Convergence
A Journal for Young Researchers

Received: [for official use only] ; Accepted: [for official use only];
Managing Editor: [for official use only];
Action Editor: [for official use only]

Assessing the development of gastroenteritis via the IL-6-mediated pathway and designing CRISPR/Cas9 gene editing

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 proposes that IL-6 genetic variation modifies the effect of environmental stressors on gastroenteritis outcomes and establishes a molecular framework and gene-editing strategy to test this hypothesis. In silico analyses were conducted using TMHMM, SignalP, Clustal Omega, and the STRING database to investigate the properties of the IL-6 protein. TMHMM revealed that IL-6 contains no transmembrane domains, indicating it is not membrane-bound. SignalP identified a 27-amino-acid signal peptide that is cleaved during secretion. Clustal Omega alignment across six species showed that IL-6 is not highly conserved, suggesting that genetic variation may contribute to differences in disease susceptibility between individuals. The STRING protein-protein interaction map identified core interactors, including IL6R, IL6ST/gp130, JAK1, and JAK2, with additional connections to IL1B, TNFRSF1A, IL10, IL2, and IL3, placing IL-6 at the center of an interconnected inflammatory module. A CRISPR/Cas9 design targeting an exon of IL-6 was generated using CRISPR Finder and SnapGene with stringent off-target criteria, and simulated DNA gel electrophoresis was used to illustrate the range of possible editing outcomes. Overall, this study provides a molecular framework for investigating the contribution of IL-6-mediated inflammation to gastroenteritis across various environmental factors and offers a gene-editing strategy for future studies to test whether IL-6 variation modifies environmental effects on disease outcomes.

Keywords: gastroenteritis, IL6, environmental factors, temperature, CRISPR/Cas9, inflammation

1. Introduction

Gastroenteritis is the inflammation of the stomach and large or small 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, this is required for a diagnosis (*Diagnosis of Viral Gastroenteritis ("Stomach Flu") - NIDDK*, n.d.). This disease occurs when the immune system activates to defend the stomach and intestinal tract against infection but fails. 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. Therefore, this paper investigates the role of IL-6-mediated inflammatory signaling in the development and progression of gastroenteritis under different environmental influences.

There are different types of gastroenteritis. Firstly, bacterial gastroenteritis, which is an infection of the intestines and stomach, 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. Infection is closely linked to the environment; therefore, I look at IL-6-mediated gastroenteritis developments together with environmental factors.

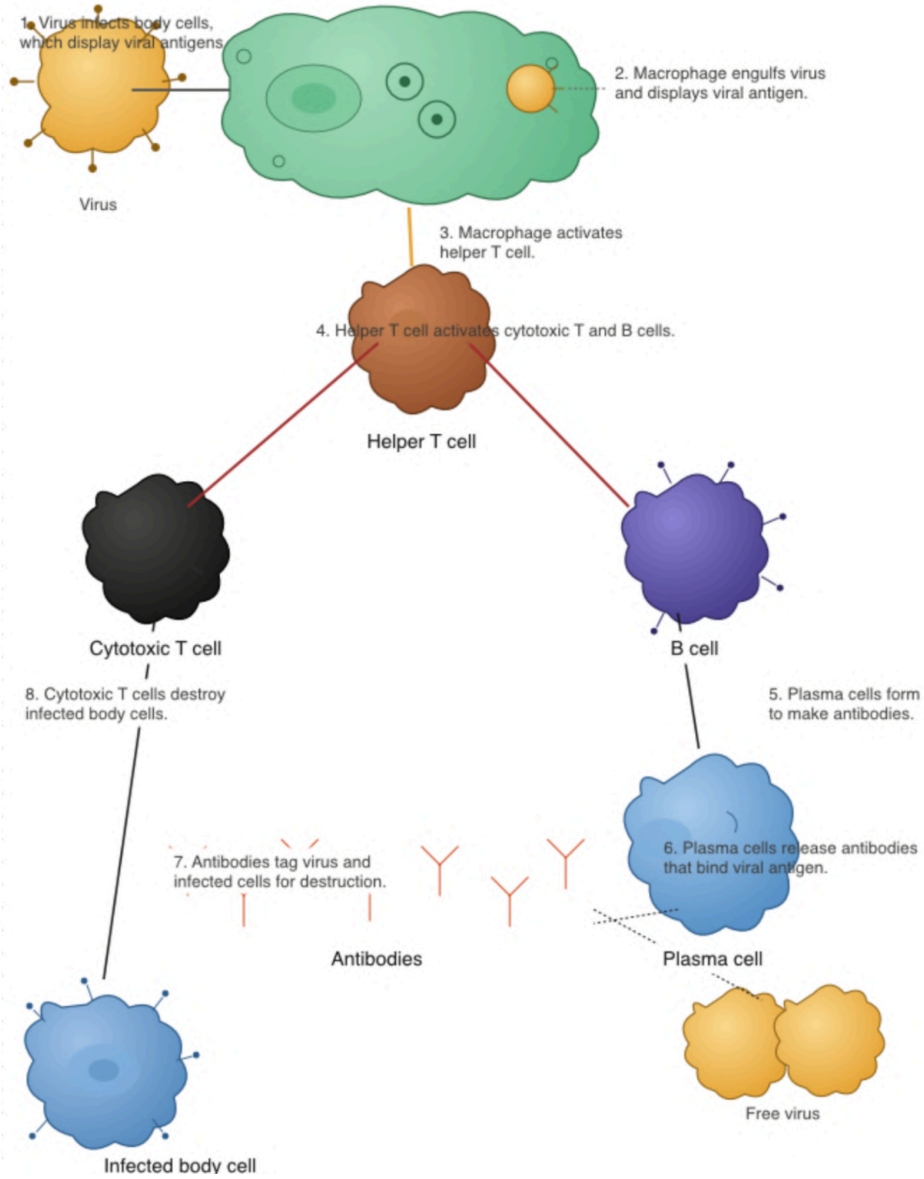


Figure 1. Steps that show the viral infection cycle linked to gastroenteritis. The diagram describes the steps in how the immune system tackles infections. Each of these has specialized responsibilities and activities. The immune system fights pathogens (bacteria and viruses) and removes them from the body, recognizes and neutralizes toxins, and fights disease-causing changes in the body. It is activated by antigens, and when antigens attach to immune system cells (steps 2-8), a series of processes are triggered in the body. Once the body has encountered a pathogen for the first time, it stores information about how to fight it, so that a subsequent encounter with the same pathogen can be recognized and fought more quickly ("In Brief," 2023).

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).

