in macrophages, for example, were mediated by *c*-Jun N-terminal kinases (JNK), which were upstream of the cytochrome *c*/caspase pathway. Therefore, GSTOs will have an important role in the onset and maintenance of COVID-19, as well as other diseases which have a redox element to their manifestation.

Predicting the susceptibility of animals to the SARS-CoV-2 virus is turning out to be very hard to do, and little sequence analysis seems to be helpful. Despite sequence alignments across a wide range of vertebrates, using the sequence of ACE29,10,22, there was little correlation with predictions and the reality of the animals reported to be SARS-CoV-2 positive. This was more recently confirmed by Djukic *et al*.25. Therefore, other proteins were also investigated by sequence alignment, including TMPRSS2, neuropilin-1 and furin10. Again, little prediction could be made, with some of the sequences being well conserved across vertebrate species. Therefore, here, the report that GST-omega gene sequences can predict COVID-19 severity in humans25 seemed like an observation worth investigating across vertebrate species. Although little could be predicted from the GSTO1 sequences, the GSTO2 genes may give an indication of viral susceptibility in animals.

There are some species which are not represented in the study here. For example, no snake sequence could be found for GSTO1, and neither mink nor reptiles were found for GSTO2. There are obviously thousands of species not analysed here: estimates suggest around 65000 vertebrate species exist. Furthermore, there will be SNPs in the animal species which are not looked at either. Therefore, the lack of finding a predictive use here, for example with the GSTO1 sequences, does not rule out that some animal GST sequences cause an animal species, or an individual animal, to be more SARS-CoV-2 susceptible. The same will be true for all the other proteins analysed in this way, including ACE2, TMPRSS2, neuropilin-1 and furin9,10,22. Of course, in the analysis shown here there was no structural analysis of the polypeptides which would potentially highlight the significance of amino acid changes. However, although structural analysis would be better here, it is both more complex and long-winded. On the other hand, such a quick prediction by sequence analysis would be useful to guide structural work, especially considering how many vertebrate species there are. Here it is suggested that further analysis of the GSTO2 sequences would be worthwhile. There are many factors which dictate the viral entry into the cells and the production of new viruses which can be subsequently released. A simple analysis of some of the proteins involved, especially GST-omega-2 isoforms, at the present time seems like a useful pursuit.

There are now at least three animal species which are thought to be able to transmit the virus back to humans. These are mink18, hamsters20 and deer21. There is a need therefore to have a way to predict which animals are likely to be susceptible. There are many circumstances in which humans and animals are in close contact, including companion animals, zoos and at conservation sites. There are also many vertebrate species which may be at risk but have had little investigation. These include the marine mammals, which have been predicted to be susceptible42, and may be exposed through human waste43,44 which is a potential source of infection. Here, the data from the GSTO2 genes may be useful to inform future monitoring and research.

Although the virus probably originated in a bat45, and then possibly infected humans through an intermediate host, this has been disputed8. Bat species are not the only ones which need to be considered as able to transmit the virus to humans. SARS-CoV-2 is the latest in a series of coronavirus outbreaks, which includes SARS and MERS46. Therefore, the COVID-19 pandemic is unlikely to be the last. Any work which can be used to predict, or be ruled out as a method for prediction, which animals (including humans) are likely to be affected by coronaviruses in the future should aid in epidemic/pandemic control. Therefore, there needs to be far more of a focus on the impact of the SARS-CoV-2 virus in animal populations, both in monitoring what is happening in the real world and how to predict what happens, or may happen, during this and future pandemics. As already suggested by others47, a One Medicine approach needs to be taken where the impact of epidemics/pandemics on human and animal populations are considered holistically. Certainly, it is likely that new coronaviruses will emerge and work is being undertaken to predict in which animal hosts this might take place48. With more genomic analysis and more awareness of the interactions of animals with other species, and with humans, on the spread of viruses, such as SARS-CoV-2, amongst humans and animals, we will be better prepared for future challenges.

**Acknowledgements:** This work was funded by the University of the West of England, Bristol, who financed the authors’ time and literature sourcing for the preparation of this manuscript.

**Competing interests statement:** J.T.H., T.J.C., D.V. and R.C.R. all declare that they have no competing financial interests.

**References**

1- [WHO Coronavirus (COVID-19) Dashboard](https://covid19.who.int/): https://covid19.who.int/ (accessed 23/02/22)

2- Tizaoui, K., Zidi, I., Lee, K.H., Ghayda, R.A., Hong, S.H., Li, H., et al. Update of the current knowledge on genetics, evolution, immunopathogenesis, and transmission for coronavirus disease 19 (COVID-19). *Int. J. Biol. Sci*., **16**, *2906–2923* (2020)

3- Huang, Y., Yang, C., Xu, X., Xu, W., Liu, S.W. Structural and functional properties of SARS-CoV-2 spike protein: Potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin*., **41**, *1141–1149* (2020)

4- Hoffman, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. SARS-CoV-2 Cell entry Depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, **181**, *271–280* (2020)

5- Daly, J.L., Simonetti, B., Klein, K., Chen, K.-E., Williamson, M.K., Antón-Plágaro, C., et al. Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science* **370**, *861-865* (2020)

6- Wang, J., Jiang, M., Chen, X., Montaner, L.J. Cytokine storm and leukocyte changes in mild versus severe SARS-CoV-2 infection: Review of 3939 COVID-19 patients in China and emerging pathogenesis and therapy concepts. *J. Leukoc. Biol*. **108**, *17-41* (2020)

7- Zhou, P. Yang, X.-L. Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**, *270–273* (2020)

8- Chan, A. Ridley, M. *Viral: The Search for the Origin of Covid-19*. Fourth Estate (2021) ISBN: 978-0008487492

9- Damas, J., Hughes, G.M., Keough, K.C., Painter, C.A., Persky, N.S., Corbo, M., et al. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc. Nat. Acad. Sci. USA*, **117**, *22311–22322* (2020)

10- Hancock, J.T., Rouse, R.C., Stone, E. and Greenhough, A. Interacting proteins, polymorphisms and the susceptibility of animals to SARS-CoV-2. *Animals*, **11**, *797* (2021)

11- Coronavirus: Four lions test positive for Covid-19 at Barcelona Zoo: https://www.bbc.co.uk/news/world-europe-55229433 (accessed 23/02/22)

12- Mathavarajah, S. Dellaire, G. Lions, tigers and kittens too: ACE2 and susceptibility to COVID-19. *Evol. Med. Public Health* **2020**, *109–113* (2020)

13- Shi, J., Wen, Z., Zhong, G., Yang, H., Wang, C., Huang, B., et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. *Science*, **368**, *1016–1020* (2020)

14- Lion at Honolulu Zoo dies after contracting COVID in case prompting broader concern: https://www.hawaiinewsnow.com/2021/10/16/lion-honolulu-zoo-dies-after-contracting-covid-case-prompting-broader-concern-facility/ (Accessed 23/02/22)

15- Second lion dies of Covid at Indian zoo: https://www.independent.co.uk/news/science/covid-india-lion-death-zoo-b1867681.html (Accessed 23/02/22)

16- Gaudreault, N.N., Trujillo, J.D., Carossino, M., Meekins, D.A., Morozov, I., Madden, D.W., et al. SARS-CoV-2 infection, disease and transmission in domestic cats. Emerg. *Microbes Infect*., **9**, *2322–2332* (2020)

17- Oude Munnink, B.B., Sikkema, R.S., Nieuwenhuijse, D.F., Molenaar, R.J., Munger, E., Molenkamp, R., et al. Jumping back and forth: Anthropozoonotic and zoonotic transmission of SARS-CoV-2 on mink farms. *bioRxiv* 2020.

