

Bioinformatic characterization of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis

Cherly Chan

Harrow International School Hong Kong, Tuen Mun, New Territories, Hong Kong

Abstract

Gastroenteritis is a significant diarrheal disease worldwide, and its incidence is associated with environmental factors including temperature and rainfall. The molecular basis of this association is poorly understood. Elevated temperature has been shown to induce IL-6 release via the aryl hydrocarbon receptor pathway, and IL-6 levels are elevated in acute gastroenteritis and correlate with disease severity. This paper does not test that relationship directly; instead, it characterizes IL-6 computationally and develops a gene-editing framework with which the hypothesis that environmental exposures modify gastroenteritis-associated inflammation through IL-6 could be tested experimentally. In silico analyses were conducted using TMHMM, SignalP, Clustal Omega, and the STRING database to investigate the properties of the IL-6 protein. TMHMM predicted no transmembrane domains in IL-6, indicating that it is not membrane-bound. SignalP identified a 27-amino-acid signal peptide that is cleaved during secretion. Clustal Omega alignment across six species identified a small number of fully conserved residues alongside extensive divergence, which may inform the selection of appropriate model organisms for studying human IL-6 biology. The STRING analysis displayed IL-6 with its first shell interactors, including IL6R, IL6ST/gp130, JAK1, and JAK2, together with IL1B, TNFRSF1A, IL10, IL2, and IL3. A CRISPR/Cas9 strategy for generating an IL6 knockout was designed using CRISPR Finder and SnapGene with stringent off-target criteria, and simulated DNA gel electrophoresis was used to illustrate the expected range of editing outcomes. No editing, expression, or infection experiments were performed. This study provides a hypothesis-generating in-silico framework and defined editing strategy for future work, testing whether IL-6 modifies environmental effects on gastroenteritis outcomes.

Keywords: IL-6, gastroenteritis, CRISPR, Cas9, inflammation, environmental factors, cytokines, bioinformatics, gene editing, signal peptide, protein-protein interaction, temperature, immune response



1. Introduction

Gastroenteritis is the inflammation of the stomach, small intestine, or large intestine, which can lead to a series of medical issues, including abdominal pain, cramping, nausea, vomiting, and diarrhea (Graves, 2013). The symptoms of gastroenteritis include loss of appetite, abdominal pain and cramping, fever, chills, fatigue, body aches, and diarrhea (*Viral Gastroenteritis (Stomach Flu)*-*Viral Gastroenteritis (Stomach Flu) - Diagnosis & Treatment*, n.d.). Diagnosis of gastroenteritis is often based on the patient's symptoms. In cases where symptoms are mild and short-lasting, a medical history, physical exam, and stool tests are not required. However, in more serious cases, a medical history, physical examination, and stool tests may be required (*Diagnosis of Viral Gastroenteritis ("Stomach Flu") - NIDDK*, n.d.). Symptoms result from a combination of pathogen-associated injury to the gut lining and the host inflammatory response mounted against it. Inflammatory cells are sent to repair damaged tissues and fight the infection, but sometimes they are not enough. There is usually no specific treatment for gastroenteritis, but it is important for patients to implement self-care measures and stay hydrated (*Viral Gastroenteritis (Stomach Flu)*-*Viral Gastroenteritis (Stomach Flu) - Diagnosis & Treatment*, n.d.). The underlying reasons for disease development are also not fully elucidated. This paper focuses on IL-6, a cytokine that is both inducible by environmental conditions and elevated during enteric infection.

There are different types of gastroenteritis. Firstly, bacterial gastroenteritis, which is an infection of the intestines and stomach, is commonly contracted through contaminated food or water. Secondly, viral gastroenteritis, which is the most common type of gastroenteritis. It is contracted when viruses infect the stomach and intestines (Figure 1), triggering sudden and uncomfortable symptoms. Because infection risk is closely linked to environmental conditions, IL-6 is examined here alongside those factors.

Viral gastroenteritis can also spread through contaminated food or water, as well as through person-to-person and surface contact. Beyond viral causes, acute gastroenteritis encompasses several distinct etiological categories (Chiejina & Samant, 2026). Non-infectious gastroenteritis is inflammation of the stomach and intestines that, unlike infectious forms, arises from direct chemical or allergic irritation of the gut lining; potential triggers are diverse, ranging from medications and environmental toxins to specific food components that the body cannot tolerate. Bacterial gastroenteritis is an infection of the stomach and intestines that typically stems from consuming contaminated food or water, with common agents including *Campylobacter*, *Salmonella*, *Shigella*, and *Escherichia coli*. Parasitic gastroenteritis, though less common, is a significant cause of digestive problems and results from infection by protozoa such as *Cryptosporidium* and *Giardia* that invade the gastrointestinal tract (Chiejina & Samant, 2026). Environmental factors such as weather variables causing temperature changes affect the infections (Figure 2). Cold temperatures heavily affect disease transmission via increased stability and the infectivity of enteric viruses, which leads to increased risk for gastroenteritis (Chong et al., 2021). Figure 2 illustrates the seasonal relationship between weather variables and pediatric gastroenteritis admissions, which motivates the focus on environmental influences developed in the following sections.



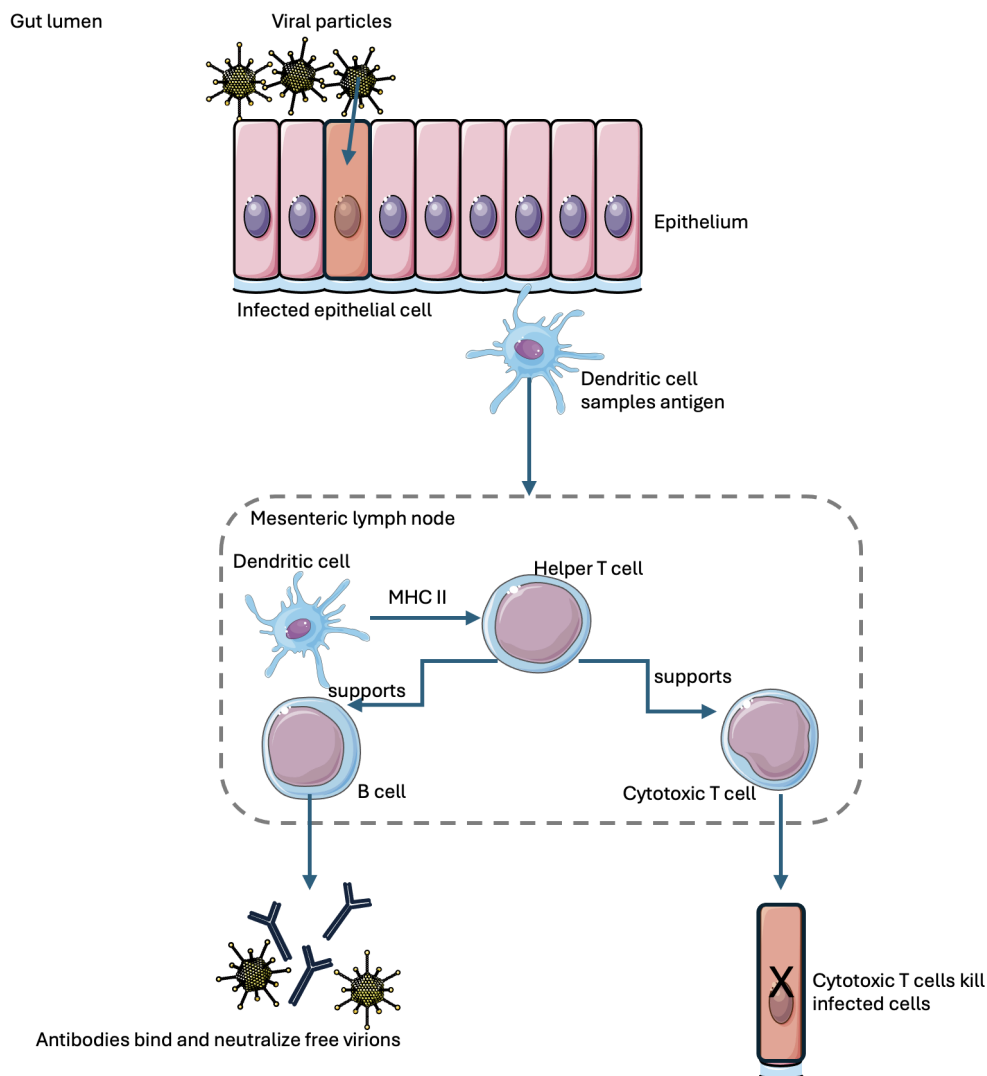


Figure 1: Immune response to enteric viral infection.

Note. Viral particles in the gut lumen infect intestinal epithelial cells. Dendritic cells sample viral antigen and migrate to the mesenteric lymph node, where antigen is presented on MHC II to helper T cells. Activated helper T cells support the responses of antigen-specific B cells and cytotoxic T cells. Antibodies produced by B cells bind and neutralize free virions, while cytotoxic T cells destroy infected epithelial cells. Adenovirus is shown as a representative enteric virus. Figure created using images from Servier Medical Art, licensed under CC BY 4.0.