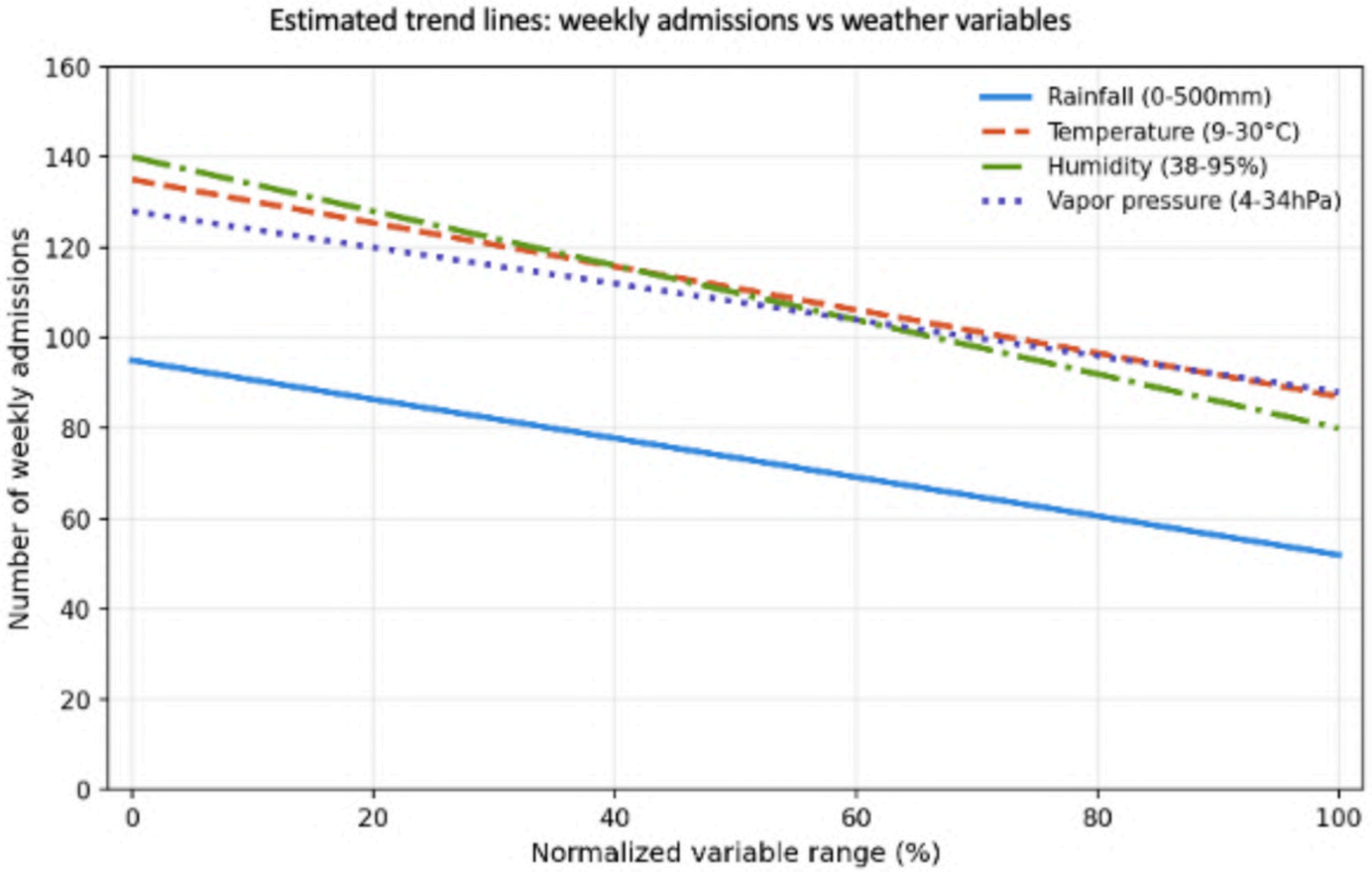


Figure 2. Statistics of pediatric patients admitted to hospital for gastrointestinal infection due to environmental factors (weather variables). A graph showing the weekly number of pediatric gastrointestinal infection admissions against different meteorological factors such as rainfall, temperature, 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 variable 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 generated with data provided by (Chong et al., 2021).

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 support the central hypothesis of this paper, 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 named osteoclasts, which break down bone tissue during bone remodeling (*TNFRSF11B Gene*, 2026).

I selected to focus on the IL-6 gene for the scope of this paper in order to understand and develop a research framework to study gastroenteritis in the future. The IL6 gene produces 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). The encoded protein has proven to be an endogenous pyrogen that can induce fever in people with autoimmune infections or diseases. IL-6 contributes to the pathogenesis of acute gastroenteritis caused by rotavirus and norovirus, and is released in response to gastrointestinal infection (Chen et al., 2014). In the context of acute gastroenteritis, both serum and fecal levels of IL-6 show significant elevation compared to healthy individuals. Critically, IL-6 levels are significantly higher in bacterial gastroenteritis than in viral cases, giving it strong diagnostic value as a biomarker to help differentiate between etiologies. IL-6 demonstrates superior sensitivity, specificity, and positive predictive value compared to C-reactive protein (CRP) and IL-8 in distinguishing bacterial from viral causes. Furthermore, increased IL-6 levels correlate with disease severity and complications, including benign seizures, complications from Shigella infection, and increased liver enzymes (ALT/AST) in rotavirus gastroenteritis. Hence, IL-6 serves not only as a diagnostic marker for etiological differentiation but also as a prognostic indicator of disease severity in pediatric acute gastroenteritis (Adumitrăchioaiei et al., 2024). Therefore, IL-6 is a good candidate to focus on for the investigation related to contributions of genetics and environmental factors to gastroenteritis.

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 IL6 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 evaluate the role of IL-6 in the development and progression of gastroenteritis in the context of environmental and genetic influences. This is achieved by investigating the structural and functional properties of the IL6 protein and proposing a CRISPR/Cas9 gene-editing strategy. Together, these analyses establish a framework for future studies investigating how environmental factors and genetic variation may influence IL-6-mediated inflammatory pathways in gastroenteritis.

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). I used the NCBI>protein and NCBI>gene sections to retrieve protein and gene sequences.

Ensembl: A bioinformatics project with a purpose to organize biological information around the sequences of large genomes (Birney et al., 2004). I used Ensembl 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. I used this analysis to search whether IL-6 has any transmembrane region.

SignalP: A widely used computational tool for predicting signal peptides from amino acid sequences (Teufel et al., 2022). I retrieved the IL6 protein sequence from the NCBI protein database in FASTA format and submitted it to the SignalP algorithm for analysis.

Clustal Omega: An algorithm for multiple sequence alignment that can deal with very large numbers of DNA/RNA due to its use of the mBED algorithm for calculating guide trees (Sievers & Higgins, 2014). I used Clustal Omega analysis to look for the degree of conservation of the IL6 protein across different 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). I used this analysis to look for the interactors of IL-6.

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 (Ansori et al., 2023).

SnapGene: A multipurpose software program used for designing and documenting sequence-based experiments and gene editing. I used it for DNA sequence alignment, annotation, editing, and visualizing the IL6 gene to present the location of the guide RNA and PCR primers for the gene editing design of IL-6 (*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. I searched the gene of interest and selected an exon, and then the algorithm produced the potential guide RNAs in a table. Then I looked for the off-target scores for 0, 1, 2, 3, and 4 mismatches. I aimed to make 0=1, 1=0, 2=0, and for 3,4 the lowest possible scores.

2. Results

3.1 IL-6 does not contain any 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 about their function/s 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
TCNKSNMCESSKEALAENNLNLPKMAEKDGCFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSEEQA
RAVQMSTKVLIQFLQKKAKNLDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKEFLQSSLRALR
```

