

# Respiratory Analysis of COVID-19 Severity-Related Genes Reveals Decreased Respiration in *Drosophila*

Fopofoluwa Ayangbayi, Alison Belonzi, Gianna-Rose DiDonato, Bryce Showell

Department of Biology, Rutgers The State University of New Jersey, Camden, N.J. 08102

## Abstract

The virus SARS-CoV-2 causes COVID-19—a disease that affects the respiratory system. It is highly contagious and transmitted via cough droplets (Jayaweera et al., 2020). To determine why some individuals are affected more severely by the disease, we examined a recent Genome Wide Association study (GWAS) which identified six genes associated with respiratory failure in COVID-19: SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1. It is not known whether the genes identified in the GWAS affect rates of respiration. An additional study identified a gene, TRL7, also associated with COVID-19. Since respiratory issues are a main symptom of COVID-19, we tested the effects of mutations in these genes on respiration. To do this, we used respirometers and *Drosophila*—which had homologues for two of the genes identified disrupted with transposons and wild type *Drosophila*—to assess whether genetic modifications would cause a decrease in respiration. Of the genes tested, we found that disruption of SLC6A20, which encodes a proline transporter and is also expressed in lung tissue, caused a decrease in respiration. The role of proline oxidation and the generation of reactive oxygen species (ROS) in the respiratory system could explain these results. The results could also explain why this gene was associated with respiratory failure. It is possible that genes like SLC6A20, which affect respiratory function, affect the severity of COVID-19 symptoms. Knowing the genes that put individuals at greater risk of more severe symptoms is critical to their health. By obtaining a better understanding of the genes and their functions, individuals could be better informed of their increased susceptibility to COVID-19.

## Introduction

Originating in Wuhan, China at the end of 2019, SARS-CoV-2 was discovered to be a novel coronavirus causing the disease known as COVID-19 (Mackenzie and Smith, 2020). Currently, risk factors for the disease are being determined, and they shed light on why some individuals die while others remain asymptomatic. There have been 16 genes identified in a

GWAS study as being associated with COVID-19 severity, however there is still no information on why these genes may increase risk (Ellinghaus et al, 2020). Of the 16 genes identified, six were on the same locus. Given that COVID-19 is a respiratory illness, we used *Drosophila* with homologs for those genes that were knocked out and tested their respiratory function.

Previous coronavirus outbreaks include middle east respiratory syndrome (MERS) in 2012 and severe acute respiratory syndrome (SARS) in 2003. The most common symptoms associated with COVID-19 are shortness of breath, fever, cough, fatigue, and loss of taste and smell. Possible complications of COVID-19 include pneumonia and acute respiratory distress syndrome (ARDS), both of which can be life threatening (Baj et al., 2020). Symptoms can range from severe to very mild, with the most severe complication involving respiratory failure (Dreher et al., 2020). Long term effects of COVID-19 are still unknown due to the relatively short amount of time that has passed since COVID began affecting humans. As of December 2020, nearly 15.3 million Americans have gotten sick, and almost 300,000 of them have died from COVID-19 (Center for Disease Control, 2020). For the other 15 million, symptoms ranged from severe to non-existent.

Some general risk factors have already been determined, such as immune function, age, sex, and chronic conditions such as diabetes and hypertension. In general, older males and those with pre-existing health conditions may have more severe symptoms (Ellinghaus et al, 2020). However, many young and healthy individuals are becoming critically ill, and it is currently unknown what causes one healthy individual to become very sick, while others exhibit only mild symptoms.

To get a more accurate understanding of who is more severely affected by this virus, genetic tests were conducted to determine if specific gene mutations play a role in the severity of the disease. A genome wide association study was conducted and found that age and sex do play roles in the severity of COVID-19 symptoms (Ellinghaus et al, 2020). Two

loci were identified, 3p21.31 and 9q34.2. Locus 9q34.2 is associated with blood groups, whereas there are multiple genes associated with the 3p21.31 locus. Six genes found at this location include: SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1. Genes SLC6A20 and LZTFL1 are both strongly expressed in lung cells. SLC6A20 encodes for SIT1, a protein which interacts with ACE-2, the receptor to which SARS-CoV-2 binds (Ellinghaus et al., 2020). In addition, CCR9 and CXCR6 encode chemokine receptors with CXCR6 acting in the lungs (Wein et al., 2019). Another genetic study identified genetic variants in the gene TLR7 which were associated with severe COVID-19, particularly in men, due to the presence of the gene on the X chromosome (van der Made et al., 2020). Due to their correlation with respiratory failure, it is likely that these genes play a role in the severity of COVID-19 symptoms.