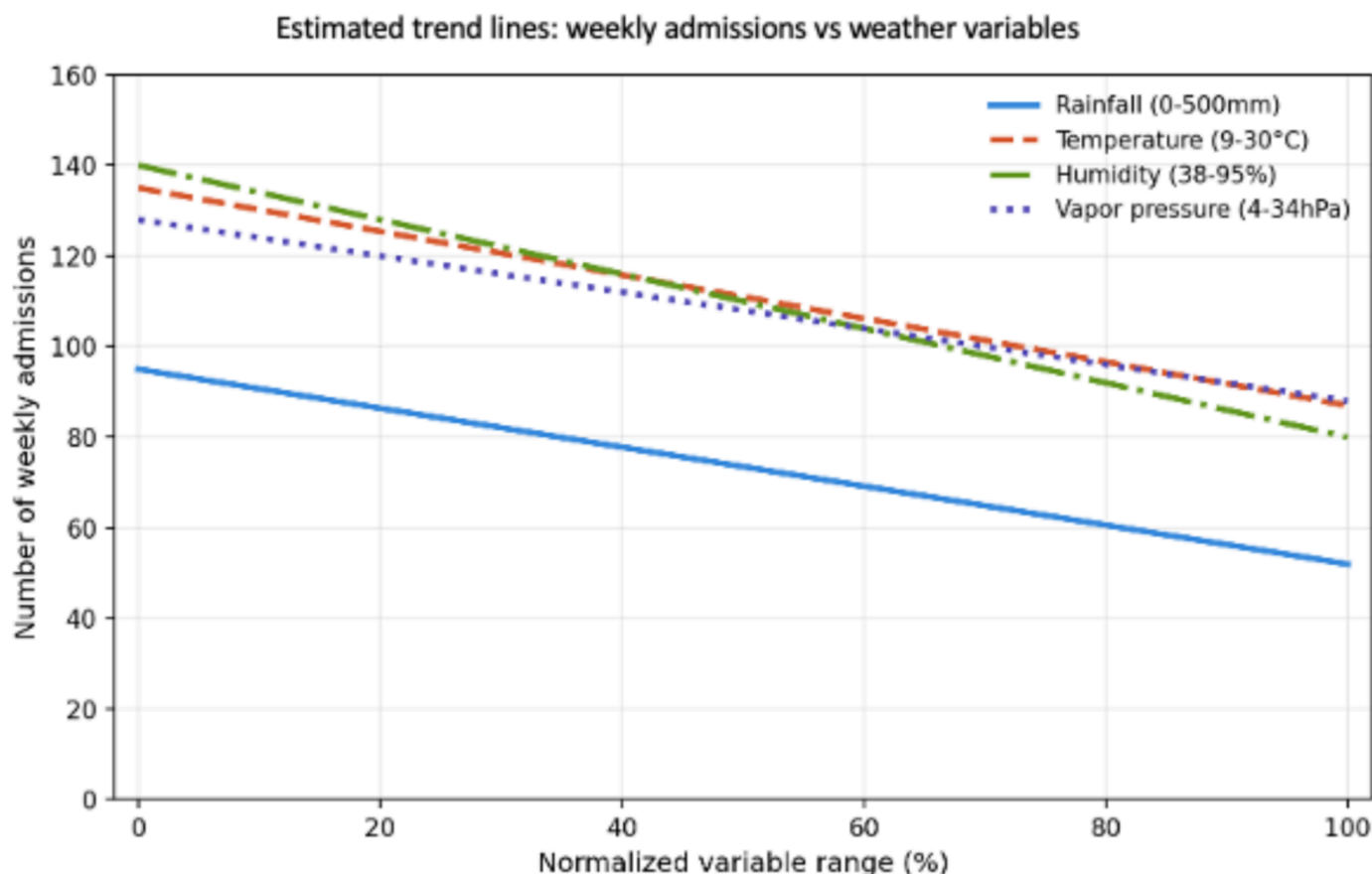


Figure 2: Statistics of pediatric patients admitted to hospital for gastrointestinal infection due to environmental factors (weather variables).

Note. A graph showing the weekly number of pediatric gastrointestinal infection admissions against different meteorological factors such as rainfall, temperature, humidity, and vapor pressure from 1998 to 2018 was normalized and plotted. The y-axis shows the number of people admitted weekly to hospitals, and the x-axis shows a normalized range of weather variables. All four show the same basic pattern: admissions trend downward as each weather variable increases, but rainfall and humidity show the steepest relative declines, while temperature's decline is the gentlest. This figure was produced by the author in Microsoft Excel using the underlying data reported by Chong et al. (2021), and is not reproduced or adapted from any published figure. Each weather variable was normalized to a 0-1 range across its observed values.

The inverse relationship between these meteorological variables and pediatric admissions shown in Figure 2 highlights the seasonal profile of childhood gastroenteritis in affluent environments. The steep decline in admissions associated with rising humidity and temperature reflects the classic winter-peak dynamics of enteric viruses like rotavirus and norovirus, which exhibit maximum stability and transmission in cold, dry conditions. Conversely, the flatter, gentler slopes for rainfall and

vapor pressure demonstrate that wet, high-temperature summer conditions do not act as primary drivers of disease. While heavy rainfall typically correlates with increased bacterial pathogen exposure in less-developed regions via water contamination pathways, these vectors are highly controlled in affluent settings with advanced water treatment infrastructure (Chong et al., 2021). Consequently, the independent influence of rainfall on transmission remains minor in comparison to the dominant, cold-weather viral drivers.

Beyond temperature's effect on pathogen stability and transmission, environmental factors may also act directly on IL-6 expression at the molecular level, providing a mechanistic basis for the link between climate and gastroenteritis. Elevated temperatures have been shown to enhance the release of IL-6 and IL-8 via the aryl hydrocarbon receptor (AhR) pathway and to augment the pro-inflammatory effects of other factors (Hu et al., 2024). In a population-based comparison, serum IL-6 was significantly upregulated during high-temperature periods relative to cooler ones, and a controlled cell model demonstrated that an increase of just 1.5 °C is sufficient to elevate IL-6 and IL-8 at both the gene and protein levels (Hu et al., 2024). This positions the AhR pathway as a mediator between environmental heat stress and immune regulation, suggesting that warming itself, independent of infection, can prime the IL-6-driven inflammatory axis. A similar pattern is observed for airborne particulate stressors: interleukin-6 levels in blood samples collected during dust-storm air-pollution conditions were significantly higher than those in a control group (Abdipour et al., 2025), reinforcing the idea that IL-6 functions as a sensitive biomarker of environmental insult. Taken together, these findings extend the environmental link beyond rainfall and low temperature to include heat and air quality, and they motivate the hypothesis examined here, that environmental stressors may modify gastroenteritis outcomes through IL-6-mediated inflammation, since the same cytokine that is elevated in acute gastroenteritis is also directly inducible by environmental conditions.

There are genetic factors involved in the development of gastroenteritis. Seventeen genes have been associated with the gastroenteritis phenotype by text-mining of GWAS publications in the HuGE Navigator Gene-Phenotype Associations dataset, retrieved via the Harmonizome database (*Gastroenteritis Gene Set – HuGE Navigator Gene-Phenotype Associations*, n.d.): ABCC2, ABO, CXCL8, FUT2, HLA-DQA1, HLA-DQB1, IFNG, IL10, IL18, IL1B, IL1RN, IL6, IMPDH1, IMPDH2, PRRT2, TNF, and TNFRSF11B. Of the genes associated with gastroenteritis, ABCC2 belongs to the ATP-binding cassette (ABC) transporter superfamily, whose members move a wide variety of substrates across cellular membranes. Its product is a multidrug resistance-associated protein (PubChem, 2026a). ABO encodes the glycosyltransferases that define the ABO system, the first human blood group to be characterized (PubChem, 2026b). CXCL8 (also known as IL-8) encodes a CXC-family chemokine that is a major mediator of the inflammatory response (PubChem, 2026c). FUT2 encodes one of the two galactoside 2-L-fucosyltransferase enzymes in humans. Its product catalyzes the terminal reaction in the synthesis pathway that produces soluble ABO blood group antigens, and it additionally participates in cell-cell interaction, cell surface expression, and cell proliferation (PubChem, 2026d). HLA-DQA1 and HLA-DQB1 provide instructions for making a protein that has a crucial role in the immune system. It is part of the human leukocyte antigen complex (a family of genes), which helps the immune system distinguish the body's own proteins from foreign ones (HLA-DQA1 Gene, 2026; HLA-DQB1 Gene, 2026). These two proteins code for the alpha and beta chains of HLA, respectively. Human leukocyte antigens (HLA) are genes located within the major histocompatibility complex (MHC) and encode proteins essential for distinguishing self from non-self. They play a key role in both immune defense and disease, offering benefits to the immune system while also having potential harmful effects (Nordquist & Jamil, 2026).

IFNG encodes a soluble cytokine protein, which is secreted by the immune cells. The active protein then binds to the interferon gamma receptor, which triggers a cellular response to viral and microbial infection (PubChem, 2026e). The protein



encoded by IL10 has pleiotropic effects in immunoregulation and inflammation, and also enhances antibody production (IL10 *Interleukin 10 [Homo Sapiens (Human)] - Gene - NCBI*, 2026). The protein encoded by IL18 is a proinflammatory cytokine of the IL-1 family that is constitutively found as a precursor within the cytoplasm of a variety of cells (PubChem, 2026h). IL1B encodes a member of the interleukin 1 cytokine family that is an important mediator of the inflammatory response (PubChem, 2026f). IL1RN encodes another interleukin 1 family member, which inhibits the activities of IL1A and IL1B and modulates interleukin 1-related immune and inflammatory responses (PubChem, 2026g). The IMPDH1 product acts as a homotetramer that regulates cell growth (PubChem, 2026i). IMPDH2 encodes the rate-limiting enzyme in de novo guanine nucleotide biosynthesis (PubChem, 2026j). PRRT2 gene codes for the PRRT2 protein, which is found in neurons in the brain and regulates signaling from neuron to neuron (PRRT2 *Gene*, 2026). TNF codes for a multifunctional proinflammatory cytokine that belongs to the TNF superfamily. This cytokine is involved in the regulation of many biological processes and has been implicated in a variety of diseases, including insulin resistance and cancer (TNF *Tumor Necrosis Factor [Homo Sapiens (Human)] - Gene - NCBI*, 2026). TNFRSF11B codes for a protein that has a critical role in bone remodeling. The TNFRSF11B protein is involved in the regulation of specialized cells called osteoclasts, which break down bone tissue during bone remodeling (TNFRSF11B *Gene*, 2026).