```
# 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
```

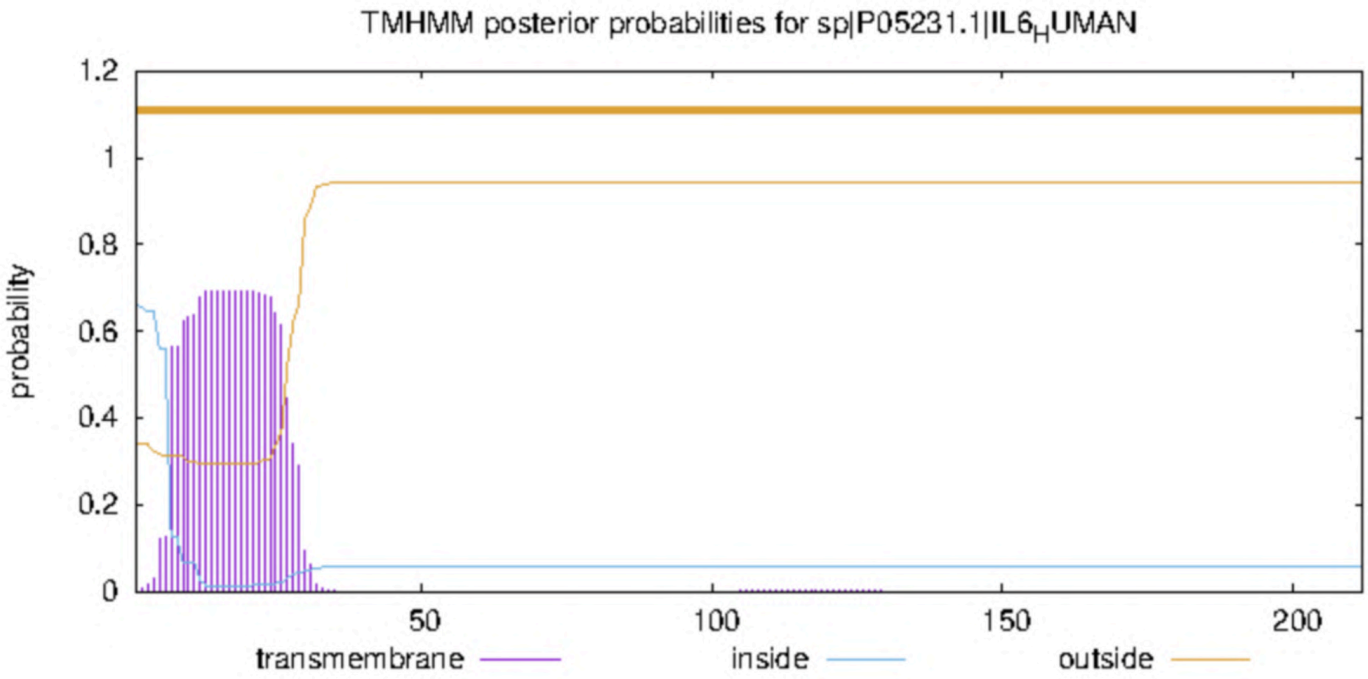


Figure 3. TMHMM transmembrane domain prediction result for the IL-6 protein. The upper text is the protein amino acid sequence retrieved from the NCBI Protein database. The bottom text and plot are the result of TMHMM analysis for IL-6. The text of the TMHMM result shows the amino acid length and the 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 for it 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.

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

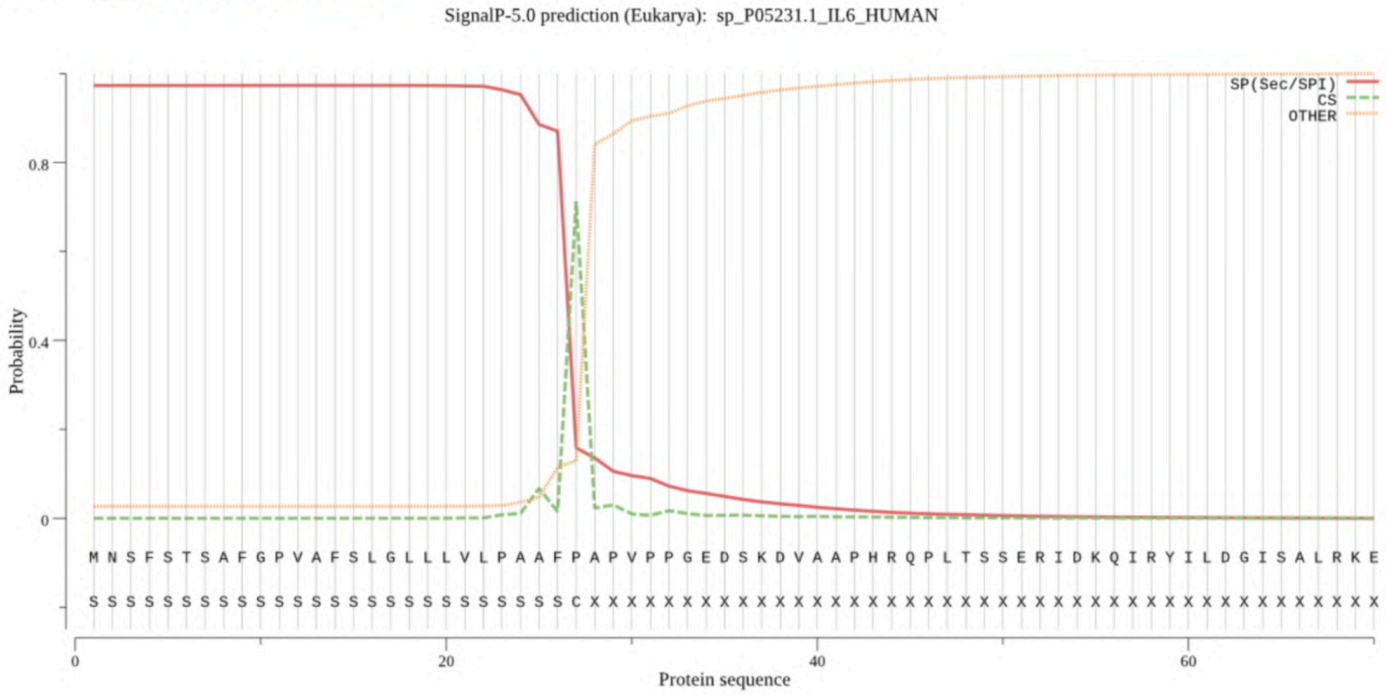


Figure 4. SignalP 5.0 analysis of IL-6. The plot shows the prediction of the presence of a signal peptide in IL6. The x-axis shows the amino acid sequence of IL-6; the y-axis shows the probability calculation between 0 and 1. The red line represents the amino acids that are signal peptides, the green 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.

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 IL6 protein, with the purpose of finding out

which parts of the IL6 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 IL6-related immune mechanisms is not conserved across the species put in this analysis.

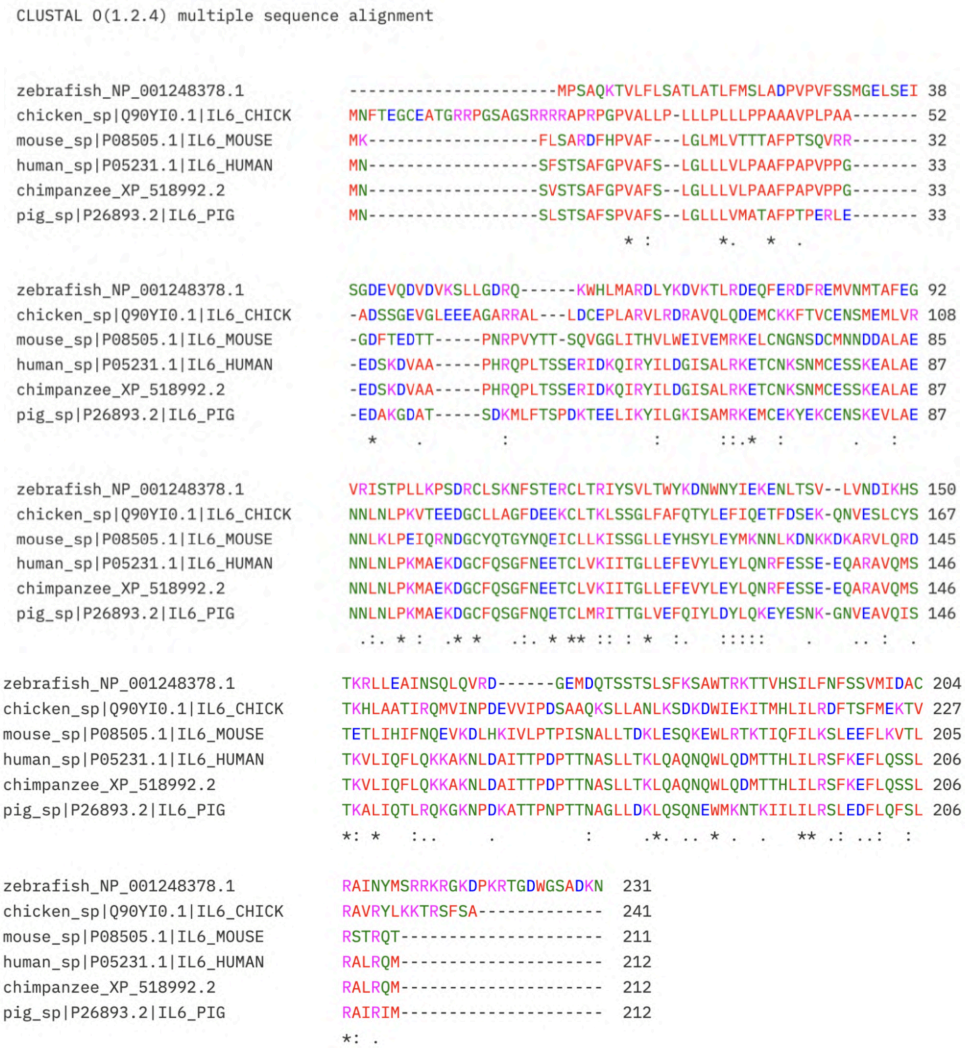


Figure 5. Multiple protein alignment of IL6 across different species. The alignment shows the species names on the left side and the amino acids aligned on the right side. There is a amino acid coloring scheme (ClustalX convention): which means RED: Small, hydrophobic amino

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 the IL6

Protein-protein IL6 interaction 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, 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 a 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). This analysis means that IL-6 is not merely a component of the inflammatory module, but rather a central node within it.

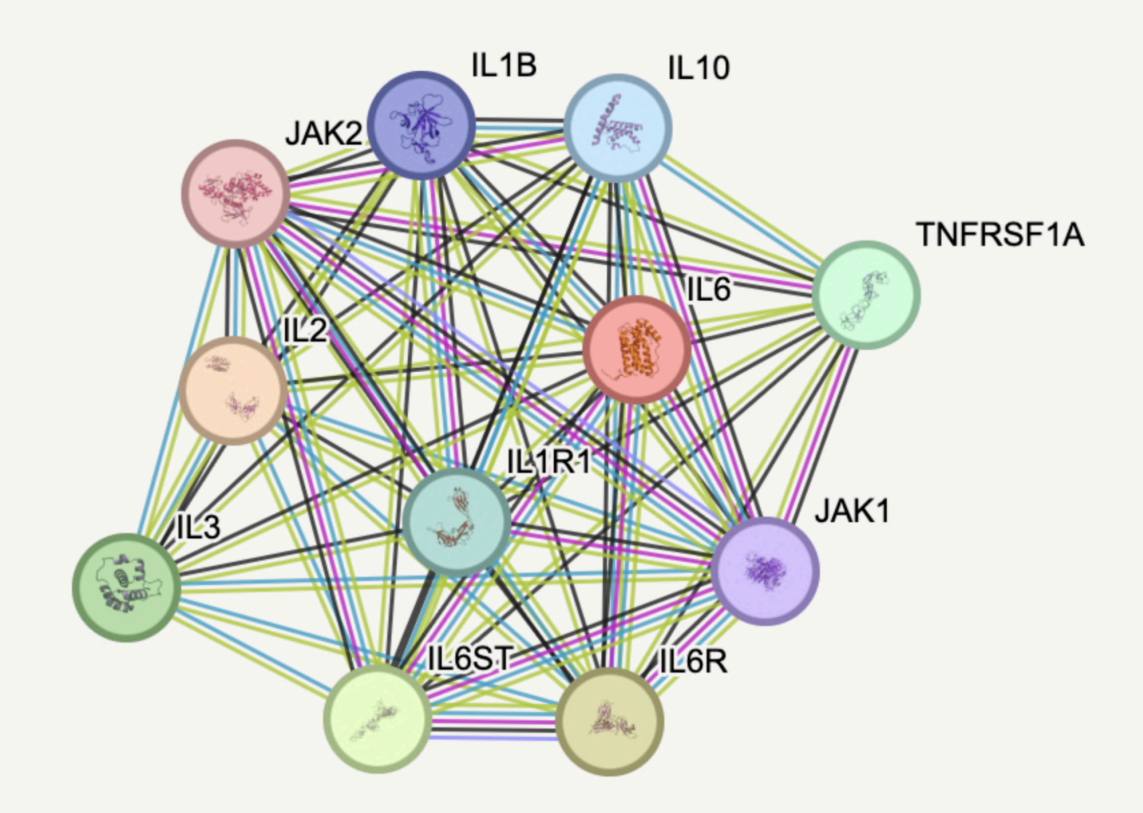


Figure 6. Protein-protein interaction map of IL6. The map shows the proteins predicted to interact with IL-6, generated using the STRING database. Nodes represent proteins: IL6 (center), its receptor subunits IL6R and IL6ST, the associated kinases JAK1 and JAK2, and further inflammatory mediators IL1B, 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: turquoise, 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 IL-6

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. Here, I proposed a gene editing strategy for generating mutants of IL-6 to study the effects of environmental factors combined with different genetic mutations on developing gastroenteritis in the future.

For the gene editing design, NCBI>GENE, Ensembl, CRISPR Finder (Sanger), and SnapGene were used. The gene sequence of IL-6 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 IL-6 on the NCBI gene database was checked to be sure that IL-6 does not collide with different genes (Figure 7). The gene sequence annotation was done by looking at Ensembl exon and intron annotations. Later, the guide RNA was found through the CRISPR Finder database, and the guide RNA that was selected had the best possible off-target scores: 0=1, 1=0, 2=0, 3 and 4= as low as possible. The guide and the PAM sequences were also marked on SnapGene Viewer (Figure 8). Then the PCR primers were designed for hypothetical follow-up experiments for training purposes. The precise outcomes need to be verified via sequencing in the actual lab practice.

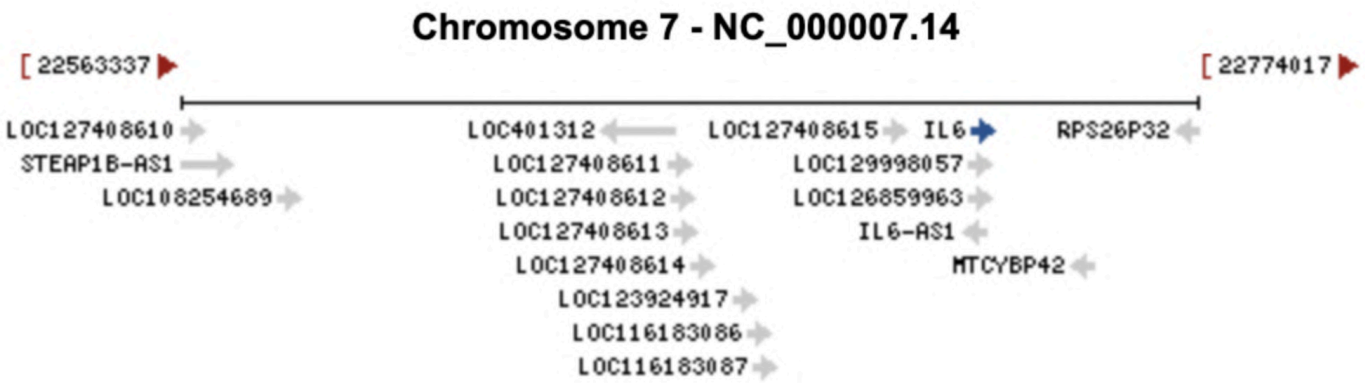


Figure 7. Chromosomal Location of IL-6. The genomic region spanning positions 22,563,337 to 22,774,017 on chromosome 7 is shown. The orientation of each gene is indicated by arrows. The purpose of this map is to identify the IL6 genomic location and guide the design of the CRISPR/Cas9 gene-editing targets within the IL6 locus.

Exons for ENST00000404625

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

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

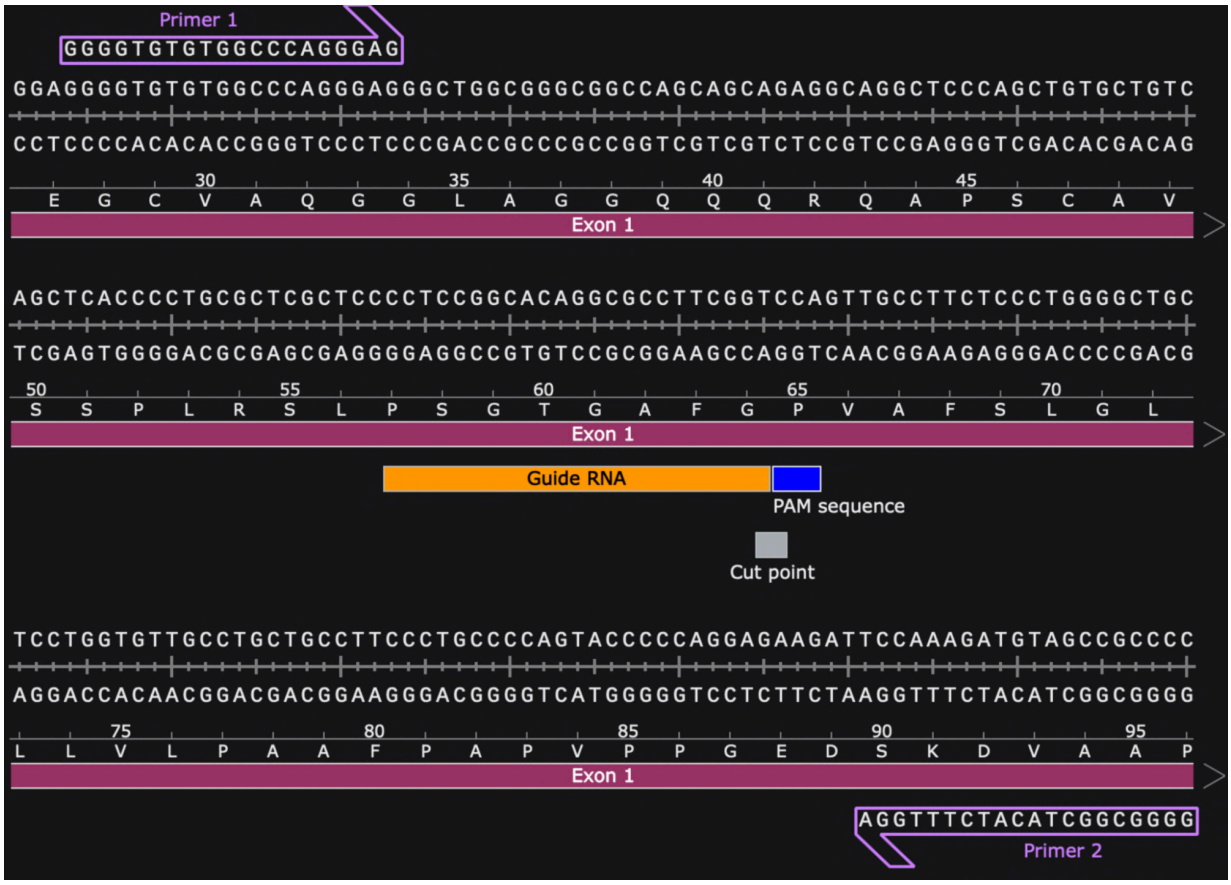


Figure 8. SnapGene view of CRISPR/Cas9 gene editing design for IL6. The upper image is the list of exons for IL6 in the CRISPR Finder database. The one highlighted is the selected exon to do gene editing. The middle image shows the selected Guide RNA (orange in Snapgene view below) and the off-target scores (the very 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 cut point for the PAM sequence (grey), 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.

The proposed strategy enables a specific follow-up experiment. Intestinal epithelial cells carrying the IL-6 edits 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). IL-6 induction would be quantified at the transcript level by qPCR and at the protein level by ELISA and compared across genotypes. If IL-6 genetic variation modifies the effect of thermal stress, the magnitude of induction should differ between edited and wild-type cells. Equivalent induction across genotypes would argue against IL-6 variation as a modifier of the environmental effect and would redirect attention to the other interactors identified in the STRING network (Figure 6).

3.6 Simulated DNA gel electrophoresis

The CRISPR/Cas9 gene-editing strategy targeting the IL-6 gene to study its role in the development of gastroenteritis was hypothetically designed, and PCR followed by DNA gel electrophoresis experiments were proposed and simulated as follow-up experiments. Different mutations in IL-6 could affect disease development in different ways, and these effects can vary when combined with environmental factors. 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.

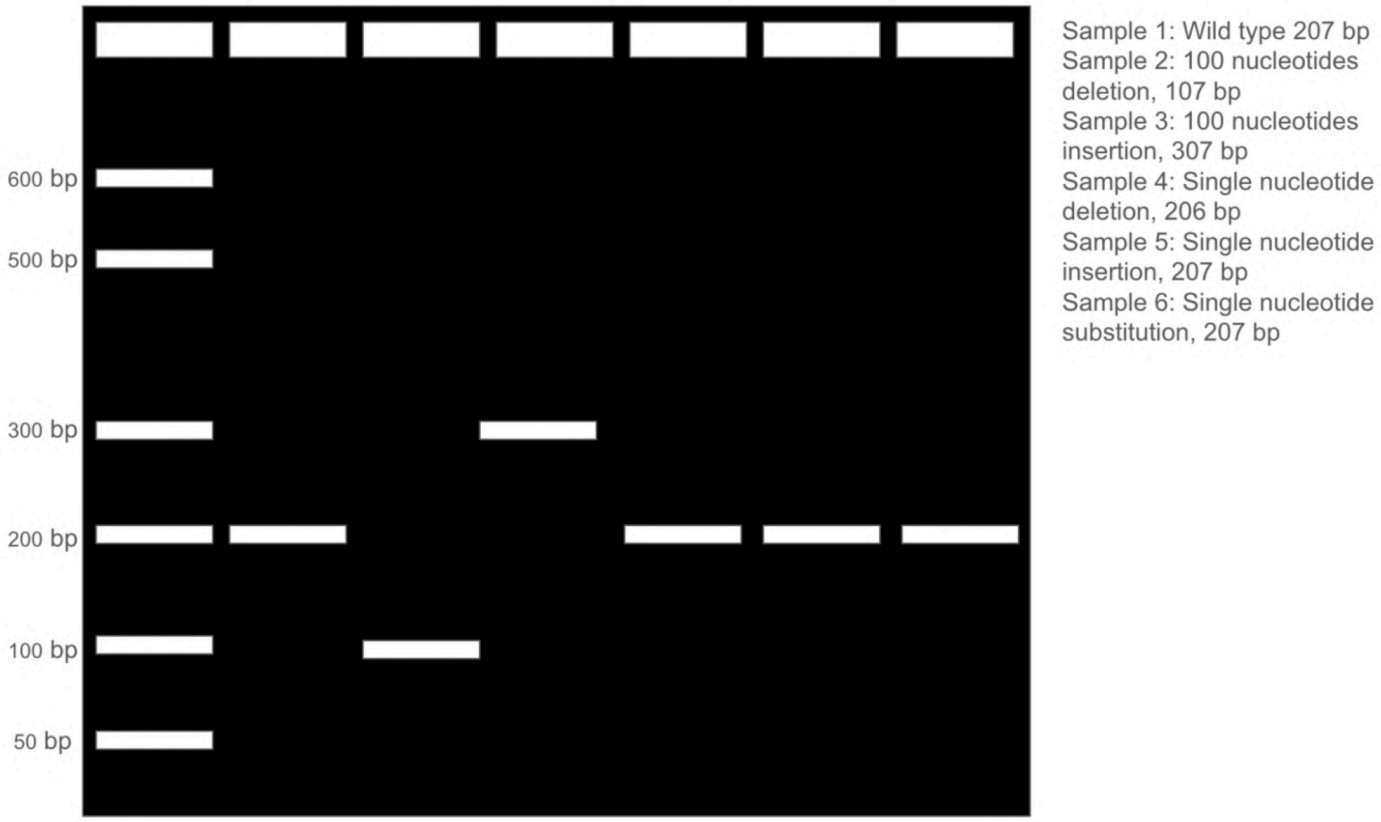