18- Oude Munnink, B.B., Sikkema, R.S., Nieuwenhuijse, D.F., Molenaar, R.J., Munger, E., Molenkamp, R., et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science*, **371**, *172-177* (2020)

19- Dyer, O. Covid-19: Denmark to kill 17 million minks over mutation that could undermine vaccine effort. *BMJ*, **371**, *m4338* (2020)

20- Hong Kong seizes hamsters from pet store for mass cull: https://www.bbc.co.uk/news/world-asia-china-60038551 (Accessed 23/02/22)

21- Pickering, B., Lung, O., Maguire, F., Kruczkiewicz, P., Kotwa, J.D., Buchanan, T., et al. Highly divergent white-tailed deer SARS-CoV-2 with potential deer-to-human transmission *bioRxiv* **2022**.02.22.481551 (2022)

22- Kumar, A., Pandey, S.N., Pareek, V., Narayan, R.K., Faig, M.A., Kumari, C. Predicting susceptibility for SARS-CoV-2 infection in domestic and wildlife animals using ACE2 protein sequence homology. *Zoo Biol*. **40**, *79-85* (2020)

23- Senapati, S., Banerjee, P., Bhagavatula, S., Kushwaha, P.P., Kumar, S. Contributions of human ACE2 and TMPRSS2 in determining host–pathogen interaction of COVID-19. *Journal of Genetics*, **100**, *1-16* (2021)

24- Captive gorillas test positive for coronavirus: https://www.sciencemag.org/news/2021/01/captive-gorillas-test-positive-coronavirus (accessed on 01/03/22).

25- Djukic, T.; Stevanovic, G., Coric, V., Bukumiric, Z., Pljesa-Ercegovac, M., Matic, M., et al. GSTO1, GSTO2 and ACE2 polymorphisms modify susceptibility to developing COVID-19. *J. Pers. Med.*, **12**, *458* (2022)

26- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol*., **7**, *539* (2011)

27- Sigrist, C.J.A., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., et al. PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform*., **3**, *265-274* (2002)

28- Vaish, S., Gupta, D., Mehrotra, R., Mehrotra, S., Basantani, M.K. Glutathione *S*-transferase: a versatile protein family. *3 Biotech*., **10**, *321* (2020)

29- Singh, R.R., Reindl, K. M. Glutathione *S*-Transferases in cancer. *Antioxidants (Basel, Switzerland)*, **10**, *701* (2021)

30- Majidi, N., Rabbani, F., Gholami, S., Gholamalizadeh, M., BourBour, F., Rastgoo, S., et al. The effect of Vitamin C on pathological parameters and survival duration of critically ill coronavirus disease 2019 patients: A randomized clinical trial. *Front. Immunol*., **12**, *717816* (2021)

31- Abbas, M., Verma, S., Verma, S., Siddiqui, S., Khan, F.H., Raza, S.T., et al. Association of GSTM1 and GSTT1 gene polymorphisms with COVID‐19 susceptibility and its outcome. *Journal of Medical Virology*, ***93***, *5446-5451* (2021)

32- Why mice are man's best friend in the fight against COVID-19, at: https://sciencenordic.com/animal-testing-epidemic-health/why-mice-are-mans-best-friend-in-the-fight-against-covid-19/1667209 (Accessed 24/03/22)

33- Kousathanas, A., Pairo-Castineira, E., Rawlik, K., Stuckey, A., Odhams, C.A., Walker, S., et al. Whole genome sequencing reveals host factors underlying critical Covid-19. *Nature*, *1-10* (2022)

34 - Fajgenbaum, D.C., June, C.H. Cytokine storm. *New England Journal of Medicine*, **383**, *2255-2273* (2020)

35 - Meftahi, G.H., Bahari, Z., Jangravi, Z., Iman, M. A vicious circle between oxidative stress and cytokine storm in acute respiratory distress syndrome pathogenesis at COVID-19 infection. *Ukrainian Biochemical Journal*, ***93***,*18-29* (2021)

36 - Bilgin, E., Can Demirdöğen, B., Türkanoğlu Özçelik, A., Demirkaya, Ş., Adalı, O. Association analysis of Glutathione *S*-transferase omega-1 and omega-2 genetic polymorphisms and ischemic stroke risk in a Turkish population. *Neurol. Res*. **41**, *118-124* (2019)

37 - Sorokin, A.V., Karathanasis, S.K., Yang, Z.H., Freeman, L., Kotani, K., Remaley, A.T. COVID-19-associated dyslipidemia: Implications for mechanism of impaired resolution and novel therapeutic approaches. *FASEB J*. **34**, *9843-9853* (2020)

38 - Meng, F., Zhang, Y., Liu, F., Guo, X., and Xu, B. Characterization and mutational analysis of omega-class GST (GSTO1) from *Apis cerana cerana*, a gene involved in response to oxidative stress. *PLoS One*, **9**, *e93100* (2014)

39 - Balakrishnan, B., Su, S., Wang, K., Tian, R., Chen, M. Identification, expression, and regulation of an Omega Class glutathione *S*-transferase in *Rhopalosiphum padi* (L.) (Hemiptera: Aphididae) under insecticide stress. *Front Physiol*., **9**, *427* (2018)

40 - Board, P.G. The omega-class glutathione transferases: structure, function, and genetics. *Drug Metabolism Reviews*, **43**, *226-235* (2011)

41 - Paul, S., Jakhar, R., Bhardwaj, M., Kang, S.C. Glutathione-*S*-transferase omega 1 (GSTO1-1) acts as mediator of signaling pathways involved in aflatoxin B1-induced apoptosis–autophagy crosstalk in macrophages. *Free Radical Biology and Medicine*, **89**, *1218-1230* (2015)