Since we are unaware if these mutations affect respiration rates in humans, the effects of these gene mutations can be studied in the lab using *Drosophila*. Two of these human genes (TRL7 and SLC6A20) have *Drosophila* homologues—as often is the case with human genes—which makes *Drosophila* a useful model organism for studying human diseases (Pandey and Nichols, 2011). The gene disruption project was used to select *Drosophila* with a disruption in each of the identified genes. A transposon is inserted to disrupt the function of the gene, which will be tested against Oregon-R *Drosophila*. The simplicity of their genome—which contains 4 chromosomes—also makes them excellent model organisms. Respiration rates can be studied using *Drosophila* in which these select genes are interrupted. The respiration rates in various genotypes of *Drosophila* will be tested. We hypothesize that these mutated genes reduce respiration rates making individuals with these genetic mutations experience worsening respiratory health with COVID-19. If the hypothesis is true, *Drosophila* with mutations in the genes associated with COVID-19 severity would also have lower than average rates of respiration when compared to wild type flies.

## Materials & Methods

### *Respirometer and Measurement Chamber Assembly*

We created a modified version of a respirometer developed by Yatsenko et al in 2014 (Yatsenko et al., 2014). The chamber used for this experiment consists of two rectangular plastic containers which create a clam shell when placed together. Vaseline is used to create an airtight seal between the containers. For the respirometers, a 1 ml pipette tip is cut to allow the insertion of the 50 µl capillary tube and then a piece of cotton is then placed and pushed down the pipette into the tip. Next, a small amount of soda lime as a CO<sub>2</sub> absorbent is added on top of the cotton and then sealed with another piece of cotton. Glue is applied where the capillary tube is inserted into the pipette tip and this must be left to dry overnight. The solution is prepared with 25 drops of food coloring per 200 mL of water. The chamber is then labeled with a centimeter scale to measure how far the dye travels.

Issues with individual respirometer construction may result in data that is much lower than expected, signifying a nonfunctioning respirometer. These could be a result of too much or too little space given for *Drosophila* in the tubes from packing the cotton, the glue not sealed completely on the pipette, or clay not sealed properly on the tubes. In addition, parafilm was found not to be effective due to its oxygen permeability. Furthermore, female *Drosophila* are known to disrupt the data due to their variability in CO<sub>2</sub> production when reproducing (Yatsenko et al., 2014). In addition, there also appear to be disruptions in respiration when 5 or more *Drosophila* are placed into one respirometer and unreliable data when 2 or less are placed into one respirometer (Fig. 1).

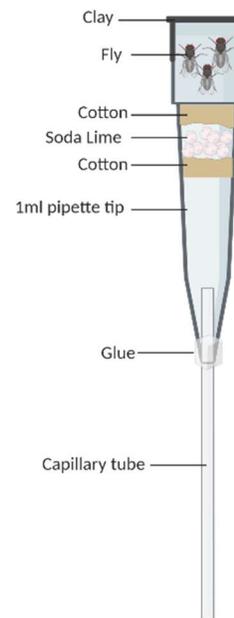


Figure 1. Diagram of respirometer which includes clay placed on the top of the 1ml pipette tube, 3-4 flies placed above a layer of soda lime which is held in place by two pieces of cotton, and a capillary tube inserted into the 1ml pipette tube and held in place by glue. The drawing was made by the tools in BioRender.com.

### *Analysis of Results*

Image J can be used to analyze the images by using the pixels as a scale to measure the distance the liquid travels. The difference in distances ( $\Delta d$ ) between the liquids at their initial levels and at the end of the experiment should be taken. With this, the amount of CO<sub>2</sub> (µl/hr/fly) produced can be determined by the following equation:  $\frac{((\pi \times R^2) \times (\Delta d) - (\Delta c)) \times 1000}{n \times h}$  where R = the radius of the capillary tube in centimeters,  $\Delta c$  = the distance the liquid has moved up the capillary tube of the negative control sample, n = the number of flies used and h = hours (Yatsenko et al., 2014). The same equation (without the denominator) can be used to calculate the total volume produced in an individual respirometer.

### Handling *Drosophila*

Five-day old male *Drosophila* are used in this experiment. *Drosophila* are anesthetized using ice packs then separated by sex based upon their size and abdomen (Pulver & Berni, 2012). For all experiments, six tubes are used. Five tubes contain three to four *Drosophila* each, and the last tube is left empty as a negative control to adjust for any baseline movement of the dye up the respirometer (c).