Figure 9. Simulated DNA gel electrophoresis to present potential outcomes of our CRISPR/Cas9 gene editing for training purposes. Lane 1: DNA molecular weight ladder /size marker (50-600bp). Lane 2 (Sample 1): Wild-type IL6 control showing a band at 207 bp. Lane 3 (Sample 2): Large-scale 100-nucleotide 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.

3. 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 established a framework by analyzing the protein properties of IL-6 and designing a gene-editing strategy to investigate the combined effects of environmental factors and IL-6 mutations in gastroenteritis. 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, 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 biological model can be proposed to explain how these molecular components may influence disease development. 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 IL6 signaling is connected to major inflammatory pathways that include: IL1B, TNFRSF1A, IL10, IL2, and IL3. The presence of IL10 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 a 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.

There are underlying genetic factors, particularly involving the IL-6 gene, that could explain the diversity seen among populations despite the shared environmental conditions for developing gastroenteritis. I wanted to see whether the IL-6 protein is a conserved protein across animals and which amino acids are conserved evolutionarily. 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 could enable researchers to introduce targeted mutations into the IL-6 gene 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.

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.

4. Conclusion

In conclusion, gastroenteritis remains a critically significant global health challenge. The IL-6-mediated inflammatory pathway plays an important role in the inflammatory response associated with gastroenteritis and may be influenced by environmental factors. 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 genetic variation modifies the effects of environmental stressors on gastroenteritis using physiologically relevant models such as intestinal organoids or animal models.

5. 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. NM., Antonius, Y., Susilo, R. JK., Hayaza, S., Kharisma, V. D., Parikesit, A. A., Zainul, R., Jakhmola, V., Saklani, T., Rebezov, M., Ullah, Md. 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., Eyras, 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

Diagnosis of Viral Gastroenteritis ("Stomach Flu")—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved April 3, 2026, from https://www.niddk.nih.gov/health-information/digestive-diseases/viral-gastroenteritis/diagnosis

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*, 40(3), 727. 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/

Staff, N. (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/

Viral gastroenteritis (stomach flu)—Diagnosis & treatment. (n.d.). Mayo Clinic. Retrieved December 30, 2025, from 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

Acknowledgements

[redacted]

Author Biography

[redacted]

Mentor Contribution Statement

[redacted]

EDITOR RECOMMENDATION:
The reviewers commend the author for an ambitious and compelling paper, and have several suggestions for ways to improve the strength of the manuscript. As the reviews are quite detailed, I would focus efforts on ways to improve the text to appropriately temper conclusions where noted, and to add clarity to the methods and their strengths/limitations.

REVIEWER 1:

This manuscript addresses an interesting and potentially relevant question regarding the relationship between IL-6, environmental factors, and inflammatory responses in the context of gastroenteritis. The author has incorporated several useful bioinformatic analyses, including protein characterization, protein–protein interaction analysis, evolutionary sequence comparison, and a proposed CRISPR/Cas9 gene-editing strategy. The manuscript also demonstrates the authors' ability to integrate computational approaches with potential experimental applications. In particular, the proposed investigation of temperature-dependent effects on IL-6 is an interesting direction that could potentially provide a basis for future experimental studies.

However, the overall scope and interpretation of the manuscript should be revised to better match the analyses actually performed. At present, the manuscript sometimes extends the interpretation beyond what can be directly supported by the current analysis. In particular, IL-6 is occasionally presented as a demonstrated mediator of gastroenteritis development and progression under environmental stress, whereas the current study primarily consists of bioinformatic characterization and a hypothetical CRISPR/Cas9 gene-editing strategy.

The manuscript would be substantially strengthened by clearly distinguishing between findings supported by the current analyses and potential mechanisms or future experimental directions. Statements regarding environmental factors, IL-6 signaling, and gastroenteritis pathogenesis should therefore be appropriately toned down throughout the Introduction, Results, Discussion, and Conclusion. The central research question and stated objectives should also be narrowed so that they accurately reflect what was actually investigated in the present study.

Overall, I believe the manuscript has a useful foundation and could become a clearer and more focused study after revision.