42- Nabi, G. Risk of COVID-19 pneumonia in aquatic mammals. *Environ. Res*., **188**, *109732* (2020)

43- Quilliam, R.S., Weidmann, M., Moresco, V., Purshouse, H., O’Hara, Z., Oliver, D.M. COVID-19: The environmental implications of shedding SARS-CoV-2 in human faeces. *Environ. Int*., **140**, *105790* (2020)

44- Maal-Bared, R., Sobsey, M., Bibby, K., Sherchan, S.P., Fitzmorris, K.B., Munakata, N., et al. Pandemic danger to the deep: The risk of marine mammals contracting SARS-CoV-2 from wastewater. *Sci. Total Environ*., **2020**, *144855* (2020)

45- Nova, N. Cross-species transmission of coronaviruses in humans and domestic mammals, what are the ecological mechanisms driving transmission, spillover, and disease emergence? *Frontiers in Public Health*, **9**(2021)

46- Hu, T., Liu, Y., Zhao, M., Zhuang, Q., Xu, L., He, Q. A comparison of COVID-19, SARS and MERS. *PeerJ*, **8**, *e9725* (2020)

47- Stout, A.E., André, N.M., Jaimes, J.A., Licitra, B.N., Whittaker, G.R. One Medicine: a comparative approach to investigating human and animal coronavirus infections. *Journal of Feline Medicine and Surgery*, **23**, *267-268* (2021)

48- Wardeh, M., Baylis, M., Blagrove, M.S. Predicting mammalian hosts in which novel coronaviruses can be generated. *Nature Communications*, **12**, *1-12* (2021)

Figure 1: Clustal Omega alignments GSTO1 across animal species. Sequence details are in Table 1. “\*” indicates sequence identity (“:” indicates a conserved change, whilst “.” Indicates less of a conserved change). Position of SNP indicated in yellow. The predicted (using Prosite and human isoform 1 sequence) N-terminal GST domain is highlighted in green, whilst the N-terminal domain is highlighted in purple. Predicted interesting sites in human isoform 1 are marked (from Prosite: protein kinase C target: grey; tyrosine phosphorylation: brown; amidation site: bright blue; Casein kinase II phosphorylation site: red; cAMP-dependent kinase: light blue: myristoylation;dark green)

CLUSTAL O(1.2.4) multiple sequence alignment

Frog-(tropical-clawed) ---------------------------------MAGPTKSLAKGSPAPGPVPDGVIRAYI 27

Zebrafish ---------------------------------MAASQKCLGKGSPAPGPVPKDHIRLYS 27

Hamster-isoform-X1 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRVYS 28

Hamster-isoform-X7 --------------------------------MSGTSSRSLGKGSSPPGPVPEGLIRVYS 28

Hamster-isoform-X5 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRIYS 28

Hamster-isoform-X2 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRVYS 28

Hamster-isoform-X3 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRIYS 28

Hamster-isoform-X4 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRVYS 28

Hamster-isoform-X6 --------------------------------MSGASSRSLGKGSSPPGPVPEGLIRIYS 28

Whale-(killer) MSGGSARSLGKGEACLPRFAATLRRVSGPGPGGWAAPERSHAGGSAPPGQVPEGLIRVYS 60

Dolphin-(common-bottlenose) MSGGSARSLGKGEACLPRFAATLRRVSRPGPGGWAAPERSHAGGSAPPGQVPEGLIRVYS 60

Cow-(cattle) MSGG--------------------------------SARSLGKGSAPPGPVPEGLIRVYS 28

Bat-(Myotis-myotis) MSGG--------------------------------SDKSLEKGSAPPGPVPEGLIRVYS 28

Bat-(Kuhl's-pipistrelle) MSAG--------------------------------SAKSLDKGSAPPGPVPQGLIRVYS 28

Squirrel MSGE--------------------------------SARSLAKGSAPPGPVPEGLIRVYS 28

Bat-(large-flying-fox) MSGA--------------------------------SARSLGKGSAPPGPVPEGLIRLYS 28

Rhesus-monkey MSRE--------------------------------SARSLGKGSAPPGPVPEGSIRVYS 28

Human-GSTO1-isoform-2 MSGE--------------------------------SARSLGKGSAPPGPVPEGSIRIYS 28

Chimpanzee MSGE--------------------------------SARSLGKGSAPPGPVPEGSIRVYS 28

Human-GSTO1-isoform-1 MSGE--------------------------------SARSLGKGSAPPGPVPEGSIRIYS 28

Human-GSTO1-isoform-3 ------------------------------------------------------------ 0

Gorilla-isoform-X1 MSGE--------------------------------SARSLGKGSAPPGPVPEGSIRVYS 28

Gorilla-isoform-X2 ------------------------------------------------------------ 0

Dog MSAG--------------------------------SARSLAKGSAPPGPVPEGLIRVYS 28

Ferret-(domestic) -MSG--------------------------------SDRSLAKGSAPPGPVPEGLIRVYS 27

Mink-(American) -MSG--------------------------------SARSLAKGSAPPGPVPEGLIRVYS 27

Mouse --------------------------------MSGESSRSLGKGSAPPGPVPEGQIRVYS 28

Rat --------------------------------MSGASARSLGKGSAPPGPVPEGQIRVYS 28

Chicken --------------------------------MAGDHSRSLGKGSAAPGPVPEGVIRLYS 28

Crocodile --------------------------------MAGAMSRSRGKGSAAPGPVPEGLIRLYS 28

Frog-(tropical-clawed) MRFCPFAQRAHLILVAKGIKHELVYINTLNKPDWFFEKSPFGQVPAIETSKGQLIYESQI 87

Zebrafish MRFCPFAQRTRLVLNAKGIKYDTININLKNKPDWFLEKNPLGLVPVLETQSGQVIYESPI 87

Hamster-isoform-X1 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X7 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X5 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X2 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X3 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X4 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Hamster-isoform-X6 MRFCPFAQRALMVLKAKGIRHEIVNINLKNKPEWFLKKNPTGLVPVLEDSKGRLISESVI 88

Whale-(killer) MRFCPFAKRTLLVLKAKGIGHEVININLKNKPEWFFKKNPFGLVPVLENSQGQVIYESAI 120

Dolphin-(common-bottlenose) MRFCPFAKRTLLVLKAKGIGHEVININLKNKPEWFFKKNPFGLVPVLENSQGQVIYESAI 120

Cow-(cattle) MRFCPYAQRTRLVLTAKGIRHEVININLKNKPEWFFKKNPSGLVPVLETSQGQLICESAI 88

Bat-(Myotis-myotis) MRFCPFAQRTRLVLKAKGIRHEVVNINLKNKPEWFFKKNPAGLVPVLENSQGHLICESAI 88

Bat-(Kuhl's-pipistrelle) MRFCPFAQRTLLVLKAKGIRHEVININLKNKPEWFFKKNPAGLVPVLENTQGQLICESAI 88