The *Drosophila* used in the present study were purchased from Bloomington *Drosophila* Stock Center. On the Bloomington site, a *Drosophila* homologue can be searched for and an accompanying stock ID is given which is unique for the type of modification to a given gene. For this experiment, we have inserted transposons into the homologues we wish to study. We were able to test four different homologues for which transposons were inserted. These homologues include: Toll-6, two modifications of TL and CG43055 for which the stock IDs are 34467, 30652, 36134 and 33157, respectively. Both Toll-6 and TL are *Drosophila* homologues for the gene TRL7 and CG43055 is the *Drosophila* homologue for the gene SLC6A20.

### Respirometry Control

To assess the reliability of the respirometry assay, a control test is run using wild type Oregon-R *Drosophila*. Six respirometers are tested with increasing numbers of *Drosophila*. In order, the respirometers contained zero, two, three, four, and five *Drosophila*. The respirometer with no *Drosophila* is the atmospheric control or the negative control (Yatsenko et al., 2014). The respirometers are placed in the chamber and held up by reusing the interior packaging from 1 microliter pipette tips. The respirometers are inserted with the capillary tube tip down into the chamber so that the tip submerges into the dye solution, but so the capillary tube is not occluded on the bottom. Vaseline is added between the lid cover and the chamber to provide stronger isolation from temperature and pressure fluctuations. The lid is closed, and the system can equilibrate for 15 min. Two photographs are taken at the start and end of each experiment: one is taken of the chamber at the start of the experiment to make sure that the level of liquid within each capillary tube is visible and another is taken after 60 minutes. Both photographs show a ruler placed next to the chambers for measurement analysis with ImageJ. When each experiment is finished, the *Drosophila* are discarded.

## Results

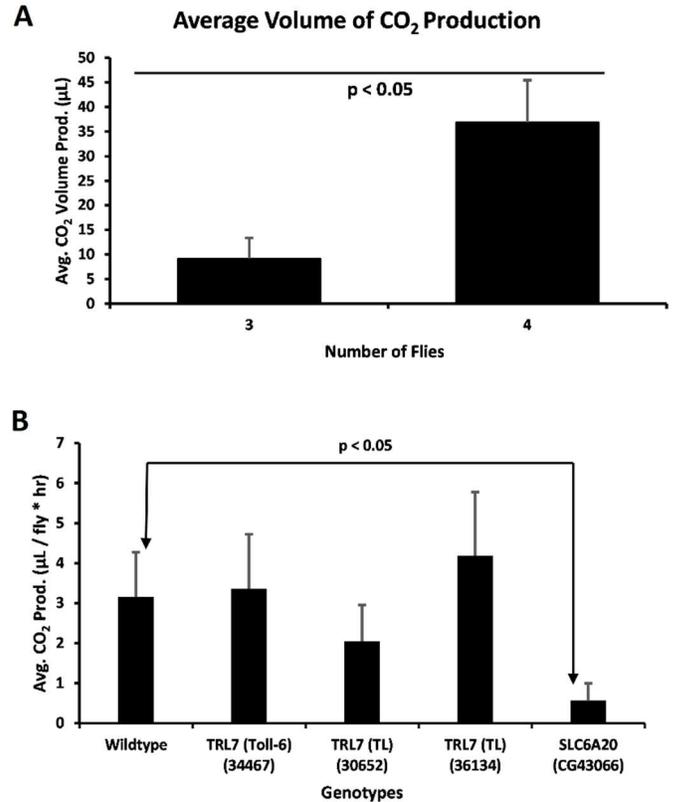


Figure 2. CO<sub>2</sub> production by flies. A) Average volume of CO<sub>2</sub> production in a vial by number of *Drosophila* in a vial (For 3 *Drosophila*, n = 4; for 4 *Drosophila*, n = 2). A two-tailed t-test was performed to compare the average volumes for 3 and 4 *Drosophila* (p < 0.05). Error bars represent the standard error of the mean. B) Average amount of CO<sub>2</sub> produced per fly per hour for wild type (Oregon-R) *Drosophila* compared to mutants (n = 7, Oregon-R; n = 4, Toll-6; n = 4, TL; n = 3, TL; n = 2, SLC6A20). A two-tailed t-test was performed to compare the WT to SLC6A20 (p < 0.05). Error bars represent the standard error of the mean.