Of these genes, IL6 was selected as the focus of this paper. Several of the others have well-characterized, pathogen-specific roles: FUT2, for instance, determines secretor status and is an established determinant of norovirus susceptibility, and its relationship to gastroenteritis is therefore already comparatively well defined. IL6 was chosen instead because it sits at the convergence of the two influences this paper is concerned with. It is induced by environmental exposures such as thermal stress and particulate air pollution, and it is also elevated during enteric infection, which makes it a plausible point of intersection between environmental conditions and infection-associated inflammation. The IL6 gene encodes a cytokine that has a role in inflammation and the maturation of B cells. It signals via its receptor, IL6R, which is also known to contribute to the development of inflammatory bowel diseases (IBD) (Parisinos et al., 2018). IL-6 is an endogenous pyrogen capable of inducing fever in people with infections or autoimmune diseases, and is released in response to gastrointestinal infection and has been implicated in the pathogenesis of acute gastroenteritis caused by rotavirus and norovirus (Chen et al., 2014). In pediatric cohorts, both serum and fecal levels of IL-6 show significant elevation compared to healthy controls, with higher levels in bacterial than in viral gastroenteritis, and with better reported sensitivity and specificity than C-reactive protein or IL-8 for distinguishing bacterial from viral etiology in those cohorts (Adumitrăchioaiei et al., 2024). Elevated IL-6 has also been reported to correlate with disease severity, including benign seizures, complications of Shigella infection, and raised liver enzymes in rotavirus gastroenteritis. These are associations reported within specific study populations rather than established clinical performance across all gastroenteritis presentations, and the direction of the relationship is not resolved: more severe infection and tissue injury may produce a larger inflammatory response, and consequently higher IL-6, rather than IL-6 driving severity. IL-6 is therefore a suitable candidate for computational characterization and for the design of a framework with which its contribution could later be tested experimentally.

As part of this paper, the protein properties of IL-6 were investigated, and a gene-editing strategy was designed in order to propose a way to generate mutant IL-6 to study this condition. For the editing, I followed the CRISPR/Cas9 gene-editing technique. CRISPR/Cas9 gene editing is a technology that requires two mandatory components: a guide RNA to pair with a desired target gene, and Cas9, which is an endonuclease that causes a double-stranded DNA break, allowing modifications to the genome to be made (Redman et al., 2016). Once Cas9 introduces this double-strand break at the target site, the cell's own



DNA repair machinery — chiefly non-homologous end joining (NHEJ) or homology-directed repair (HDR) — resolves the break, and the choice of repair pathway determines the resulting edit (Yang et al., 2020).

The aim of this study is to characterize the structural and functional properties of IL-6 using publicly available bioinformatic tools, and to design a defined CRISPR/Cas9 strategy for generating an IL-6 knockout in intestinal epithelial cells. The study does not investigate gastroenteritis directly and does not test whether IL-6 mediates the relationship between environmental exposure and disease outcome. Its purpose is to establish a framework with which that hypothesis could be tested in future experimental work.

2. Materials and Methods

For this study, the databases and tools listed below were used:

2.1. Collecting protein and gene information:

NCBI: A national resource for molecular biology information. It develops new information technologies to aid our understanding of the fundamental molecular and genetic processes that control health and disease. Creates and maintains medical and scientific databases (Staff, 2023). Protein and gene sequences for IL6 were retrieved from the NCBI Protein and NCBI Gene databases.

Ensembl: A bioinformatics project with the purpose of organizing biological information around the sequences of large genomes (Birney et al., 2004). Ensembl was used to annotate the exon and intron regions of IL6.

2.2. Investigating protein properties: transmembrane domain and conservation

TMHMM: The amino acid sequence for the human IL-6 protein was retrieved from NCBI Protein in FASTA format and submitted as the query input to the TMHMM web server. The analysis was executed with default parameters, and the server-generated graphical output, which depicts the probability of transmembrane helices, intracellular, and extracellular regions across the protein length, was exported.

SignalP: A widely used computational tool for predicting signal peptides from amino acid sequences (Teufel et al., 2022). The IL-6 protein sequence was retrieved from the NCBI protein database in FASTA format and submitted to the SignalP algorithm for analysis.

Clustal Omega: An algorithm for multiple sequence alignment that can deal with very large numbers of DNA/RNA sequences due to its use of the mBED algorithm for calculating guide trees (Sievers & Higgins, 2014). Clustal Omega was used to assess the degree of conservation of IL-6 across species.

STRING Protein: The STRING database collects, integrates, and scores all the sources of protein-protein information available to the public, and complements these with computational predictions to generate protein-protein interactions (Szklarczyk et al., 2019). This analysis was performed to identify predicted interactors of IL-6. The query was submitted as "IL-6" and resolved to the *Homo sapiens* IL6 entry. The full STRING network was selected, in which edges indicate both functional and physical associations, and network edges were displayed in evidence mode, in which line color indicates the type of



supporting evidence. All seven interaction sources were active: text mining, experiments, databases, co-expression, neighborhood, gene fusion, and co-occurrence. The minimum required interaction score was set to medium confidence (0.400), the first shell was limited to a maximum of 10 interactors, and no second shell was requested.

2.3. Designing CRISPR/Cas9 strategy:

CRISPR/Cas9 targets and cleaves specific DNA sequences, with a single guide RNA directing the Cas9 nuclease to the target site to produce a double-strand break. The intended modification in this study is a loss-of-function IL6 knockout. A single guide RNA directs Cas9 to a coding exon, and the resulting double-strand break is resolved by non-homologous end joining, which characteristically produces small insertions and deletions. Frameshifting indels are expected to disrupt the reading frame and abolish production of functional IL-6 protein. No donor template was designed and no HDR, base-editing, or prime-editing strategy was used, because the objective is gene disruption rather than reconstruction of a defined naturally occurring IL-6 variant.

SnapGene: A multipurpose software program used for designing and documenting sequence-based experiments and gene editing. SnapGene was used for sequence alignment, annotation, editing, and visualization (*SnapGene Software Application*, n.d.)

CRISPR Finder (Sanger): A tool produced by the Wellcome Sanger Institute for selecting guide RNAs for human and mouse genomes. The gene of interest was searched and an exon selected, after which the algorithm produced candidate guide RNAs. Guides were selected based on off-target scores, with the criteria 0 mismatches = 1, 1 mismatch = 0, 2 mismatches = 0, and the lowest available values for 3 and 4 mismatches

3. Results

3.1. IL-6 does not contain a transmembrane domain

To predict the subcellular localization of the Interleukin-6 (IL-6) protein, I utilized TMHMM 2.0. Knowing the subcellular localization of a protein gives an indication of its functions and potential consequences of mutations, as mutations could change this localization. TMHMM predicts transmembrane domains, referring to sequential hydrophobic amino acids in a given protein sequence. Proteins are located inside membranes via these domains. This predicted topology plot was directly incorporated as a figure panel to visually confirm the absence of transmembrane domains in IL-6, indicating that it cannot be a membrane-bound protein (Figure 3). The plot shows the presence of hydrophobic amino acids at the very beginning of the IL-6 protein; however, the plot does not mark this as transmembrane because these amino acids are potentially signal peptides, and this was tested via the SignalP 5.0 algorithm for IL-6 (Figure 4). These signal peptides will be cleaved after directing the protein to a certain location.



RecName: Full=Interleukin-6; Short=IL-6; AltName: Full=B-cell stimulatory factor 2; Short=BSF-2; AltName: Full=CTL differentiation factor; Short=CDF; AltName: Full=Hybridoma growth factor; AltName: Full=Interferon beta-2; Short=IFN-beta-2; Flags: Precursor

UniProtKB/Swiss-Prot: P05231.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

```
>sp|P05231.1|IL6_HUMAN RecName: Full=Interleukin-6; Short=IL-6; AltName: Full=B-
cell stimulatory factor 2; Short=BSF-2; AltName: Full=CTL differentiation factor;
Short=CDF; AltName: Full=Hybridoma growth factor; AltName: Full=Interferon beta-2;
Short=IFN-beta-2; Flags: Precursor
MNSFSTSAFGPVAFSLGLLLVLPAAFPAPVPPGEDSKDVAAPHRQPLTSSERIDKQIRYILDGISALRKE
TCNKSNMCESSKEALAENNLNLPKMAEKDGCFCQSGFNEETCLVKIITGLLEFEVYLEYLQNRFSSEEQA
RAVQMSTKVLIIQLQKKAKNLDAITTPDPTNASLLTKLQAQNQLQDMTTHLILRSFKEFLQSSLRALR
```

```
# sp|P05231.1|IL6_HUMAN Length: 212
# sp|P05231.1|IL6_HUMAN Number of predicted TMHs: 0
# sp|P05231.1|IL6_HUMAN Exp number of AAs in TMHs: 14.78953
# sp|P05231.1|IL6_HUMAN Exp number, first 60 AAs: 14.78288
# sp|P05231.1|IL6_HUMAN Total prob of N-in: 0.66119
# sp|P05231.1|IL6_HUMAN POSSIBLE N-term signal sequence
sp|P05231.1|IL6_HUMAN TMHMM2.0 outside 1 212
```

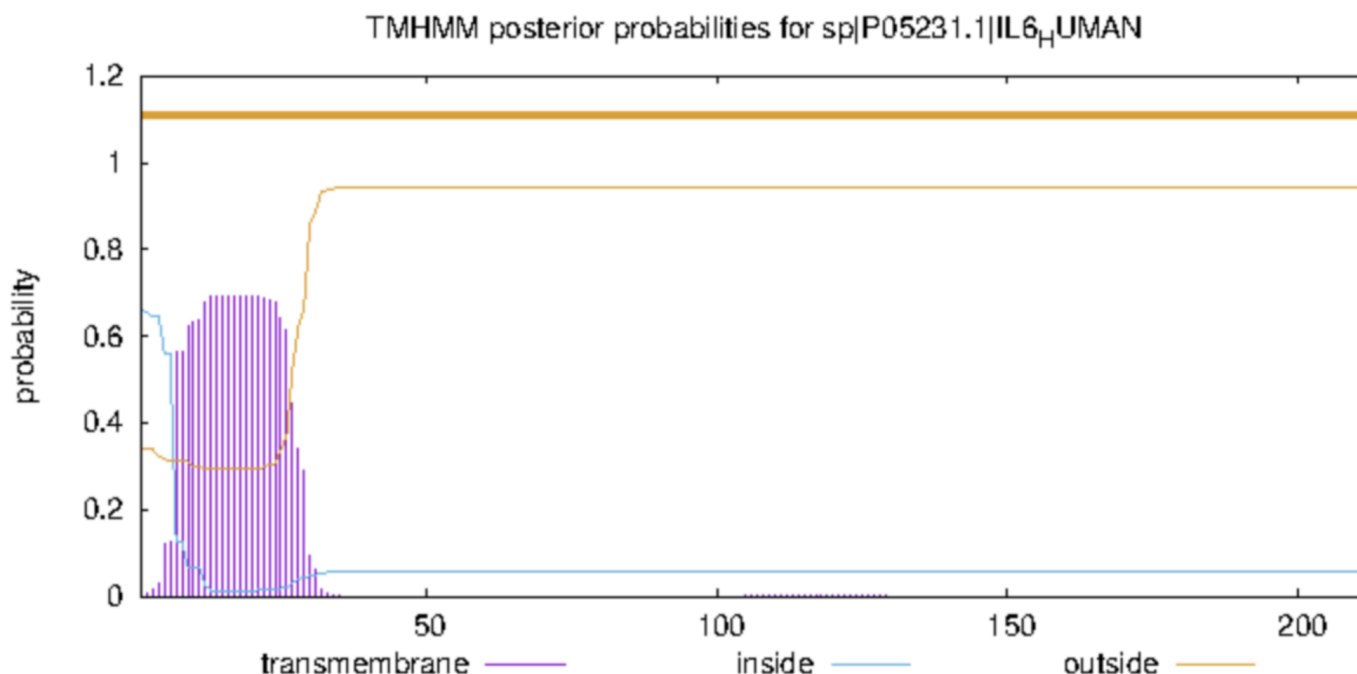


Figure 3: TMHMM transmembrane domain prediction result for the IL-6 protein.