Squirrel MRFCPFAQRTLLVLKAKGIEHEVININLRNKPEWFFKKNPLGLVPVLENSKGQLIYESTI 88

Bat-(large-flying-fox) MRFCPFAQRTRLVLKAKGIRHEIININLKNKPEWFFKKNPFGLVPVLENSKGQLVCDSAI 88

Rhesus-monkey MRFCPFAERTLLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESPI 88

Human-GSTO1-isoform-2 MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 88

Chimpanzee MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 88

Human-GSTO1-isoform-1 MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 88

Human-GSTO1-isoform-3 MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 60

Gorilla-isoform-X1 MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 88

Gorilla-isoform-X2 MRFCPFAERTRLVLKAKGIRHEVININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESAI 60

Dog MRFCPFAQRTLLVLKAKGIRHEIININLKNKPEWFFKKNPFGLVPVLENSQGQLIYESPI 88

Ferret-(domestic) MRFCPYAQRTLLVLKAKGIRHEIININLKSKPEWFFKKNPFGLVPVLENSQGQLIYESAI 87

Mink-(American) MRFCPYAQRTLLVLTAKGIRHEIININLKSKPEWFFKKNPFGLVPVLENSQGQLIYESAI 87

Mouse MRFCPFAQRTLMVLKAKGIRHEVININLKNKPEWFFEKNPLGLVPVLENSQGHLVTESVI 88

Rat MRFCPFAQRTLMVLKAKGIRHEIININLKNKPEWFFEKNPFGLVPVLENTQGHLITESVI 88

Chicken MRFCPFAQRTRLVLRAKGIRHEVVNINLKNKPDWIFEKNPDGLVPVLETSKGQLIYESPI 88

Crocodile MRFCPFAERTRLVLKAKGINHEIININLKNKPDWFFEKNPFGLVPVLETSKGQLIYESPI 88

\*\*\*\*\*:\*:\*: ::\* \*\*\*\* :: : \*\* .\*\*:\*:::\*.\* \* \*\*.:\* .\*::: :\* \*

Frog-(tropical-clawed) VCDYLDEVFPGKKLTPQDPFQKAQQKMLLEHFSKASTVAFKIVGAKKNNEDISALKAEFL 147

Zebrafish TCEYLDEVYPEKKLLPFDPFERAQQRMLLELFSKVTPYFYKIPVNRTKGEDVSALETELK 147

Hamster-isoform-X1 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPPLVT-RFIRG-NKEDRSALKEELR 146

Hamster-isoform-X7 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPPLVT-RFIRG-NKEDRSALKEELR 146

Hamster-isoform-X5 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPLLVT-RFVRG-KKEDHPALKEELT 146

Hamster-isoform-X2 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPLLVT-RFVRG-KKEDHPALKEELT 146

Hamster-isoform-X3 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPLLVT-RFVRG-KKEDHPALKEELT 146

Hamster-isoform-X4 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPLLVT-RFVRG-KKEDHPALKEELT 146

Hamster-isoform-X6 ICEYLDEAYPDKKLFPDDPYEKARQKMTLESFSKVPLLVT-RFVRG-KKEDHPALKEELT 146

Whale-(killer) TCEYLDEAYPGKKLLPDDPYEKACQKMVFELFSKVPPLLL-SFLRRQNKEDCSGLKEELH 179

Dolphin-(common-bottlenose) TCEYLDEAYPGKKLLPDDPYEKACQKMVFELFSKVPPLLL-SFLRRQNKEDCSGLKEELH 179

Cow-(cattle) TCEYLDEAYPGKKLLPGDPYEKACQKMVLESFSKVPPLIL-KILRTQNKEDCSGLKEELH 147

Bat-(Myotis-myotis) TCEYLDEAYPEKKLLPADPYEKACQKMVFELFSKVPPLVG-RFTRSKNKEDCSDLKEEFR 147

Bat-(Kuhl's-pipistrelle) TCEYLDEAYPEKRLLPADPYEKACQKMVFELFSKVPLLVG-RFTRSKTKEDRSGLKEEFR 147

Squirrel TCEYLDEAYPGKKLLPDDPYEKACQKMIFELFSKVPPLLA-NFIRGQNKNDWPGLKEELK 147

Bat-(large-flying-fox) ICEYLDEAYPGKKLLPEDPYEKACQKMSFELFSKVPSLGG-SYIRNKNKEDCSGIKEELR 147

Rhesus-monkey TCEYLDEAYPGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 147

Human-GSTO1-isoform-2 TCEYLDEAYPGKKLLPDDPYEKACQKMILELFSK-------------------------- 122

Chimpanzee TCEYLDEAYPGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 147

Human-GSTO1-isoform-1 TCEYLDEAYPGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 147

Human-GSTO1-isoform-3 TCEYLDEAYPGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 119

Gorilla-isoform-X1 TCEYLDEAYTGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 147

Gorilla-isoform-X2 TCEYLDEAYTGKKLLPDDPYEKACQKMILELFSKVPSLVG-SFIRSQNKEDYAGLKEEFR 119

Dog TCEYLDEAYPGKKLLPDDPYEKACQKMVFELFSKVPSLVT-GFLRRQNKEDGSGLKEELR 147

Ferret-(domestic) TCEYLDDVYPGKKLLPDDPYEKARQKMVFELFSKVPSLVI-SLLRKQNEEDCSGQKEELR 146

Mink-(American) TCEYLDDVYPGKKLLPDDPYEKARQKMVFELFSKVPSLVT-SLLRKQNEEDCSGQKEELR 146

Mouse TCEYLDEAYPEKKLFPDDPYKKARQKMTLESFSKVPPLIA-SFVRSKRKEDSPNLREALE 147

Rat TCEYLDEAYPEKKLFPDDPYEKACQKMTFELFSKVPSLVT-SFIRAKRKEDHPGIKEELK 147

Chicken TCEYLDEAFPGRKLMPSDPYERALQKMLLEHFSKITSVISK---ALKEGGDLTALTAELA 145

Crocodile TCDYLDEAYPGKKLLPEDPYERAYQKMLLEHFSKLTPLAYKHFMAIQNGEDSTALKAEFS 148

\*:\*\*\*:.: ::\* \* \*\*:::\* \*:\* :\* \*\*\*

Frog-(tropical-clawed) EKLVQFDQVVAKLNTPYVGGSSVSMADYMILPIFERFDIFGVKDCLEKTPHLLQWYQLML 207