Figure 2: The rationale for including this figure should be clarified. In particular, the connection between the environmental factors discussed in the manuscript and IL-6 signaling is currently weak. The authors should provide a clearer biological rationale explaining how and why these environmental factors are expected to influence IL-6-related pathways and gastroenteritis. Also, is this figure adapted or reproduced from a previously published paper? If so, please clarify whether permission to use the figure has been obtained.

Line 153: The statement that the IL-6 gene was selected "to understand and develop a research framework to study gastroenteritis in the future" is problematic. The study did not experimentally investigate gastroenteritis or establish such a framework. Future applications should not be presented as the primary objective or main thesis of the current study. The authors should revise this statement to clearly distinguish between what was actually investigated in the present study and what is proposed as a potential future direction.

Line 177: The stated aim of the study—to evaluate the role of IL-6 in the development and progression of gastroenteritis under environmental and genetic influences—does not appear to

be supported by the actual analyses presented in the manuscript. The main analyses focus on the molecular characteristics of IL-6 and a proposed gene-editing strategy, rather than directly investigating gastroenteritis. The aim should therefore be revised to more accurately reflect the scope and results of the current study, while any potential relevance to gastroenteritis should be presented as a future direction rather than as a demonstrated finding.

Methods: The Methods section should be revised to use passive voice consistently.

Multiple protein alignment of IL-6 among different species: The rationale for including this analysis is unclear. Please explain how the interspecies conservation of IL-6 is directly relevant to the research question and its relationship to gastroenteritis.

Section 3.5 – CRISPR/Cas9 Gene Editing Strategy for IL-6:
The idea of investigating temperature sensitivity in relation to IL-6 is potentially interesting; however, the current section reads more like a research proposal or future experimental plan than a completed research study. A research paper should have a clearly defined research question and an analysis that provides its own scientific contribution, rather than primarily proposing what could be done in the future. More importantly, the intended CRISPR/Cas9 editing strategy is unclear. It is not specified whether the authors aim to generate an IL-6 knockout, a specific point mutation, or an insertion/deletion mutation. These are fundamentally different experimental objectives and would require different CRISPR/Cas9 designs and validation strategies. The authors should clearly define the intended genetic modification and provide a specific rationale for why that particular modification is relevant to the research question. If the purpose is only to propose a potential future experiment, this should be clearly presented as a future direction rather than as a central component of the current study.

3.6 Simulated DNA gel electrophoresis
Please clarify what type of IL-6 mutation you are intending to generate: a deletion, an insertion, or a specific point mutation. These are different types of genetic modifications and require different CRISPR/Cas9 editing strategies and experimental designs. Therefore, you should first clearly define the intended mutation and its biological purpose. The CRISPR/Cas9 design, expected editing outcome, PCR strategy, and validation method should then be developed based on that specific editing objective. At present, the manuscript appears to describe several possible outcomes without clearly specifying which mutation is actually being generated. Please revise this section so that the proposed CRISPR/Cas9 strategy leads to one clearly defined genetic modification rather than presenting multiple hypothetical possibilities.

Recommendation: Accept with major revisions

REVIEWER 2:

This is an ambitious and creative manuscript that attempts to connect environmental epidemiology, cytokine biology, bioinformatics and CRISPR/Cas9 experimental design. For a high-school researcher, the range of concepts and computational tools used is impressive. The author has made a serious effort to build a testable research framework, and the inclusion of TMHMM, SignalP, sequence alignment, STRING analysis and guide-RNA design demonstrates considerable intellectual curiosity and engagement.

The manuscript also appropriately acknowledges that the CRISPR strategy and gel electrophoresis are hypothetical and that experimental validation would be required. These are important strengths. However, the central biological claims currently extend beyond what the presented analyses can demonstrate. The manuscript would become substantially stronger if it were reframed as a hypothesis-generating, in-silico framework, rather than an investigation that has already established an IL-6-mediated relationship between environmental conditions and gastroenteritis.

Major comments

1. The title, aims and conclusions should more accurately reflect the study performed

The manuscript does not directly assess the development or progression of gastroenteritis, nor does it perform CRISPR editing. Instead, it characterises selected properties of IL-6 computationally and proposes a future gene-editing workflow. Statements such as “this paper investigates the role of IL-6” and “IL-6 genetic variation modifies the effect of environmental stressors” therefore sound more definitive than the evidence allows, where they should say that the study develops a framework for future investigation of this possibility. A more accurate title would be something like: Bioinformatic characterisation of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis.