Note. The upper text is the protein amino acid sequence retrieved from the NCBI Protein database. The bottom text and plot are the results of TMHMM analysis for IL-6. The text of the TMHMM result shows the amino acid length and the result of the prediction, as stated: TMHs: 0. The plot shows the probability of the amino acids being hydrophobic or not. The x-axis shows the length of the amino acids, and the y-axis shows the level of probability (0-1). The yellow line presents the region that is outside of a membrane, the blue line presents the region that passes after the inner side of a membrane, and purple lines refer to the transmembrane domain. The purple lines at the beginning of the protein refer to the signal peptides; therefore, they are not marked as transmembrane helices for this analysis.

3.2. IL-6 contains a signal peptide

The TMHMM analysis revealed that the IL-6 protein does not contain a transmembrane domain (Figure 3). However, the plot shows a series of hydrophobic amino acids at the beginning of the protein. One possible explanation is that if a protein contains a signal peptide to be localized to its correct position, then that peptide will be cleaved. To confirm this, I did Signal peptide analysis (Figure 4). SignalP 5.0 is a specialized bioinformatics tool that uses neural networks to analyze a protein's amino acid sequence. It predicts the presence and location of an N-terminal signal peptide, a short sequence tag that directs a newly synthesized protein for secretion from the cell. These results are consistent with IL-6 being synthesized as a precursor (pre-IL-6) with a 27-amino-acid signal peptide, which is cleaved off during secretion to release the mature, active IL-6 protein starting at position 28.

3.3. Multiple protein alignment of IL6 among different species

Clustal Omega is an algorithm that aligns the protein sequence of a protein of interest from different species to measure the level of conservation. This analysis was done to identify conserved regions of the IL-6 protein, with the purpose of finding out which parts of the IL-6 protein are identical or similar across evolution. The results could point out which amino acids are most likely to be important for function. This analysis is generally used to predict the conservation level of a protein, detect domains, and understand the impact of a mutation by seeing if it falls in a conserved residue or not. Sometimes this could be the answer for the diversity observed across patients, and this is also valid for gastroenteritis patients. IL-6 across zebrafish, chicken, mouse, chimpanzee, pig, and human is not found to be highly conserved (Figure 5). There are a couple of amino acids marked with * which means that they are conserved across all species, but the majority of the amino acids are not conserved based on the Clustal Omega analysis. Knowing whether a protein has a lot of conserved amino acids across different species would show us the evolutionary importance and similarity of a given protein and its amino acids. This analysis could hint that the molecular way of IL-6-related immune mechanisms is not conserved across the species put in this analysis.



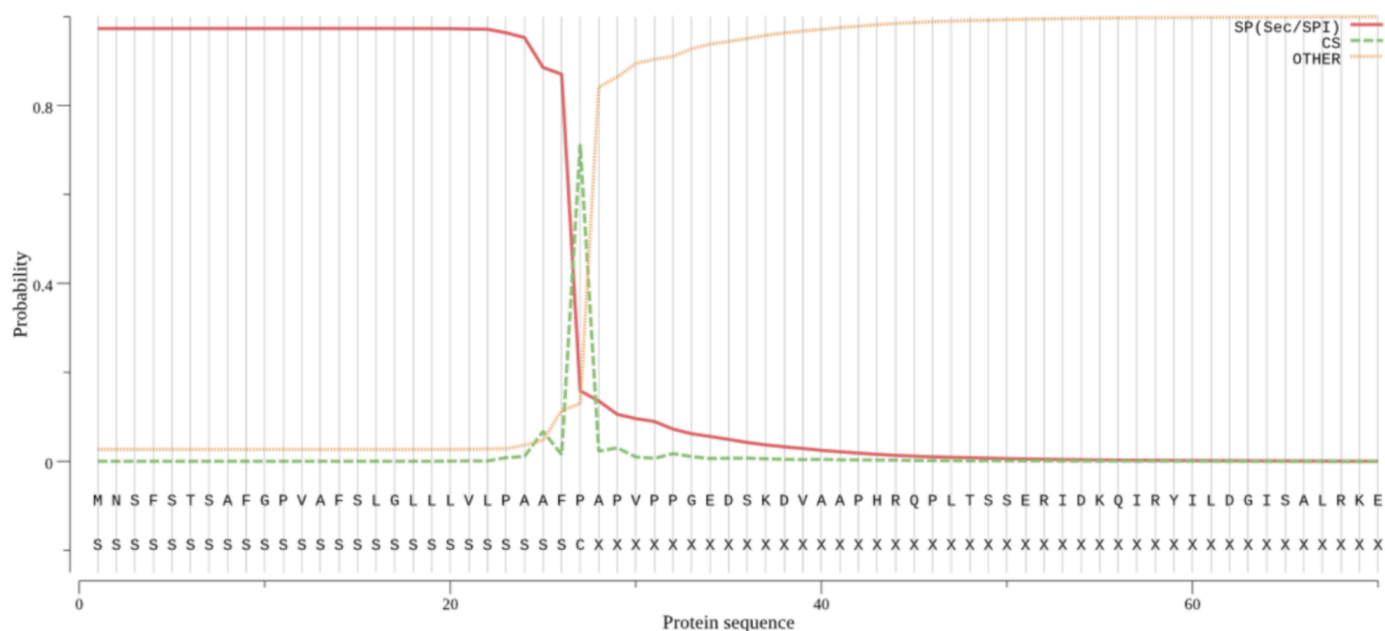
sp_P05231.1_IL6_HUMAN**Prediction:** Signal peptide (Sec/SPI)

Cleavage site between pos. 27 and 28: AFP-AP. Probability: 0.7122

Protein type	Signal Peptide (Sec/SPI)	Other
Likelihood	0.9733	0.0267

Download: [PNG](#) / [EPS](#) / [Tabular](#)

SignalP-5.0 prediction (Eukarya): sp_P05231.1_IL6_HUMAN

**Figure 4:** SignalP 5.0 analysis of IL-6.

Note. The plot shows the prediction of the presence of a signal peptide in IL6. The x-axis shows the amino acid sequence of IL-6; the y-axis shows the probability calculation between 0 and 1. The red line represents the amino acids that are signal peptides, the green line represents the cleavage site, and the yellow line represents the rest of the protein. The IL6 signal peptide will be cleaved between positions 27 and 28.



CLUSTAL O(1.2.4) multiple sequence alignment

```

zebrafish_NP_001248378.1      -----MPSAQKTVLFLSATLATLFMSLADPVPVFSSMGLSEI 38
chicken_sp|Q90YI0.1|IL6_CHICK MNFTGECATGRRPGSAGSRRRRAPRPGVALLP-LLLPLLLPPAAAVPLPAA----- 52
mouse_sp|P08505.1|IL6_MOUSE  MK-----FLSARDFHPVAF---LGLMLVTTTAFPTSQVRR----- 32
human_sp|P05231.1|IL6_HUMAN  MN-----SFSTSAFGPVAFS--LGLLLVLPAAFPAPVPPG----- 33
chimpanzee_XP_518992.2       MN-----SVSTSAFGPVAFS--LGLLLVLPAAFPAPVPPG----- 33
pig_sp|P26893.2|IL6_PIG      MN-----SLSTSAFSPVAFS--LGLLLVMATAFPTPERLE----- 33
                                * :      * . * .

zebrafish_NP_001248378.1      SGDEVQDQVDVKSLLGDRQ-----KWHLMARDLYKDVKTLRDEQFERDFREMNMTAFEG 92
chicken_sp|Q90YI0.1|IL6_CHICK -ADSSGEGVLEEEAGARRAL---LDCEPLARVLRDRAVLQDEMCKFTVCENSMMLVR 108
mouse_sp|P08505.1|IL6_MOUSE  -GDFTEDTT-----PNRPVYTT-SQVGGLITHVLWEIVEMRKELCNGNSDCMNDALAE 85
human_sp|P05231.1|IL6_HUMAN  -EDSKDVAA-----PHRQPLTSSERIDKQIRYILDGISALRKETCNKSNMCESSKEALAE 87
chimpanzee_XP_518992.2       -EDSKDVAA-----PHRQPLTSSERIDKQIRYILDGISALRKETCNKSNMCESSKEALAE 87
pig_sp|P26893.2|IL6_PIG      -EDAKGDAT-----SDKMLFTSPDKTEELIKYILGKISAMRKEMCKEYKCEKNSKEVLAE 87
                                * . :      : : : * : . :

zebrafish_NP_001248378.1      VRISTPLLKPSDRCLSKNFSTERCLTRIYSVLTWYKDNWNYIEKENLTSV--LVNDIKHS 150
chicken_sp|Q90YI0.1|IL6_CHICK NNLNLPKVTEEDGCLLAGFDEEKCLTKLSSGLFAFQTYLEFIQTFDSEK-QNVESLCYS 167
mouse_sp|P08505.1|IL6_MOUSE  NNLKLPETQRNDGCGYQTGYNQEICLLKISSGLLEYHSYLEYMKNNLKDNKKDKARVLQD 145
human_sp|P05231.1|IL6_HUMAN  NNLNLPKMAEKDGCFSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSE-EQARAVQMS 146
chimpanzee_XP_518992.2       NNLNLPKMAEKDGCFSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSE-EQARAVQMS 146
pig_sp|P26893.2|IL6_PIG      NNLNLPKMAEKDGCFSGFNETCLMRITTTGLVFEFYLYDYLYQKEYESNK-GNVEAVQIS 146
                                . : * : . * * . : . * * * : : * . : : : . . : .

zebrafish_NP_001248378.1      TKRLLLEANSQLQVRD-----GEMDQTSSTLSFKSAWTRKTTVHSILFNFSVIMDAC 204
chicken_sp|Q90YI0.1|IL6_CHICK TKHLAATIRQMVINPDEVVIPSAAQKSLLANLKSQDKWIEKITMHLIRDFTSFMEKTV 227
mouse_sp|P08505.1|IL6_MOUSE  TETLIHIFNQEVKDLHKIVLPTISNALLTDKLESQKEWLRTKTIQFILKSLEEFVKVTL 205
human_sp|P05231.1|IL6_HUMAN  TKVLIQFLQKKAKNLDATTPDPTTNASLLTKLQAQNWQLQDMTHLILRSFKEFLQSSL 206
chimpanzee_XP_518992.2       TKVLIQFLQKKAKNLDATTPDPTTNASLLTKLQAQNWQLQDMTHLILRSFKEFLQSSL 206
pig_sp|P26893.2|IL6_PIG      TKALIQTLRQKGNPKATTPNPTNAGLLDKLQSQNEWMKNTKIILILRSLEDFLQFSL 206
                                *: * : . . :      : . * . . * . . * * : : : :

zebrafish_NP_001248378.1      RAINYMSRRRGKDPKRTGDWGSADKN 231
chicken_sp|Q90YI0.1|IL6_CHICK RAVRYLKKTFSFSA----- 241
mouse_sp|P08505.1|IL6_MOUSE  RSTRQT----- 211
human_sp|P05231.1|IL6_HUMAN  RALRQM----- 212
chimpanzee_XP_518992.2       RALRQM----- 212
pig_sp|P26893.2|IL6_PIG      RAIM----- 212
                                *: .