Zebrafish DKLSQFNEILLKKKSKFFGGDSITMIDYMMWPWFERLETMNLKHCLDGTPELKKWTERMM 207

Hamster-isoform-X1 KEFYKLEEALTDCQSEFRIGDAVSMTDYLMWPWFQRLEALELNECVAHTPKLKGWMAAMQ 206

Hamster-isoform-X7 KEFYKLEEALTDCQSEFRIGDAVSMTDYLMWPWFQRLEALELNECVAHTPKLKGWMAAMQ 206

Hamster-isoform-X5 KEFYKLEEAVTKYKKDFRVGDAVSMTDYLMWPWFQWLEALELKECVAHTPKLKGWMAAMQ 206

Hamster-isoform-X2 KEFYKLEEAMTKYKKDFRVGDAVSMTDYLMWPWFQWLEALELKECVAHTPKLKGWMAAMQ 206

Hamster-isoform-X3 KEFYKLEEAMTKYKKDFRVGDAVSMTDYLMWPWFQWLEALELKECVAHTPKLKGWMAAMQ 206

Hamster-isoform-X4 KEFYKLEEAMTKYKKDFRVGDAVSMTDYLMWPWFQWLEALELKECLSIRSSTT------- 199

Hamster-isoform-X6 KEFYKLEEAVTKYKKDFRVGDAVSMTDYLMWPWFQWLEALELKECLSIRSSTT------- 199

Whale-(killer) KEFSKLEEVLTNKKTTFFGGSSLSMIDYLIWPWFEWLVALELNEYVNHTPNLKLWMEAMM 239

Dolphin-(common-bottlenose) KEFSKLEEVLTNKKTTFFGGSSLSMIDYLIWPWFEWLVALELNEYVNHTPNLKLWMEAMM 239

Cow-(cattle) KEITKLEEVLTDKKTTFFGGNSLSMIDYLIWPWFERLEALELNECVDHAPTLKLWMAAMK 207

Bat-(Myotis-myotis) QEFSKLEEVLTNKKTTFFGGNSLSMIDYLIWPWFERLEALELNECIDHTPKLKLWMAAMR 207

Bat-(Kuhl's-pipistrelle) QEFSKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEALELNECIDHTPKLKLWVAAMR 207

Squirrel KEFSKLEEVLTNKKTTFFGGNSLSMIDYLIWPWFERLEALELNECVAHTPKLKLWIAAMK 207

Bat-(large-flying-fox) KEFSKLEEVLTNKKTTFFGGNSLSMIDYLIWPWFERLEALELSECVDHTPKLKLWMAAMR 207

Rhesus-monkey KEFTKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLYECVDHTPKLKLWMAAMK 207

Human-GSTO1-isoform-2 --------VLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 174

Chimpanzee KEFTKLEEVLTNKTTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 207

Human-GSTO1-isoform-1 KEFTKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 207

Human-GSTO1-isoform-3 KEFTKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 179

Gorilla-isoform-X1 KEFTKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 207

Gorilla-isoform-X2 KEFTKLEEVLTNKKTTFFGGNSISMIDYLIWPWFERLEAMKLNECVDHTPKLKLWMAAMK 179

Dog KEFSKLEEVLTNKKTTFFGGNSLSMIDYLIWPWFERLEALELNDCVDHTPKLKLWMAAMR 207

Ferret-(domestic) KEISKLEEVLTNMKTTFFGGNSLSMIDYLIWPWFERMEILELNDCVDHTPKLKLWMAAMR 206

Mink-(American) KEISKLEEVLTNKKTTFFGGNSLSMIDYLIWPWFERMEILELNDCVDHTPKLQLWMAAMR 206

Mouse NEFKKLEEGMDNY-KSFLGGDSPSMVDYLTWPWFQRLEALELKECLAHTPKLKLWMAAMQ 206

Rat KEFSKLEEAMAKKRTAFFGGNSLSMIDYLIWPWFQRLEALELNECIDHTPKLKLWMATMQ 207

Chicken EKFGKLDEILSQRNTVFYGGDSTSLIDYMMWPWFERLEAFQLKDVLTRTPKLQRWMEAMR 205

Crocodile EKLGKFEEILANRQSVFFGGDSVSMFDYLIWPWFERLEAFQLYDCLGHTPKLKCWMDAMK 208

: . . : \*.: :: \*\*: \* \*: : : : . :

Frog-(tropical-clawed) QDPAVKATHIKPEALEGFFKLYLQGNPESVDYDI 241

Zebrafish EDPTVKATMFSTETYMVFYKSYMEGNP-NYDYGL 240

Hamster-isoform-X1 KDPTVSSHLIDAKTYRGFVNLYLQDNPEACDYGL 240

Hamster-isoform-X7 KDPTVSSHLIDAKTYRGFVNLYLQDNPEACDYGL 240

Hamster-isoform-X5 KDPTVSSHLIDAKTYRGFVNLYLQDSPEACDYGL 240

Hamster-isoform-X2 KDPTVSSHLIDAKTYRGFVNLYLQDSPEACDYGL 240

Hamster-isoform-X3 KDPTVSSHLIDAKTYRGFVNLYLQDSPEACDYGL 240

Hamster-isoform-X4 ---------------------------------- 199

Hamster-isoform-X6 ---------------------------------- 199

Whale-(killer) KDPAVSSLFIDPKAFRGFLDLYLQNNLEACDYGL 273

Dolphin-(common-bottlenose) KDPAVSSLFIDPKAFRGFLDLYLQNNLEACDYGL 273

Cow-(cattle) KDPTVSSLLTDVKTFQGFFNLYLQNNPEAYDYGL 241

Bat-(Myotis-myotis) KDPTVSALLTDVKTFQGFLNLYLQNSLEACDYGL 241

Bat-(Kuhl's-pipistrelle) KDPTVSALITDGKTFQGFLNLYLQNSVEACDYGL 241

Squirrel EDPPVSALLHDVKTHQGFLGLYLQNSLEAFDYGL 241

Bat-(large-flying-fox) EDPTVSALLTDAKTFRGYLDLYLQNSVEACDYGL 241

Rhesus-monkey EDPTVSALLISGKDWQGFLELYLQNSPEACDYGL 241

Human-GSTO1-isoform-2 EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 208

Chimpanzee EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 241

Human-GSTO1-isoform-1 EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 241

Human-GSTO1-isoform-3 EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 213

Gorilla-isoform-X1 EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 241

Gorilla-isoform-X2 EDPTVSALLTSEKDWQGFLELYLQNSPEACDYGL 213

Dog EDPAVSALLNEANTLRGFLNLYLQNSPEACDYGL 241

Ferret-(domestic) KDPAVSALLMEPKALRGFLNLYLQNSPEACDYGL 240

Mink-(American) KDPAVSALLMEPKALRGFLNLYLQNSPE------ 234

Mouse QDPVASSHKIDAKTYREYLNLYLQDSPEACDYGL 240

Rat EDPVASSHFIDAKTYRDYLSLYLQDSPEACDYGL 241

Chicken KDPAVKDTITDTQTFRSFLQLYFKNSPEACDYGL 239

Crocodile QDPAVKATMTDFQTFKGYLQLYVKNSPEACDYGL 242

Figure 2: Alignments of seven isoforms of human GSTO1 and GSTO2, showing positioning of relevant SNPs. rs4925 in yellow and rs156697 in purple. Conserved residues in that region are highlighted in blue. Sequences from Tables 1 & 2.