2. Association, causation and mediation require clearer separation

The manuscript correctly reports that IL-6 can be elevated in acute gastroenteritis and may correlate with disease severity. However, elevated IL-6 does not necessarily cause the disease to become more severe. More severe infection and tissue injury may instead produce a greater inflammatory response and consequently higher IL-6 levels. Similarly, the fact that IL-6 is elevated after environmental exposure and during gastroenteritis does not demonstrate that IL-6 mediates the relationship between the two. IL-6 is a relatively nonspecific inflammatory cytokine and may be independently increased in both settings.

To demonstrate mediation, evidence would be needed for the complete sequence: Environmental exposure → change in IL-6 → change in a gastroenteritis outcome. This would require showing that the environmental exposure precedes IL-6 induction and that experimentally altering IL-6 subsequently changes outcomes such as pathogen burden, epithelial injury, barrier function, symptom severity or recovery, which requires experimental validation.

The central claim should therefore be framed as: ‘Environmental exposures can induce IL-6 and may consequently modify gastroenteritis-associated inflammation.’ It should not yet be framed as: ‘Environmental exposures modify gastroenteritis outcomes through IL-6.’

3. The rationale for selecting IL-6 and the interpretation of genetic variation should be strengthened

The Introduction lists 17 genes associated with gastroenteritis and then describes each individually, but the transition to IL6 remains abrupt. The author should add a concise

explanation of why IL6 was selected over genes with established pathogen-specific relevance, such as FUT2.

Regarding the cross-species alignment, this does not demonstrate that human IL6 genetic variation contributes to differences in gastroenteritis susceptibility. Sequence divergence between zebrafish, chickens, rodents and humans is not equivalent to variation among human individuals. Nor does low sequence identity necessarily mean that IL-6 function or its broader pathway is not conserved.

A more defensible interpretation would be: The alignment identifies conserved and divergent residues and may help assess which species are appropriate models for human IL-6 biology.

If the manuscript wishes to make claims about human susceptibility, it should analyse actual human variants, such as IL6 SNPs, allele frequencies, eQTL evidence or variants previously associated with inflammatory or infectious phenotypes. Otherwise, the statement linking the alignment to differences among patients should be more cautiously worded.

4. The STRING network analysis is worded in a circular logic

The claim that IL-6 is a “central node” is also partly circular because IL-6 was entered as the query protein and its first-shell associations were displayed. Network centrality would require a defined network and a quantitative centrality analysis. For the sake of transparency, the STRING version, organism, confidence threshold, number of interactors and evidence settings should be reported.

5. The STRING network analysis figure legend appears incorrect

The Figure 6 colour legend also appears incorrect. In STRING evidence mode, yellow represents text-mining, black represents co-expression, blue represents gene co-occurrence, purple represents experimental evidence and light blue represents curated database evidence.

6. The CRISPR strategy needs a more clearly defined biological purpose

The manuscript alternates between discussing “genetic variation,” “mutations” and IL6 disruption, but these are not equivalent experimental goals. This is especially relevant, since Cas9 cleavage followed by non-homologous end joining generally creates a mixture of insertions and deletions and is most suitable for disrupting or knocking out a gene. It would not accurately reproduce a specific naturally occurring IL6 variant. Modelling a defined human variant would instead require a specified donor-based HDR strategy, base editing or prime editing.

The author should therefore state whether the intended experiment is an IL6 knockout experiment, introduction of a particular human variant, or generation of nonspecific mutants as a preliminary functional screen.

7. The proposed future experiment does not yet test gastroenteritis outcomes

The proposed experiment compares IL-6 transcript and protein levels between edited and wild-type intestinal epithelial cells at 37°C and 38.5°C. This could test whether the edit changes temperature-associated IL-6 expression, but it would not test whether IL-6 modifies gastroenteritis.

In future works sections, A future experiment addressing the central hypothesis would need an infectious challenge and relevant endpoints. For example, edited and control intestinal organoids could be exposed to an enteric pathogen, to measure variables like pathogen replication or burden, epithelial barrier integrity, cell death and viability, inflammatory cytokine production, etc.

8. Positioning of IL6 in the manuscript comes across ‘harmful’

The manuscript must also recognise that IL-6 is not simply harmful. It can contribute to inflammation, but it also participates in host defence, epithelial survival and intestinal repair. Experimental IL-6 deficiency has worsened pathogen control and epithelial injury in some intestinal infection models. Complete suppression could therefore produce undesirable consequences rather than improve disease. This dual role should be discussed before presenting IL6 editing as a potentially beneficial intervention.

9. Figure 1 needs revision