```

Figure 5: Multiple protein alignment of IL6 across different species.

Note. The alignment shows the species names on the left side and the amino acids aligned on the right side. There is an amino acid coloring scheme (ClustalX convention): which means RED: Small, hydrophobic amino acids (A, V, L, I, M, F, W), BLUE: Positively charged amino acids (K, R), MAGENTA: Negatively charged amino acids (D, E), GREEN: Hydroxyl-, sulfhydryl-, amine-containing amino acids (S, T, N, Q) and glycine (G), PINK: Proline (P), YELLOW: Cysteine, ORANGE: Tyrosine (Y), CYAN: Histidine (H), WHITE/GRAY: Unconserved or gapped positions. The conservation is marked with symbols: “*” means highly conserved, “.” means semi-conserved, and “:” means the least conserved. If there is no symbol, the level of conservation is not applicable.



3.4. Protein-protein interaction map of IL-6

Protein-protein interaction analysis of IL-6 was performed using the STRING database. This tool maps functional and physical interactions between proteins across thousands of organisms. It integrates evidence from 7 distinct “channels”: Genomic Context (3 channels), High-Throughput Experiments, Co-expression, automated Text Mining, and Database Knowledge to identify protein associations. The key interactors for IL-6 include: IL6R - the primary binding subunit; IL6ST - the common signal-transducing subunit shared with other cytokines; and JAK1, JAK2 - kinases activated upon receptor binding. It also shows how IL6 signaling is connected to major inflammatory pathways that include: IL1B, TNFRSF1A, IL10, IL2, IL3. This analysis was done to understand its role as part of the immune system's integrated communication network, and it is generally used to study developmental phenotypes, as phenotypes are outcomes of multiple proteins that are interacting (Wimalagunasekara et al., 2023). IL-6 was the query protein for this analysis, so its position in the network reflects the search strategy rather than an independent finding of centrality. What the network does show is that IL-6's predicted first-shell interactors are dominated by cytokine signaling and receptor components, consistent with its established role in inflammatory signaling.

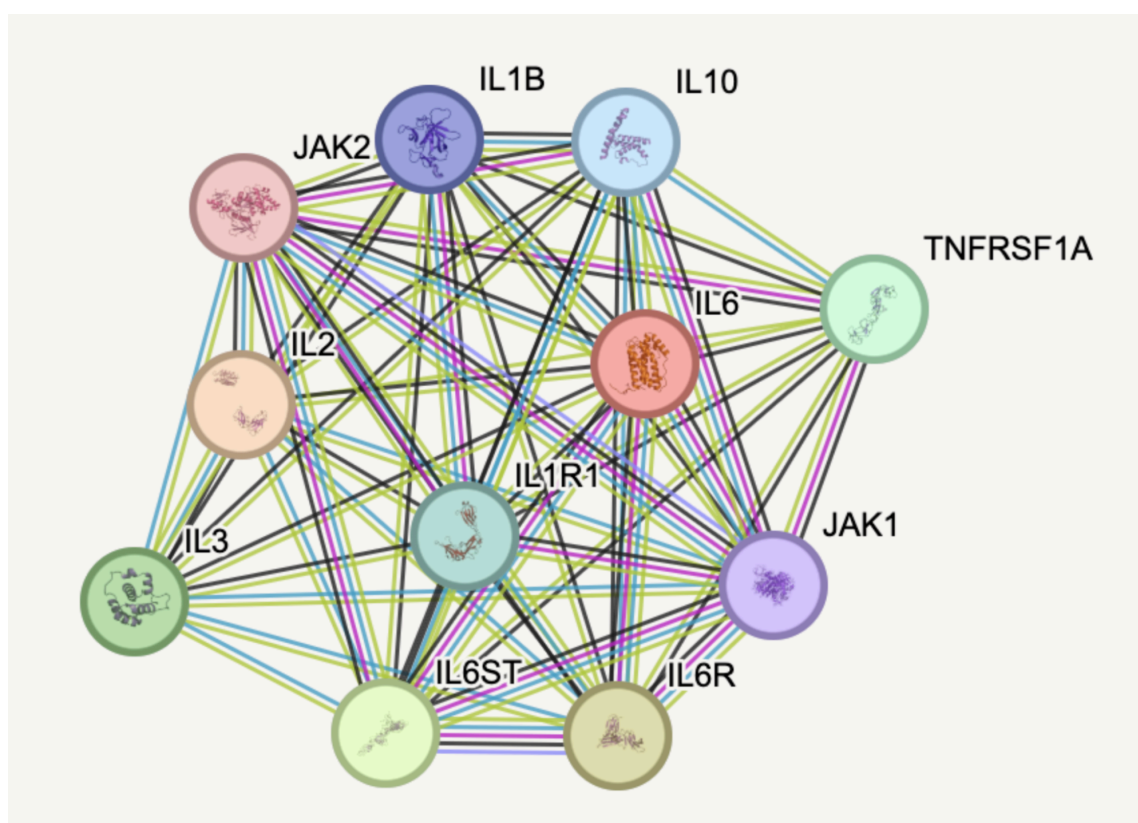


Figure 6: Protein-protein interaction map of IL-6.

Note. The map shows the proteins predicted to interact with IL-6, generated using the STRING database. Nodes represent proteins: IL6 (center), its receptor subunits IL6R and IL6ST, the associated kinases JAK1 and JAK2, and further inflammatory mediators IL1B, IL1R1, IL10, IL2, IL3, and TNFRSF1A. Node colors are assigned arbitrarily by STRING and do not encode function. The color of each connecting line indicates the type of evidence supporting the interaction: light blue, curated databases; magenta, experimentally determined interactions; green, gene neighborhood or gene fusion; red, gene co-occurrence across genomes; blue, text mining; yellow, co-expression; black, protein homology.

3.5. CRISPR/Cas9 gene-editing strategy for IL6

In this study, a hypothetical CRISPR/Cas9 gene-editing strategy has been designed. For training purposes, the gene-editing design was followed by a hypothetical polymerase chain reaction (PCR), a laboratory nucleic acid amplification technique using polymerase, primers, nucleotides, and a DNA template to amplify regions of interest (Khehra et al., 2026). This is followed by simulated DNA gel electrophoresis. Hypothetically, to define the outcome of the CRISPR/Cas9 gene editing, one could do DNA gel electrophoresis upon completion of PCR. The strategy proposed here is designed to generate a loss-of-function IL6 knockout in intestinal epithelial cells, which would provide the starting material for future work examining how the loss of IL-6 signaling alters the epithelial response to environmental and infectious challenge.

For the gene-editing design, NCBI>GENE, Ensembl, CRISPR Finder (Sanger), and SnapGene were used. The gene sequence of IL6 was retrieved through the NCBI>gene database and used to generate a SnapGene file to visualize the sequence and design the editing strategy. First, the chromosomal location of IL6 on the NCBI gene database was checked on the NCBI database to confirm that the target region does not overlap with the neighboring genes (Figure 7). The gene sequence annotation was done by looking at Ensembl exon and intron annotations, and a coding exon was selected as the target. Candidate guide RNAs were then identified using the CRISPR Finder database. The selected guide had the most favorable available off-target profile, with one exact match (0 mismatches), no sites with 1 or 2 mismatches, and the lowest available counts at 3 and 4 mismatches. The guide and the PAM sequences were marked on SnapGene Viewer (Figure 8). The PCR primers flanking the cut site were then designed for hypothetical follow-up experiments. The precise outcomes would need to be verified by sequencing in actual laboratory practice.