CLUSTAL O(1.2.4) multiple sequence alignment

Human-GSTO1-isoform-2 MSGESARSLGKGSAPPGPVPEGSIRIYSMRFCPFAERTRLVLKAKGIRHEVININLKNKP 60

Human-GSTO1-isoform-1 MSGESARSLGKGSAPPGPVPEGSIRIYSMRFCPFAERTRLVLKAKGIRHEVININLKNKP 60

Human-GSTO1-isoform-3 ----------------------------MRFCPFAERTRLVLKAKGIRHEVININLKNKP 32

human-GSTO2-isoform-1 MSGDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

human-GSTO2-isoform-3 ----------------------------MRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 32

human-GSTO2-isoform-2 MSGDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

human-GSTO2-isoform-4 ----------------------------MRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 32

\*\*\*\*\*::.\*\*\*\*\*\*\*\*\*.\*\*\*\*\*:\*\*\*\*:\*\*\*

Human-GSTO1-isoform-2 EWFFKKNPFGLVPVLENSQGQLIYESAITCEYLDEAYPGKKLLPDDPYEKACQKMILELF 120

Human-GSTO1-isoform-1 EWFFKKNPFGLVPVLENSQGQLIYESAITCEYLDEAYPGKKLLPDDPYEKACQKMILELF 120

Human-GSTO1-isoform-3 EWFFKKNPFGLVPVLENSQGQLIYESAITCEYLDEAYPGKKLLPDDPYEKACQKMILELF 92

human-GSTO2-isoform-1 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 120

human-GSTO2-isoform-3 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 92

human-GSTO2-isoform-2 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 120

human-GSTO2-isoform-4 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 92

\*\*::.\*:\*\*\* :\*\*\*\*.\*\* \*\*\*\*\*\*.\*:\*\*\*\*\*:\*\*\*\*:\*\*:\* \*\*\*\*:\* \*\*\*:\*\*\*\*

Human-GSTO1-isoform-2 SK----------------------------------VLTNKKTTFFGGNSISMIDYLIWP 146

Human-GSTO1-isoform-1 SKVPSLVGSFIR-SQNKEDYAGLKEEFRKEFTKLEEVLTNKKTTFFGGNSISMIDYLIWP 179

Human-GSTO1-isoform-3 SKVPSLVGSFIR-SQNKEDYAGLKEEFRKEFTKLEEVLTNKKTTFFGGNSISMIDYLIWP 151

human-GSTO2-isoform-1 CKVPHLTKECLVALRCGRECTNLKAALRQEFSNLEEILEYQNTTFFGGTCISMIDYLLWP 180

human-GSTO2-isoform-3 CKVPHLTKECLVALRCGRECTNLKAALRQEFSNLEEILEYQNTTFFGGTCISMIDYLLWP 152

human-GSTO2-isoform-2 CK----------------------------------ILEYQNTTFFGGTCISMIDYLLWP 146

human-GSTO2-isoform-4 CK----------------------------------ILEYQNTTFFGGTCISMIDYLLWP 118

.\* :\* ::\*\*\*\*\*\*..\*\*\*\*\*\*\*:\*\*

Human-GSTO1-isoform-2 WFERLEAMKLNECVDHTPKLKLWMAAMKEDPTVSALLTSEKDWQGFLELYLQNSPEACDY 206

Human-GSTO1-isoform-1 WFERLEAMKLNECVDHTPKLKLWMAAMKEDPTVSALLTSEKDWQGFLELYLQNSPEACDY 239

Human-GSTO1-isoform-3 WFERLEAMKLNECVDHTPKLKLWMAAMKEDPTVSALLTSEKDWQGFLELYLQNSPEACDY 211

human-GSTO2-isoform-1 WFERLDVYGILDCVSHTPALRLWISAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDF 240

human-GSTO2-isoform-3 WFERLDVYGILDCVSHTPALRLWISAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDF 212

human-GSTO2-isoform-2 WFERLDVYGILDCVSHTPALRLWISAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDF 206

human-GSTO2-isoform-4 WFERLDVYGILDCVSHTPALRLWISAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDF 178

\*\*\*\*\*:. : :\*\*.\*\*\* \*:\*\*::\*\*\* \*\*\*\*.\*\*\* .:. :\*\*\*\*:\*\*:\*\*.\*:\* \*:

Human-GSTO1-isoform-2 GL- 208

Human-GSTO1-isoform-1 GL- 241

Human-GSTO1-isoform-3 GL- 213

human-GSTO2-isoform-1 GLC 243

human-GSTO2-isoform-3 GLC 215

human-GSTO2-isoform-2 GLC 209

human-GSTO2-isoform-4 GLC 181

\*\*

Figure 3: Clustal Omega alignments GSTO2. Sequence details are in Table 2. “\*” indicates sequence identity (“:” indicates a conserved change, whilst “.” indicates less of a conserved change). Position of SNP indicated in yellow. The predicted (using Prosite) N-terminal GST domain (of human GST isoform 1) is highlighted in green, whilst the N-terminal domain is highlighted in purple. Predicted phosphorylation sites are marked (from Prosite: protein kinase C: grey; tyrosine phosphorylation: brown; amidation site: bright blue; Casein kinase II phosphorylation site: red; N-glycosylation site: teal; myristoylation;dark green).