The figure provides a useful simplified overview of the antiviral immune response, but several labels should be revised before publication. Dendritic cells, rather than macrophages alone, should be shown as the main cells initiating T-cell activation. The figure should also clarify that helper T cells support, rather than directly activate, antigen-specific cytotoxic T cells and B cells. Finally, antibodies mainly bind and neutralise free virus, whereas cytotoxic T cells destroy infected body cells. Finally, some of the lines have been out of order as well as the Y shapes of the antibodies.

Minor comments

1. Section numbering is inconsistent: the Results heading is numbered “2,” while its subsections begin with 3.1; the Discussion is then numbered 3.
2. Gene and protein nomenclature should be standardised. IL6 should refer to the gene, whereas IL-6 should refer to the protein.
3. “Autoimmune infections or diseases” should be corrected to “infections or autoimmune diseases.”
4. The statement that gastroenteritis occurs because the immune system “activates but fails” is inaccurate. Infection and symptoms result from pathogen-associated injury together with the host response; the presence of disease does not necessarily mean that the immune system has failed.
5. Statements describing IL-6 as having “strong diagnostic value” should be qualified as findings reported in particular cohorts, rather than accepted clinical performance across all gastroenteritis populations.
6. The manuscript overall would benefit from removal of repeated definitions.
7. The position of the cut marker in Figure 8 should be checked. For SpCas9, the expected cleavage position is approximately three nucleotides upstream of the PAM, within the guide-matched protospacer. The caption should describe this as the predicted Cas9 cleavage site, not the “cut point for the PAM sequence.”

Final recommendation

This is a promising and unusually ambitious project for a high-school researcher. The author has assembled several relevant computational approaches and has developed a thoughtful starting point for future experimental work. The most important next step is not to add more analyses, but to ensure that the claims precisely match what the current evidence demonstrates. With clearer separation of association from causation, correction of the figures and CRISPR details, and reframing as an in-silico research framework, the manuscript could become a strong and educational contribution suitable for a young-researcher journal.

Convergence
A Journal for Young Researchers

Received: [for official use only] ; Accepted: [for official use only];
Managing Editor: [for official use only];
Action Editor: [for official use only]

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

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 In silico analyses were conducted using TMHMM, SignalP, Clustal Omega, and the STRING database to investigate the properties of the IL-6 protein. TMHMM SignalP identified a 27-amino-acid signal peptide that is cleaved during secretion. Clustal Omega alignment across six species The STRING interactors, including IL6R, IL6ST/gp130, JAK1, and JAK2, IL1B, TNFRSF1A, IL10, IL2, and IL3 A CRISPR/Cas9 his study provides a

Keywords: gastroenteritis, IL6, environmental factors, temperature, CRISPR/Cas9, inflammation

1. Introduction

Gastroenteritis is the inflammation of the stomach and large or small 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, this is required for a diagnosis (*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, 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.

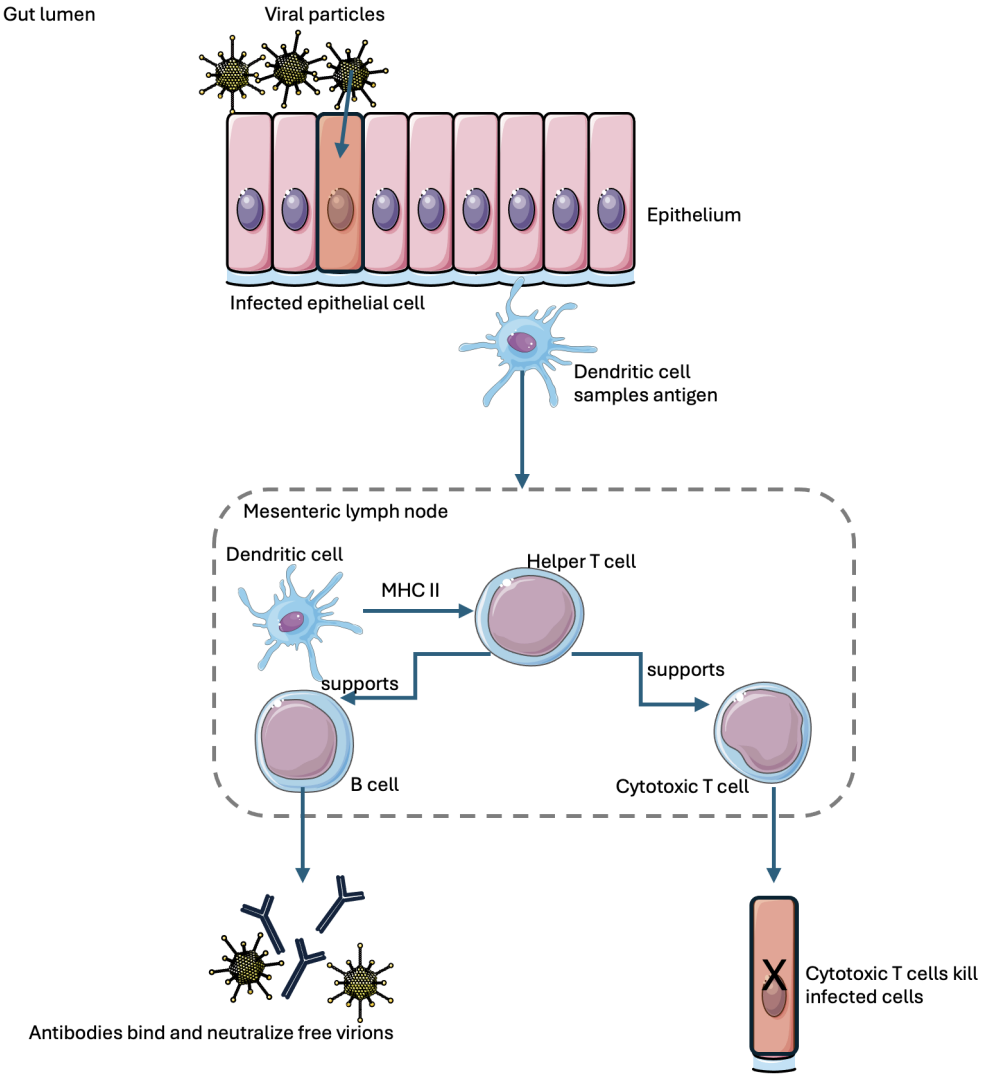


Figure 1. Immune response to enteric viral infection. Viral particles in the gut lumen infect intestinal epithelial cells. Dendritic cells sample viral antigen and migrate to the mesenteric lymph node, where antigen is presented on MHC II to helper T cells. Activated helper T cells support the responses 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.

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.

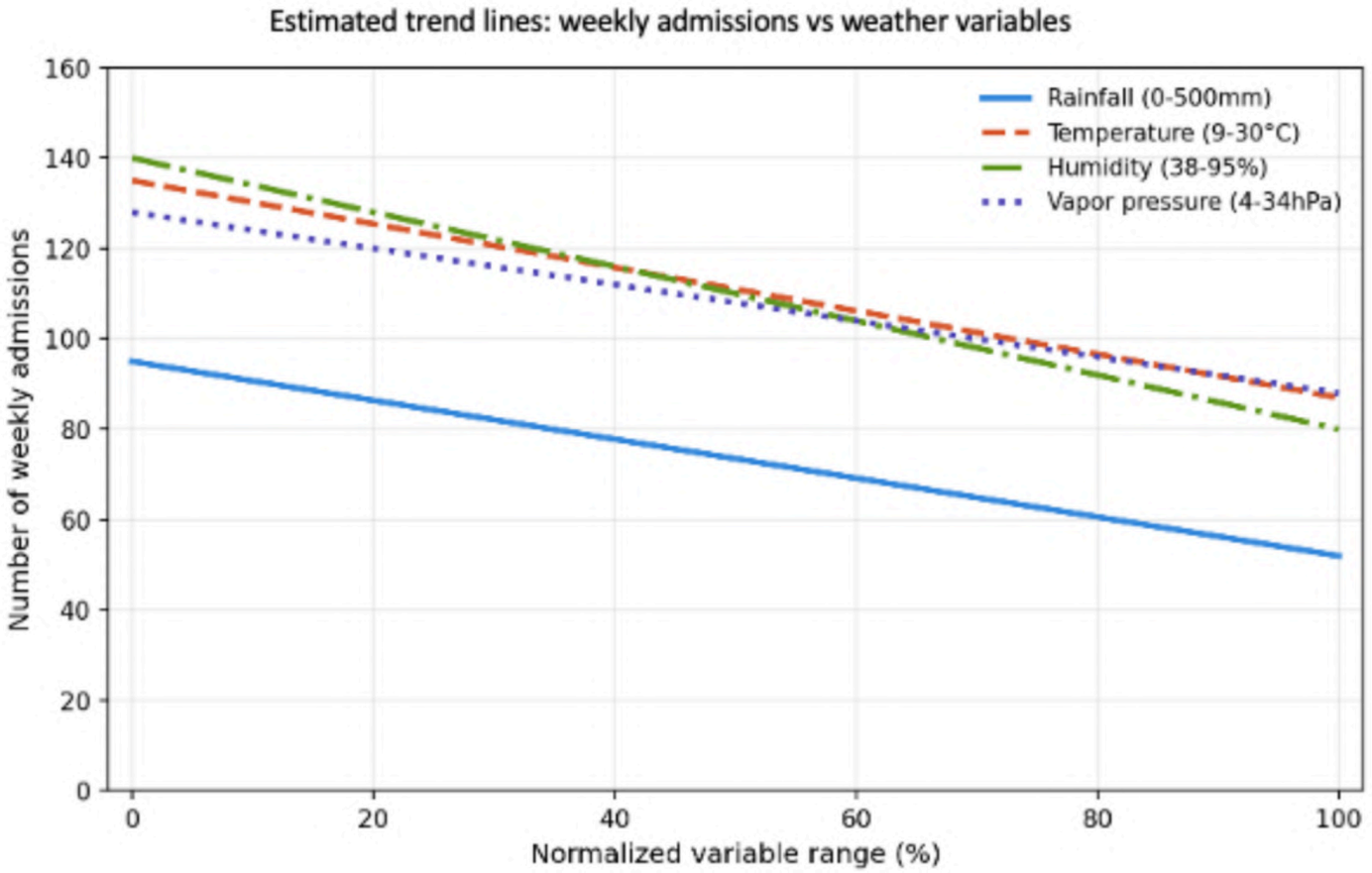


Figure 2. Statistics of pediatric patients admitted to hospital for gastrointestinal infection due to environmental factors (weather variables). A graph showing the weekly number of pediatric gastrointestinal infection admissions against different meteorological factors such as rainfall, temperature, 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 variable 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 named 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 norivirus 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 IL6 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 IL6 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 a purpose to organize 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. This analysis was performed using STRING v12.0 to identify predicted interactors of IL-6

SignalP: A widely used computational tool for predicting signal peptides from amino acid sequences (Teufel et al., 2022). The IL6 protein sequence was retrived 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 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: textmining, 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 IL6 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 on the basis of 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 any 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 about their function/s 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
TCNKSNMCESSKEALAENNLNLPKMAEKDGCFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSEEQA
RAVQMSTKVLIQFLQKKAKNLDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKEFLQSSLRALR
```

```
# 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
```