The proposed strategy enables a specific follow-up experiment. Intestinal epithelial cells carrying the IL6 knockout described above, alongside unedited wild-type controls, would be cultured at 37 °C and at 38.5 °C for 24 h, replicating the 1.5 °C increase shown to induce IL-6 via the AhR pathway (Hu et al., 2024). Because the knockout removes functional IL-6, the readout cannot be IL-6 itself. Instead, the response would be measured downstream: phosphorylation of STAT3 by Western blot, transcript levels of other cytokines in the STRING first shell such as IL1B and IL10 by qPCR, and epithelial barrier integrity by transepithelial electrical resistance. To connect the model to gastroenteritis rather than to thermal stress alone, both genotypes would additionally be challenged with an enteric pathogen or viral surrogate at each temperature, with barrier disruption and cytokine output as endpoints. If IL-6 is required for the thermal component of the epithelial inflammatory response, the temperature-associated difference in these endpoints should be reduced or absent in knockout cells relative to wild-type. Comparable responses in both genotypes would argue that IL-6 is not required for that effect and would redirect attention to the other interactors identified in the STRING network (Figure 6).



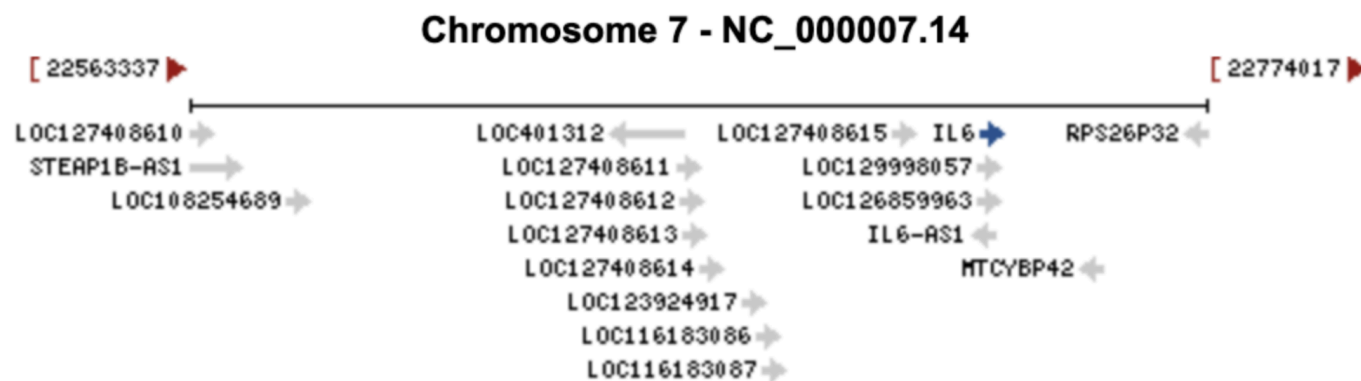


Figure 7: Chromosomal Location of IL-6.

Note. The genomic region spanning positions 22,563,337 to 22,774,017 on chromosome 7 is shown. The orientation of each gene is indicated by arrows. The purpose of this map is to identify the IL6 genomic location and guide the design of the CRISPR/Cas9 gene-editing targets within the IL6 locus.

3.6. Simulated DNA gel electrophoresis

PCR followed by DNA gel electrophoresis was proposed and simulated as a method for screening the outcome of the editing strategy described above. Because non-homologous end joining produces a range of insertion and deletion sizes at the cut site, edited cells would need to be genotyped before any downstream comparison could be made. To analyze the potential results of the gene-editing attempt, DNA gel electrophoresis, a technique that separates DNA fragments based on their size as they move through an agarose gel within an electric field generated by an electric current, can be applied, as simulated here. The purpose of this was to understand what may have happened due to CRISPR editing. Molecules of DNA are separated by size according to their uniform mass/charge ratio in the agarose gel, so that the molecules are arranged in a pattern where the distance traveled is inversely proportional to the log of their molecular weight (Lee et al., 2012). The separated fragments are then stained with a fluorescent nucleic acid dye and visualized under UV illumination on a transilluminator, with their sizes estimated by comparison against a co-electrophoresed DNA ladder marker (Motohashi, 2019). Several outcomes were hypothetically defined as specific sizes of deletion and insertion mutations to simulate the gel view (Figure 9). The deletion would produce a fragment that is smaller than the wild type, and an insertion would produce a fragment bigger than the wild type.



Exons for ENST00000404625

1. ENSE00001560586 (length 375)
2. ENSE00001556856 (length 103)
3. ENSE00003892968 (length 191)
4. ENSE00003681793 (length 114)
5. ENSE00000041007 (length 114)

1021639965	7:22727432-22727454	CCTCCGGCACAGGCGCCTTCGGT	0: 1
			1: 0
			2: 0
			3: 1
			4: 54

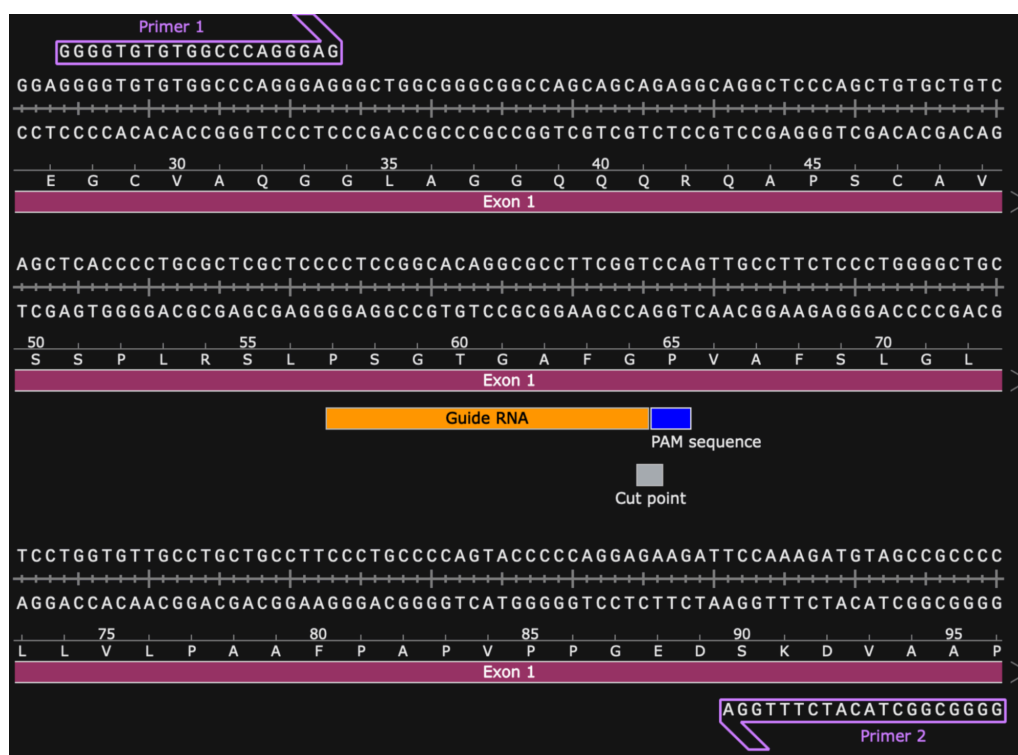


Figure 8: SnapGene view of CRISPR/Cas9 gene-editing design for IL6.

Note. The upper image is the list of exons for IL6 in the CRISPR Finder database. The highlighted one is the selected exon for gene editing. The middle image shows the selected Guide RNA (orange in the SnapGene view below) and the off-target scores (the far-right column). The lower image shows the SnapGene view of the CRISPR/Cas9 design for IL6's exon (purple), highlighting the position of the guide RNA (orange), the predicted Cas9 cleavage site (grey), approximately 3 nt upstream of the PAM, the PAM sequence (blue), and primers 1 and 2 framed in pink arrows. Primer 1 (framed in pink arrows) and Primer 2 (framed in pink arrows) produce a product of 207 bp in length.

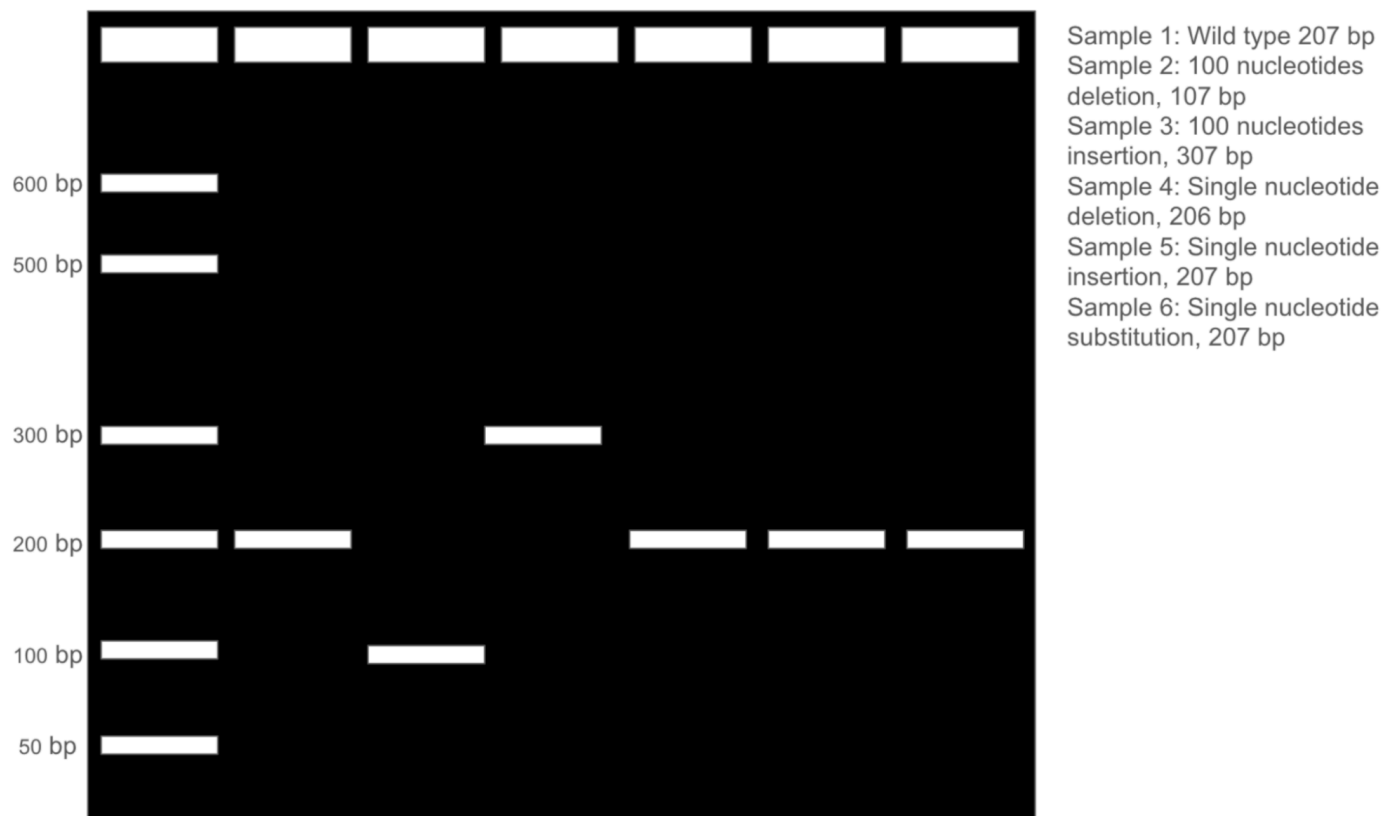


Figure 9: Simulated DNA gel electrophoresis to present potential outcomes of our CRISPR/Cas9 gene editing for training purposes.