CLUSTAL O(1.2.4) multiple sequence alignment

Frog -MTGSEKSLAKGSPAPGPVSEETIRVYSMRFCPYAQRARLVLAAKGIKHEVININLKNKP 59

Hamster-isoform-X2 ----------------------------MEI-----DVWGCDQKLGERYEVININLKNKP 27

Hamster-isoform-X1 MSGDATRSLGRGSSPPGPVPEGVIRIYSMRFCPYSHRARLVLKAKGIRYEVININLKNKP 60

Hamster-isoform-X3 MSGDATRSLGRGSSPPGPVPEGVIRIYSMRFCPYSHRARLVLKAKGIRYEVININLKNKP 60

Bat-(big-brown) MGEDASRSLGKGSVPPGPVPEGLIRIYSMRFCPFAHRTRLVLLAKGISHEVININLRNKP 60

Bat-(greater-spear-nosed) MAEDASRALGKGSHPPGPVPEGLIRIYSMRFCPYAHRTRLVLQAKGISHEVININLRNKP 60

Bat-(common-vampire) MAEDASRALGKGSHPPGPVPEGLIRIYSMRFCPYAHRTRLVLQAKGIRHEVININLRNKP 60

Cow MTDDATRTLGKGSIPPGPVPEGVIRLYSMRFCPYAHRTRLVLRAKGIRHEVININLRNKP 60

Whale MTDDAARTLGRGSAPPGPVPEGLIRLYSMRFCPYAHRTRLVLQAKGIRHEVININLRNKP 60

Dolphin MTDDAARTLGRGSAPPGPVPEGLIRLYSMRFCPYAHRTRLVLQAKGIRHEVININLRNKP 60

Dog-isoform-X1 -MEDASRTFGKGSLPPGPVPEGLIRIYSMRFCPYAHRTRLVLRAKGIRHEVININLRNKP 59

Dog-isoform-X2 -MEDASRTFGKGSLPPGPVPEGLIRIYSMRFCPYAHRTRLVLRAKGIRHEVININLRNKP 59

Ferret ------------------------------------------------HEVVNINLRNKP 12

Cat-isoform-X1 -MEDATRTFGKGSCPPGPVPEGLIRVYSMRFCPFAHRTRLVLRAKGIRHEVININLRNKP 59

Cat-isoform-X2 -MEDATRTFGKGSCPPGPVPEGLIRVYSMRFCPFAHRTRLVLRAKGIRHEVININLRNKP 59

Cat-isoform-X3 -MEDATRTFGKGSCPPGPVPEGLIRVYSMRFCPFAHRTRLVLRAKGIRHEVININLRNKP 59

Rhesus-monkey MSQDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

Human-GSTO2-isoform-2 MSGDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

Human-GSTO2-isoform-4 ----------------------------MRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 32

Gorilla MSEDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

Human-GSTO2-isoform-1 MSGDATRTLGKGSQPPGPVPEGLIRIYSMRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 60

Human-GSTO2-isoform-3 ----------------------------MRFCPYSHRTRLVLKAKDIRHEVVNINLRNKP 32

Mouse MSGDLSRCLGKGSCPPGPVPEGVIRIYSMRFCPYSHRARLVLKAKGIRHEVININLKSKP 60

Rat MSGDLTRCLGKGSCPPGPVPEGVIRIYSMRFCPYSHRTRLVLKAKSIRHEIININLKNKP 60

:\*::\*\*\*\*:.\*\*

Frog DWFIEKSPFGLVPSLETSSGQVIYESPIVCDYLDEVYPGKKLTPVDPFQKAQQKMIVEHF 119

Hamster-isoform-X2 EWYFTKHPFGQIPVLENSQGQLIYESVIACEYLDDLYPGRKLFPFDPYERARQKMLLELF 87

Hamster-isoform-X1 EWYFTKHPFGQIPVLENSQGQLIYESVIACEYLDDLYPGRKLFPFDPYERARQKMLLELF 120

Hamster-isoform-X3 EWYFTKHPFGQIPVLENSQGQLIYESVIACEYLDDLYPGRKLFPFDPYERARQKMLLELF 120

Bat-(big-brown) DWYYTKHPFGQIPVLENSRCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 120

Bat-(greater-spear-nosed) DWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPFDPYERARQKMLLELF 120

Bat-(common-vampire) DWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPFDPYERARQKMLLELF 120

Cow EWYFTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 120

Whale EWYFTKHPFGKIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPCDSYERARQKMLLDLF 120

Dolphin EWYFTKHPFGKIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPCDSYERARQKMLLDLF 120

Dog-isoform-X1 EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 119

Dog-isoform-X2 EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 119

Ferret EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 72

Cat-isoform-X1 EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 119

Cat-isoform-X2 EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 119

Cat-isoform-X3 EWYYTKHPFGQIPVLENSKCQLIYESVIACEYLDDAYPGRKLYPYDPYERARQKMLLELF 119

Rhesus-monkey EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPHDPYERARQKMLLELF 120

Human-GSTO2-isoform-2 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 120

Human-GSTO2-isoform-4 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 92

Gorilla EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 120

Human-GSTO2-isoform-1 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 120

Human-GSTO2-isoform-3 EWYYTKHPFGHIPVLETSQCQLIYESVIACEYLDDAYPGRKLFPYDPYERARQKMLLELF 92

Mouse DWYYTKHPFGQIPVLENSQCQLVYESVIACEYLDDVYPGRKLFPYDPYERARQKMLLELF 120

Rat DWYYTKHPFGQVPVLENSQCQLIYESVIACEYLDDVFPGRKLFPYDPYERARQKMLLELF 120

:\*: \* \*\*\* :\* \*\*.\* \*::\*\*\* \*.\*:\*\*\*: :\*\*:\*\* \* \* :::\*:\*\*\*::: \*

Frog SKISTLFYKILLAKKNNEDVSGVKAEVQEKLVKLDE------------------------ 155

Hamster-isoform-X2 CKVPHLAKECLVALRCGRECMDLKTALRQEFRNLEE------------------------ 123

Hamster-isoform-X1 CKVPHLAKECLVALRCGRECMDLKTALRQEFRNLEE------------------------ 156

Hamster-isoform-X3 CKV--------------------------------------------------------- 123

Bat-(big-brown) YKVPHLTKECLVALRCGKECCDLKLALREEFCNLEE------------------------ 156

Bat-(greater-spear-nosed) YKIPHLTKECLVALRCGKECANLKLALREEFCNLEE------------------------ 156

Bat-(common-vampire) YKIPHLTKECLVALRCGKECGNLKLALREEFCNLEE------------------------ 156

Cow YKVPHLTKECLVALRCGRDCGDLKLALRQEFCNLEE------------------------ 156

Whale YKVPHLTKECLIASRCGRECADLKLALRQEFCNLEE------------------------ 156

Dolphin YKVPHLTKECLIASRCGRECADLKLALRQEFCNLEE------------------------ 156

Dog-isoform-X1 YKVPHLTKECLVALRCGRECTDLKLALRQEFCNLEE------------------------ 155

Dog-isoform-X2 YK---------------------------------------------------------- 121

Ferret YKVPHLTKECLVALRCGRECADLKLALRQEFCNLEE------------------------ 108

Cat-isoform-X1 YKVPHLTKECLVAIRCGRECADLKLALRQEFCNLEEVFPVVISSTAFTNRKSSTPSGEME 179

Cat-isoform-X2 YKVPHLTKECLVAIRCGRECADLKLALRQEFCNLEE------------------------ 155

Cat-isoform-X3 YK---------------------------------------------------------- 121

Rhesus-monkey CKVPHLTKECLVALRCGRECTDLKAALRQEFCNLEE------------------------ 156

Human-GSTO2-isoform-2 CK---------------------------------------------------------- 122