To confirm that the respirometers were reliable, we ran an assay containing increasing numbers of wild type *Drosophila*. Ideally, the total volume of CO<sub>2</sub> should increase with the number of *Drosophila*, indicating a functioning respirometer (Fig. 2A). The assays were run with different trials to calculate an average value. As previously stated, we found that when 5 or more *Drosophila* were placed in one respirometer, it resulted in malfunctioning respirometers and sometimes dead *Drosophila* (data not reported). In addition, if 2 or less were placed in one respirometer the data was unreliable. Therefore, 3-5 *Drosophila* is the best amount for conducting these experiments. The assay was then confirmed with 3 and 4 *Drosophila* tested, for which the values in µL were 9.18 and 36.9 respectively (Fig. 2A). These average values were

significantly different from one another, confirming that the respirometers are functional (Fig. 2A).

Once the respirometers were tested and functioning, we ran the assay using wild type Oregon-R *Drosophila*, and *Drosophila* with the genes Toll-6, TL (30652), TL (36134), and CG43066 disrupted. Both TL and Toll-6 are homologues for the human gene TRL7 while CG43066 is a homologue for the gene SLC6A20. Since there was variability in the data, we ruled out the outliers in the data by determining the percentage difference for each data point. If a data point was either greater than 60% or lower than 40% compared to the average, it was considered an outlier. Average rates of respiration in  $\mu\text{l}$  per fly per hour for Oregon-R, Toll-6, TL (30652), TL (36134) and CG43066 were 3.15, 3.36, 2.04, 4.18, and 0.56 respectively (Fig. 2B). First a one-way ANOVA was performed to compare the differences between groups. There was found to be a *p* value of 0.029 for the difference between all the groups. Next, a two-tailed *t*-test revealed that the wild type and CG43066 were significantly different from each other. (Fig. 2B).

## Discussion

Our results show a significant decrease in average respiration between the wild type, Oregon-R, and the *Drosophila* mutant for CG43066, which was disrupted with the insertion of a transposon.

Issues with individual respirometer construction resulted in data that was much lower than intended, signifying a nonfunctioning respirometer. This could have been a result of too much or too little space given for *Drosophila* in the tubes from packing the cotton, the glue not sealed completely on the pipette, or clay not sealed properly on the tubes. In addition, parafilm was found not to be effective due to its oxygen permeability, and female *Drosophila* are known to disrupt the data due to their variability in CO<sub>2</sub> production when reproducing (Yatsenko et al., 2014). In addition, there also appear to be disruptions in respiration when 5 or more *Drosophila* are placed into one respirometer and unreliable data when 2 or less are placed into one respirometer that we constructed (Fig. 1).

SLC6A20 (solute carrier family 6, member 20) is known to encode the protein SIT1 which, as previously stated, interacts with ACE2. SLC6A20 is expressed mainly in the kidney and the small intestine and its protein product functions as a proline transporter. It is also expressed in low levels in other tissues such as the lungs (Muhanna et al., 2020). The role of proline oxidation in the electron transport chain (ETC) could potentially explain the lower levels of respiration found in *Drosophila*. Proline oxidase is an enzyme that functions in the inner mitochondrial membrane which oxidizes proline and allows eventual breakdown into alpha ketoglutarate, a molecule of the citric acid cycle (Pandhare et al., 2009). However, the presence of proline leads to decreased ROS within cells, so if the gene responsible for proline transportation is disrupted, this would likely lead to a decrease in ROS production due to excess proline available (Krishnan

et al., 2008). In the respiratory system, increased ROS production is linked to several respiratory diseases such as COPD and asthma (Boukhenouna et al., 2018). Due to this mechanism, it is unclear why there is a decrease in respiration rates with the disruption in the gene responsible for proline transportation. More studies would also need to be done to confirm if the results in *Drosophila* are the same as in humans.

For the data collected from CG43066, *n*=2, and given the low sample size, further research should be done to ensure more *Drosophila* are tested to achieve a more stable average. If a new average were obtained, it could be further analyzed statistically to determine if any significant differences remain. In addition, the transposons used to disrupt the genes may have a range of effects on gene expression. The transposon may completely disrupt a gene, partially disrupt a gene, or may even cause the gene to malfunction.

Overall, we were able to find a gene that does cause a decrease in respiration when that gene is disrupted as we expected. Due to limited time and resources, more tests should be done on SLC6A20 and other genes to determine if there are more genes tied to respiration. Using *Drosophila* was helpful and effective when testing the connection between genetics and respiration and would be recommended for future studies. This study has the potential in helping present and future COVID-19 patients with being able to provide them another way for determining their risk for contracting severe COVID-19.

## Acknowledgements

We would like to thank the Biology department at Rutgers University – Camden for resources and funding. We would also like to thank Dr. Lee and Dr. Fried for guidance throughout this entire research project.