Note. Lane 1: DNA molecular weight ladder /size marker (50–600bp). Lane 2 (Sample 1): Wild-type IL6 control showing a band at 207 bp. Lane 3 (Sample 2): Large-scale 100-nucleotide deletion resulting in a faster-migrating band at 107 bp. Lane 4 (Sample 3): Large-scale 100-nucleotide insertion resulting in a slower-migrating band at 307 bp. Lanes 5–7 (Samples 4–6): Single-nucleotide mutations (deletion, insertion, and substitution). Small-scale edits (206–207 bp) are visually indistinguishable from the wild-type control due to the resolution limits of standard agarose gel electrophoresis. Samples 4–6 are included to illustrate the resolution limits of agarose gel electrophoresis rather than to represent expected NHEJ

outcomes; a single-nucleotide substitution would not arise from NHEJ repair and would require a donor template or base editing. Sequencing would therefore be required to confirm the precise edit in any real application.

4. Discussion

Gastroenteritis remains a major public health problem worldwide, accounting for significant morbidity and healthcare burden, particularly in pediatric and elderly populations. In addition to host genetics, environmental factors play important roles in the transmission and progression of gastroenteritis; however, the molecular mechanisms underlying these effects remain incompletely understood. Environmental factors, including rainfall and temperature, are associated with gastroenteritis incidence (Chong et al., 2021), and temperature has been shown to act directly on IL-6 expression through the AhR pathway (Hu et al., 2024). In this study, I characterized the protein properties of IL-6 computationally and designed a gene-editing strategy that could be used in future work to test whether IL-6 is required for the epithelial inflammatory response to environmental and infectious challenge. The protein-protein interaction map of IL-6 provides insight into its key molecular interactors (Figure 6). The interactors that were found are IL6R - the primary binding subunit, IL6ST - the common signal-transducing subunit shared with other cytokines, and JAK1 and JAK2 - kinases activated upon receptor binding. The identification of these core interactors provides a molecular framework for understanding how environmental factors, particularly rainfall and temperature, may influence the clinical trajectory of gastroenteritis. Based on this interaction network, a hypothetical model can be outlined, which the present work does not test. Environmental factors may engage this axis through two distinct routes. Low temperature increases the stability and infectivity of enteric viruses, raising transmission and thus infection-driven IL-6 release (Chong et al., 2021), whereas elevated temperature induces IL-6 directly in host cells through AhR signaling, independent of infection (Hu et al., 2024). Either route converges on the same inflammatory module, and the other interactors identified in the map could modulate the outcome. It also shows how IL-6 signaling is connected to major inflammatory pathways that include: IL-1B, TNFRSF1A, IL-10, IL-2, and IL-3. The presence of IL-10 suggests that IL-6 signaling also engages negative feedback loops, which may become dysregulated under chronic or severe environmental stress. This analysis was done to understand its role as part of the immune system's integrated communication network.

The TMHMM analysis for IL-6 (Figure 3) provides insight into the subcellular localization of the protein, as the localization is important for its functions, which may be altered by disease-associated mutations. Therefore, it is essential to analyze the features of IL-6 in this regard. TMHMM results showed a hydrophobic region at the beginning of the protein, but this was confirmed as a signal peptide with another analysis, as the hydrophobic amino acids appeared at the very beginning of the protein, which is an indicator of being a signal peptide. (Figure 3). This analysis was limited to the transmembrane region, and alone it cannot identify the subcellular localization of a protein, but it can exclude localization within transmembrane regions because no transmembrane domain was predicted, as we know the absence of the transmembrane domain. The subcellular localization of a protein could indicate the type of function a protein could perform and the type of loss in the case that this localization is disrupted due to a mutation.

The Clustal Omega alignment was performed to assess which residues of IL-6 are conserved across species. Highly conserved amino acids are often functionally important, and this information is relevant both to interpreting the consequences of disruption and to selecting appropriate model organisms. Highly conserved amino acids are often functionally important, and mutations affecting these residues may have greater biological consequences. The Clustal Omega multiple protein alignment of IL-6 protein sequences across zebrafish, chicken, mouse, chimpanzee, pig, and human revealed that IL-6 is not highly



conserved across evolution among those selected species. Only a few amino acid positions showed complete conservation. These findings suggest that IL-6-associated molecular pathways have diverged across species rather than remaining highly conserved, and should be considered when selecting model organisms for future studies investigating IL-6 signaling.

I also designed a gene-editing strategy to develop a framework related to IL-6 and gastroenteritis; for this, CRISPR/Cas9 gene-editing methodology was applied. After retrieving the IL-6 human gene sequence, the annotated version of the gene was plotted in SnapGene, and a selected guide RNA via CRISPR Finder was added. The guide RNA was selected with stringent off-target criteria (0 mismatches = 1, 1 mismatch = 0, 2 mismatches = 0, and the lowest possible scores for 3 and 4 mismatches). This selection minimizes the risk of unintended edits at off-target genomic loci, which is a major concern in CRISPR-based experiments. This proposed strategy would enable researchers to generate IL6 knockout and observe the phenotypic consequences in cellular or animal models that may be exposed to environmental stressors such as low temperature and viral pathogens to study the development of gastroenteritis. To complete the proposed workflow, a DNA gel electrophoresis was simulated. The simulated DNA gel electrophoresis serves as a tool to demonstrate how different mutational outcomes would appear experimentally. Large-scale deletions (100 bp) produce a faster-migrating band at 107 bp; large-scale insertions produce a slower-migrating band at 307 bp; and single-nucleotide mutations (deletions, insertions, or substitutions) produce fragments that are visually indistinguishable from wild-type at 207 bp due to the resolution limits of standard agarose gel electrophoresis. In experimental practice, gene-editing outcomes would require verification using DNA sequencing methods.

One consideration is central to interpreting any such experiment. IL-6 is not straightforwardly pathogenic in the gut. It contributes to epithelial repair, supports the acute-phase response, and participates in controlling enteric pathogens, so its elevation during gastroenteritis reflects an active host response as well as a potential contributor to tissue damage. Removing IL-6 might therefore worsen outcomes rather than improve them. This is precisely why a knockout is informative as an experimental tool rather than a therapeutic proposition: it tests whether IL-6 is required for a given response, not whether reducing it would benefit patients.

An important gap in current gastroenteritis research is how emerging technologies such as CRISPR/Cas 9 could be incorporated to develop more physiologically relevant intestinal models that better mimic in vivo conditions (Zhang et al., 2026). This study demonstrates the feasibility of designing such a workflow and provides a framework for future experimental implementation. Future studies investigating how environmental factors influence IL-6-mediated inflammatory pathways in gastroenteritis may help address this knowledge gap and ultimately support the development of improved diagnostic and therapeutic strategies.

Although this study integrates bioinformatic analyses with a proposed CRISPR/Cas9 gene-editing strategy, its findings are based on computational analyses and literature evidence rather than experimental validation. Consequently, the predicted effects of IL6 modification on inflammatory signaling and gastroenteritis remain hypothetical until experimentally validated in physiologically relevant biological systems. Furthermore, considerations such as gene-editing efficiency, off-target effects, and delivery methods were beyond the scope of this study but will be important in future investigations.



5. Conclusion

In conclusion, gastroenteritis remains a critically significant global health challenge. IL-6 is elevated during gastroenteritis and is inducible by environmental conditions, which makes it a candidate point of intersection between the two, though this study does not establish that it functions as one. This study provides a molecular framework for understanding how IL-6-mediated pathways can be studied further and provides a foundation for future experimental investigations. Using bioinformatics tools, I have characterized the key protein properties of the IL-6 protein, mapped its core protein-protein interactions (IL6R, IL6ST, JAK1, JAK2, and connections to TNF, IL1B, and IL10), and designed a hypothetical CRISPR/Cas9 gene-editing strategy to enable future functional studies.

Future studies should now move from correlation to causation by experimentally testing whether IL-6 is required for the effects of environmental stressors on gastroenteritis-associated inflammation, using physiologically relevant models such as intestinal organoids or animal models.