Human-GSTO2-isoform-4 CK---------------------------------------------------------- 94

Gorilla CKVPDLTKECLVALRCGRECTDLKAALRQEFGNLEE------------------------ 156

Human-GSTO2-isoform-1 CKVPHLTKECLVALRCGRECTNLKAALRQEFSNLEE------------------------ 156

Human-GSTO2-isoform-3 CKVPHLTKECLVALRCGRECTNLKAALRQEFSNLEE------------------------ 128

Mouse CKVPPLSKECLIALRCGRDCTDLKVALRQELCNMEE------------------------ 156

Rat CKVPQLSKECLVALRCGRDCTDLKVALRQELCNLEE------------------------ 156

\*

Frog ------------ILAKQNGLFFGGSDVSMVDYMIWPWFERLIIFDSKDCLNKTPHIDKWY 203

Hamster-isoform-X2 ------------VLEYQNITFFGGDRISMIDYLFWPWFERLDVYGLSDCVSHTPMLRLWI 171

Hamster-isoform-X1 ------------VLEYQNITFFGGDRISMIDYLFWPWFERLDVYGLSDCVSHTPMLRLWI 204

Hamster-isoform-X3 -------------LEYQNITFFGGDRISMIDYLFWPWFERLDVYGLSDCVSHTPMLRLWI 170

Bat-(big-brown) ------------ILSYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCVNHTPALRLWI 204

Bat-(greater-spear-nosed) ------------ILSYQNTVFFGGPCISMIDYLFWPWFERLDVYGIADCVNHTPALRLWI 204

Bat-(common-vampire) ------------ILSYQNTVFFGGPCISMIDYLFWPWFERLDVYGIADCVNHTPALRLWI 204

Cow ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLEVYGIADCVNHTPALRLWI 204

Whale ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCVNHTPALRLWI 204

Dolphin ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCVNHTPALRLWI 204

Dog-isoform-X1 ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLEVYGIADCLNHTPALRLWT 203

Dog-isoform-X2 ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLEVYGIADCLNHTPALRLWT 169

Ferret ------------ILGYQNTIFFGGDCISMIDYLFWPWFERLDVYGIADCLNHTPALRLWT 156

Cat-isoform-X1 NIFFHPRSITLKILGYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCLNHTPALRLWT 239

Cat-isoform-X2 ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCLNHTPALRLWT 203

Cat-isoform-X3 ------------ILGYQNTVFFGGDCISMIDYLFWPWFERLDVYGIADCLNHTPALRLWT 169

Rhesus-monkey ------------ILEYQNTTFFGGTCTSMIDYLLWPWFERLDVYGIADCVSHTPALRLWI 204

Human-GSTO2-isoform-2 ------------ILEYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPALRLWI 170

Human-GSTO2-isoform-4 ------------ILEYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPALRLWI 142

Gorilla ------------ILEYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPALRLWI 204

Human-GSTO2-isoform-1 ------------ILEYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPALRLWI 204

Human-GSTO2-isoform-3 ------------ILEYQNTTFFGGTCISMIDYLLWPWFERLDVYGILDCVSHTPALRLWI 176

Mouse ------------ILEYQNTTFFGGDCISMIDYLVWPWFERLDVYGLADCVNHTPMLRLWI 204

Rat ------------ILEYQNTTFFGGDSISMIDYLVWPWFERLDVYGLADCVNHTPMLRLWI 204

\* \*\* \*\*\*\* \*\*:\*\*:.\*\*\*\*\*\*\* ::. \*\*:.:\*\* : \*

Frog QQMLQDPAVKATYIEPDLLLGFFKLYSQNDVEACDYGL------ 241

Hamster-isoform-X2 STMKQDPTVCALLTDKNIFLGFLHLYFQNNPCAFDFGLCAPAIR 215

Hamster-isoform-X1 STMKQDPTVCALLTDKNIFLGFLHLYFQNNPCAFDFGLCAPAIR 248

Hamster-isoform-X3 STMKQDPTVCALLTDKNIFLGFLHLYFQNNPCAFDFGLCAPAIR 214

Bat-(big-brown) EAMKQDPTVCALLIDKNIFLGFLNLYFQNHPDAFDYGLSC---- 244

Bat-(greater-spear-nosed) AAMKRDPTVCALLIDKNIFLGFLNLYFQNHPEAFDYGLSC---- 244

Bat-(common-vampire) AAMKRDPTVCSLLIDKNIFLGFLNLYFQNHPEAFDYGLSC---- 244

Cow AAMKQDPTVCSLLTDKNTFLGFLNLYFQNNPGAFDYGLSC---- 244

Whale AAMKQDPTVCALLIDKNIFLGFLNLYFQNNPDAFDYGLSC---- 244

Dolphin AAMKQDPTVCALLIDKNIFLGFLNLYFQNNPDAFDYGLSC---- 244

Dog-isoform-X1 AAMKQDPTVCALLIDKSVFSGFLNLYFQNNPDAFDYGLIC---- 243

Dog-isoform-X2 AAMKQDPTVCALLIDKSVFSGFLNLYFQNNPDAFDYGLIC---- 209

Ferret AAMKQDPTVCALLIDKNIFLGFLNLYFQNNPDAFDYGLA----- 195

Cat-isoform-X1 AAMKQDPTVCALLIDRSIFLGFLNLYFQNNPDAFDYGLTC---- 279

Cat-isoform-X2 AAMKQDPTVCALLIDRSIFLGFLNLYFQNNPDAFDYGLTC---- 243

Cat-isoform-X3 AAMKQDPTVCALLIDRSIFLGFLNLYFQNNPDAFDYGLTC---- 209

Rhesus-monkey SAMKWDPTVCALLTDKSIFQGFLNLYFQNNPNAFDFGLC----- 243

Human-GSTO2-isoform-2 SAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDFGLC----- 209

Human-GSTO2-isoform-4 SAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDFGLC----- 181

Gorilla SAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDFGLC----- 243

Human-GSTO2-isoform-1 SAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDFGLC----- 243

Human-GSTO2-isoform-3 SAMKWDPTVCALLMDKSIFQGFLNLYFQNNPNAFDFGLC----- 215

Mouse ASMKQDPAVCALHTDKSVFLGFLNLYFQNNPCAFDFGLCNPIIR 248

Rat SSMKQDPAVCALHIDKNIFLGFLNLYFQNNPCAFDFGLCGPIVR 248

\* \*\*:\* : : . : \*\*::\*\* \*\*. \* \*:\*\*