## References:

- Boukhenouna, S., Wilson, M.A., Bahmed, K., and Kosmider, B. (2018). Reactive Oxygen Species in Chronic Obstructive Pulmonary Disease. *Oxid. Med. Cell. Longev.* 2018.
- Center for Disease Control. (2020, November 16). CDC COVID Data Tracker. [https://covid.cdc.gov/covid-data-tracker/#cases\\_casesinlast7days](https://covid.cdc.gov/covid-data-tracker/#cases_casesinlast7days)
- DeBiasi, R.L., Song, X., Delaney, M., Bell, M., Smith, K., Pershad, J., Ansusinha, E., Hahn, A., Hamdy, R., Harik, N., et al. (2020). Severe Coronavirus Disease-2019 in Children and Young Adults in the Washington, DC, Metropolitan Region. *J. Pediatr.* 223, 199-203.e1.
- Dreher, M., Kersten, A., Bickenbach, J., Balfanz, P., Hartmann, B., Cornelissen, C., Daher, A., Stoehr, R., Kleines, M., Lemmen, S., Brokmann, J., Mueller, T., Mueller-Wieland, D.,

- Marx, G., & Miklaus, M. (2020). The Characteristics of 50 Hospitalized COVID-19 Patients With and Without ARDS. *Deutsches Aertzblatt International*, 117(16), 271–278. <https://doi.org/0.3238/arztebl.2020.0271>
- Ellinghaus, D., Degenhardt, F., Bujanda, L., Buti, M., Albillos, A., Invernizzi, P., Fernández, J., Prati, D., Baselli, G., Asselta, R., et al. (2020). Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N. Engl. J. Med.*
- Janssen, R., Bont, L., Siezen, C.L.E., Hodemaekers, H.M., Ermers, M.J., Doornbos, G., Slot, R., van 't, Wijmenga, C., Goeman, J.J., Kimpen, J.L.L., et al. (2007). Genetic Susceptibility to Respiratory Syncytial Virus Bronchiolitis Is Predominantly Associated with Innate Immune Genes. *J. Infect. Dis.* 196, 826–834.
- Krishnan, N., Dickman, M.B., and Becker, D.F. (2008). Proline modulates the intracellular redox environment and protects mammalian cells against oxidative stress. *Free Radic. Biol. Med.* 44, 671–681.
- Kuba, K., Imai, Y., Ohto-Nakanishi, T., and Penninger, J.M. (2010). Trilogy of ACE2: A peptidase in the renin–angiotensin system, a SARS receptor, and a partner for amino acid transporters. *Pharmacol. Ther.* 128, 119–128.
- Mackenzie, J. S., & Smith, D. W. (2020). COVID-19: a novel zoonotic disease caused by a coronavirus from China: what we know and what we don't. *Microbiology Australia*. <https://doi.org/10.1071/MA2001>
- Pulver, S.R., and Berni, J. (2012). The Fundamentals of Flying: Simple and Inexpensive Strategies for Employing *Drosophila* Genetics in Neuroscience Teaching Laboratories. *J. Undergrad. Neurosci. Educ.* 11, A139–A148.
- Wein, A.N., McMaster, S.R., Takamura, S., Dunbar, P.R., Cartwright, E.K., Hayward, S.L., McManus, D.T., Shimaoka, T., Ueha, S., Tsukui, T., et al. (2019). CXCR6 regulates localization of tissue-resident memory CD8 T cells to the airways. *J. Exp. Med.* 216, 2748–2762.
- Yatsenko, A., Marrone, A., Kucherenko, M., and Schherbata, H. (2014). Measurement of Metabolic Rate in *Drosophila* using Respirometry. *J. Vis. Exp.*
- Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* 395, 1054–1062.
- van der Made, C.I., Simons, A., Schuurs-Hoeijmakers, J., van den Heuvel, G., Mantere, T., Kersten, S., van Deuren, R.C., Steehouwer, M., van Reijmersdal, S.V., Jaeger, M., et al. (2020). Presence of Genetic Variants Among Young Men With Severe COVID-19. *JAMA* 324, 663.
- Boopathi, S., Poma, A.B., and Kolandaivel, P. (2020). Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *J. Biomol. Struct. Dyn.* 1–10.
- Pandhare, J., Cooper, S.K., Donald, S.P., and Phang, J.M. (2009). Regulation and Function of Proline Oxidase under Nutrient Stress. *J. Cell. Biochem.* 107, 759–768.
- Pandey, U.B., and Nichols, C.D. (2011). Human Disease Models in *Drosophila melanogaster* and the Role of the Fly in Therapeutic Drug Discovery. *Pharmacol. Rev.* 63, 411–436.