6. References

- Abdipour, H., Assadian Narenji, M., Seyed Mousavi, S. E., Khammari, A., Dahghani Ghnataghestani, M., Alijani, E., Kamani, H., & Hosseinzehi, M. (2025). The relationship between biomarker interleukin-6 and particulate matter of dust storms (cohort study in 2024 – Zahedan city). *Toxicologie Analytique et Clinique*, 37(4), 504–512. <https://doi.org/10.1016/j.toxac.2025.05.001>
- Adumitrăchioaiei, H., Săsăran, M. O., & Mărginean, C. O. (2024). The diagnostic and prognostic role of interleukin 6 and interleukin 8 in childhood acute gastroenteritis—A review of the literature. *International Journal of Molecular Sciences*, 25(14), 7655. <https://doi.org/10.3390/ijms25147655>
- Ansori, A. N. M., Antonius, Y., Susilo, R. J. K., Hayaza, S., Kharisma, V. D., Parikesit, A. A., Zainul, R., Jakhmola, V., Saklani, T., Rebezov, M., Ullah, M. E., Maksimiuk, N., Derkho, M., & Burkov, P. (2023). Application of CRISPR-Cas9 genome editing technology in various fields: A review. *Narra J*, 3*(2), e184. <https://doi.org/10.52225/narra.v3i2.184>
- Birney, E., Andrews, T. D., Bevan, P., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cuff, J., Curwen, V., Cutts, T., Down, T., Eyra, E., Fernandez-Suarez, X. M., Gane, P., Gibbins, B., Gilbert, J., Hammond, M., Hotz, H.-R., Iyer, V., ... Clamp, M. (2004). An overview of Ensembl. *Genome Research*, 14(5), 925–928. <https://doi.org/10.1101/gr.1860604>
- Chen, S.-M., Lin, C.-P., Tsai, J.-D., Chao, Y.-H., & Sheu, J.-N. (2014). The significance of serum and fecal levels of interleukin-6 and interleukin-8 in hospitalized children with acute rotavirus and norovirus gastroenteritis. *Pediatrics & Neonatology*, 55(2), 120–126. <https://doi.org/10.1016/j.pedneo.2013.05.008>
- Chiejina, M., & Samant, H. (2026). Viral diarrhea. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK470525/>
- Chong, K. C., Chan, E. Y. Y., Lee, T. C., Kwok, K. L., Lau, S. Y. F., Wang, P., Lam, H. C. Y., Goggins, W. B., Mohammad, K. N., Leung, S. Y., & Chan, P. K. S. (2021). A 21-year retrospective analysis of environmental impacts on paediatric acute gastroenteritis in an affluent setting. *Science of The Total Environment*, 764, 142845.



<https://doi.org/10.1016/j.scitotenv.2020.142845>

National Institute of Diabetes and Digestive and Kidney Diseases. (2018). Diagnosis of viral gastroenteritis ("stomach flu"). U.S. Department of Health and Human Services.

Gastroenteritis Gene Set—HuGE Navigator Gene-Phenotype Associations. (n.d.). [Dataset]. Harmonizome. Retrieved July 15, 2026, from https://maayanlab.cloud/Harmonizome/gene_set/Gastroenteritis/HuGE+Navigator+Gene-Phenotype+Associations

Graves, N. S. (2013). Acute gastroenteritis. Primary Care: Clinics in Office Practice, *40*(3), 727–741. <https://doi.org/10.1016/j.pop.2013.05.006> HLA-DQA1 gene: MedlinePlus Genetics. (2026). <https://medlineplus.gov/genetics/gene/hla-dqa1/>

HLA-DQB1 gene: MedlinePlus Genetics. (2026). <https://medlineplus.gov/genetics/gene/hla-dqb1/>

Hu, X., Fu, H., Zhang, L., Zhang, Q., Xu, T., Chen, Y., Chen, G., Jiang, S., Ji, J., Xie, H. Q., Xu, L., & Zhao, B. (2024). Effect of elevated temperatures on inflammatory cytokine release: An in vitro and population-based study. *Environment & Health*, 2(10), 721–728. <https://doi.org/10.1021/envhealth.4c00059>

IL10 interleukin 10 [Homo sapiens (human)]—Gene—NCBI. (2026). <https://www.ncbi.nlm.nih.gov/gene/3586>

In brief: How does the immune system work? (2023). In *InformedHealth.org [Internet]*. Institute for Quality and Efficiency in Health Care (IQWiG). <https://www.ncbi.nlm.nih.gov/books/NBK279364/>

Khehra, N., Padda, I. S., & Zubair, M. (2026). Polymerase Chain Reaction (PCR). In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK589663/>

Lee, P. Y., Costumbrado, J., Hsu, C.-Y., & Kim, Y. H. (2012). Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments : JoVE*, (62), 3923. <https://doi.org/10.3791/3923>

Motohashi, K. (2019). Development of highly sensitive and low-cost DNA agarose gel electrophoresis detection systems, and evaluation of non-mutagenic and loading dye-type DNA-staining reagents. *PLOS ONE*, 14(9), e0222209. <https://doi.org/10.1371/journal.pone.0222209>

Nordquist, H., & Jamil, R. T. (2026). Biochemistry, HLA antigens. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK546662/>

Parisinos, C. A., Serghiou, S., Katsoulis, M., George, M. J., Patel, R. S., Hemingway, H., & Hingorani, A. D. (2018). Variation in interleukin 6 receptor gene associates with risk of Crohn's disease and ulcerative colitis. *Gastroenterology*, 155(2), 303–306.e2. <https://doi.org/10.1053/j.gastro.2018.05.022>

PRRT2 gene: MedlinePlus Genetics. (2026). <https://medlineplus.gov/genetics/gene/prrt2/>



- PubChem. (2026a). ABCC2—ATP binding cassette subfamily C member 2 (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/ABCC2/human>
- PubChem. (2026b). ABO - ABO, alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/ABO/human>
- PubChem. (2026c). CXCL8—C-X-C motif chemokine ligand 8 (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/CXCL8/human>
- PubChem. (2026d). FUT2—Fucosyltransferase 2 (H blood group) (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/FUT2/human>
- PubChem. (2026e). IFNG - interferon gamma (human). <https://pubchem.ncbi.nlm.nih.gov/gene/IFNG/human>
- PubChem. (2026f). IL1B - interleukin 1 beta (human). <https://pubchem.ncbi.nlm.nih.gov/gene/IL1B/human>
- PubChem. (2026g). IL1RN - interleukin 1 receptor antagonist (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/IL1RN/human>
- PubChem. (2026h). IL18—Interleukin 18 (human). <https://pubchem.ncbi.nlm.nih.gov/gene/IL18/human>
- PubChem. (2026i). IMPDH1—Inosine monophosphate dehydrogenase 1 (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/IMPDH1/human>
- PubChem. (2026j). IMPDH2—Inosine monophosphate dehydrogenase 2 (human).
<https://pubchem.ncbi.nlm.nih.gov/gene/IMPDH2/human>
- Redman, M., King, A., Watson, C., & King, D. (2016). What is CRISPR/Cas9? *Archives of Disease in Childhood. Education and Practice Edition*, 101(4), 213–215. <https://doi.org/10.1136/archdischild-2016-310459>
- Sievers, F., & Higgins, D. G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. In D. J. Russell (Ed.), *Multiple Sequence Alignment Methods* (pp. 105–116). Humana Press.
https://doi.org/10.1007/978-1-62703-646-7_6
- SnapGene Software Application. (n.d.). Retrieved March 1, 2026, from <https://www.snapgene.com/>
- NCBI Staff. (2023, February 23). What is NCBI and who works here? NCBI Insights.
<https://ncbiinsights.ncbi.nlm.nih.gov/2023/02/23/what-and-who-ncbi/>
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(Database issue), D607–D613. <https://doi.org/10.1093/nar/gky1131>
- Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von



- Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, 40(7), 1023–1025. <https://doi.org/10.1038/s41587-021-01156-3>
- TNF tumor necrosis factor [*Homo sapiens* (human)]—Gene—NCBI. (2026). <https://www.ncbi.nlm.nih.gov/gene/7124>
- TNFRSF11B gene: MedlinePlus Genetics. (2026). <https://medlineplus.gov/genetics/gene/tnfrsf11b/>
- Mayo Clinic Staff. (2025, April 29). Viral gastroenteritis (stomach flu) - Diagnosis & treatment. Mayo Clinic. <https://www.mayoclinic.org/diseases-conditions/viral-gastroenteritis/diagnosis-treatment/drc-20378852>
- Wimalagunasekara, S. S., Weeraman, J. W. J. K., Tirimanne, S., & Fernando, P. C. (2023). Protein-protein interaction (PPI) network analysis reveals important hub proteins and sub-network modules for root development in rice (*Oryza sativa*). *Journal of Genetic Engineering & Biotechnology*, 21, 69. <https://doi.org/10.1186/s43141-023-00515-8>
- Yang, H., Ren, S., Yu, S., Pan, H., Li, T., Ge, S., Zhang, J., & Xia, N. (2020). Methods favoring homology-directed repair choice in response to CRISPR/Cas9 induced-double strand breaks. *International Journal of Molecular Sciences*, 21(18), 6461. <https://doi.org/10.3390/ijms21186461>
- Zhang, K., Li, H., & Zhou, Y. (2026). Progress in human intestinal organoid research: Applications to acute gastroenteritis viruses. *Stem Cell Research & Therapy*, 17(1), 156. <https://doi.org/10.1186/s13287-026-04953-2>

Code and Availability Statement

All data analyzed in this study were obtained from publicly available databases. The human IL6 gene and IL-6 protein sequences were retrieved from NCBI Gene (<https://www.ncbi.nlm.nih.gov/gene/3569>) and NCBI Protein (UniProt accession P05231). Exon and intron annotations were obtained from Ensembl (<https://www.ensembl.org>). Protein-protein interaction data were retrieved from the STRING database (<https://string-db.org>). Guide RNA candidates were identified using CRISPR Finder (<https://www.sanger.ac.uk/tool/crispr-finder/>). No custom code or scripts were developed for this study. All analyses were performed using the web-based tools cited in the Methods section, and all parameters used are described therein. The author confirms that no new datasets were generated.

Acknowledgements

I thank my mentor for her guidance throughout the paper. I also thank my family and friends who supported me during this time and before and after.

Author Biography

Cherly Chan is a Y13 student at Harrow International School Hong Kong. She is interested in pursuing a career in healthcare and technology. Outside of academics, she enjoys playing the guzheng and badminton.



Mentor Contribution Statement

Zeynep Öztürk, PhD, Postdoctoral Fellow at the University of Oxford, Nuffield Department of Medicine, supported the author throughout this paper. She discussed and helped refine the scope of the paper, taught the author the fundamentals of cell biology and genetics, and introduced the bioinformatics databases used in the analysis. The gene-editing strategy was designed by the author under Dr. Öztürk's direct guidance, and she provided feedback on drafts of the writing. The author selected IL-6 as the focus gene and conducted the in silico analyses.