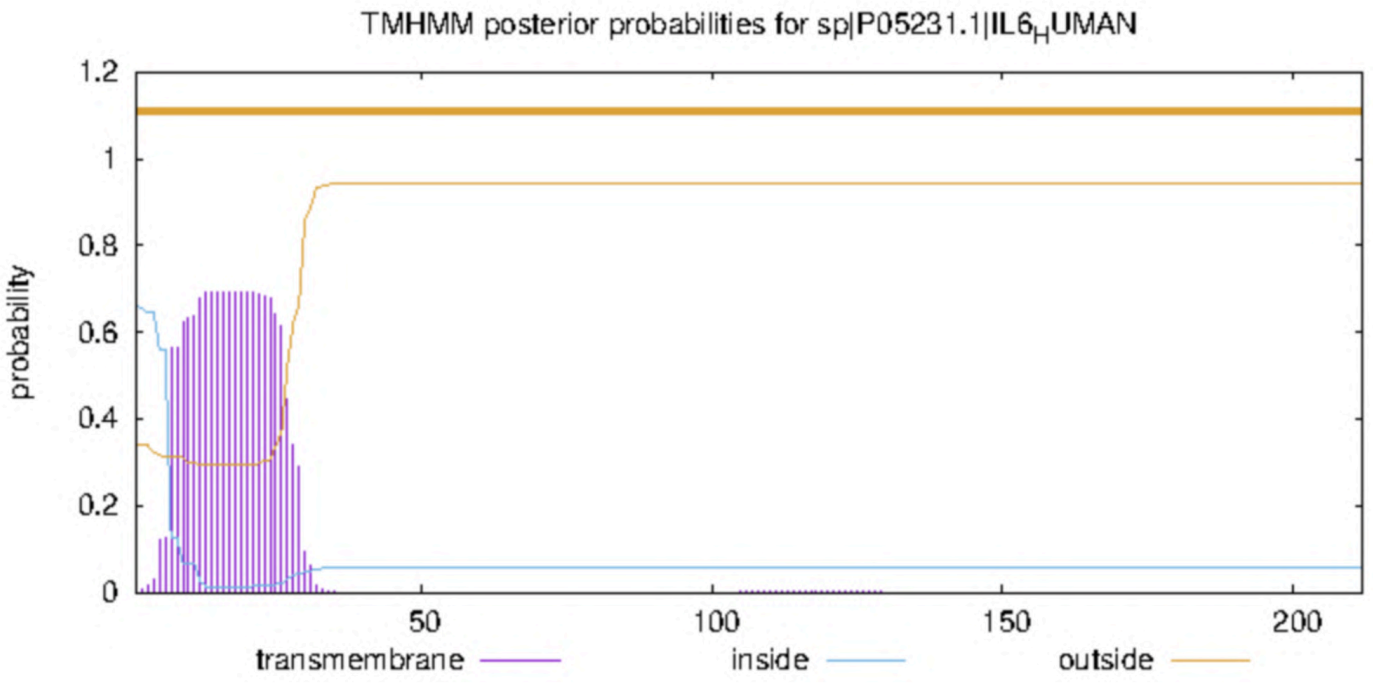


Figure 3. TMHMM transmembrane domain prediction result for the IL-6 protein. The upper text is the protein amino acid sequence retrieved from the NCBI Protein database. The bottom text and plot are the result of TMHMM analysis for IL-6. The text of the TMHMM result shows the amino acid length and the 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 for it 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.

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

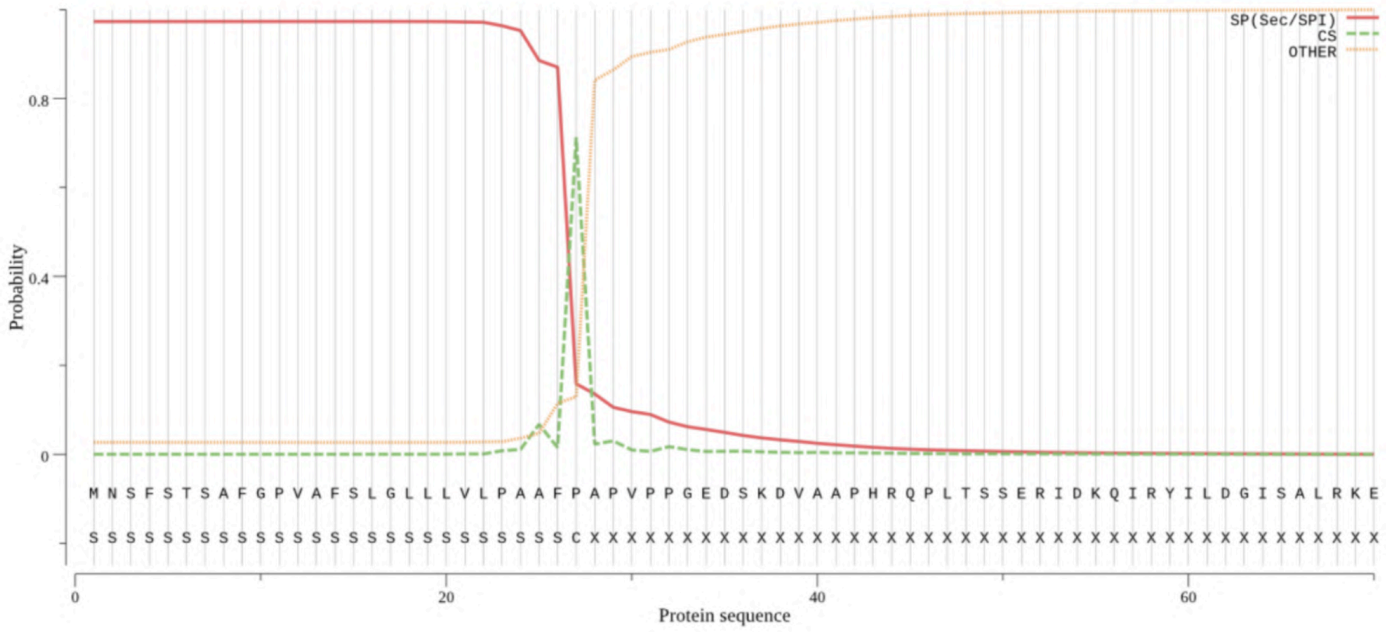


Figure 4. SignalP 5.0 analysis of IL-6. The plot shows the prediction of the presence of a signal peptide in IL6. The x-axis shows the amino acid sequence of IL-6; the y-axis shows the probability calculation between 0 and 1. The red line represents the amino acids that are signal peptides, the green 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.

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 IL6 protein, with the purpose of finding out which parts of the IL6 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 IL6-related immune mechanisms is not conserved across the species put in this analysis.

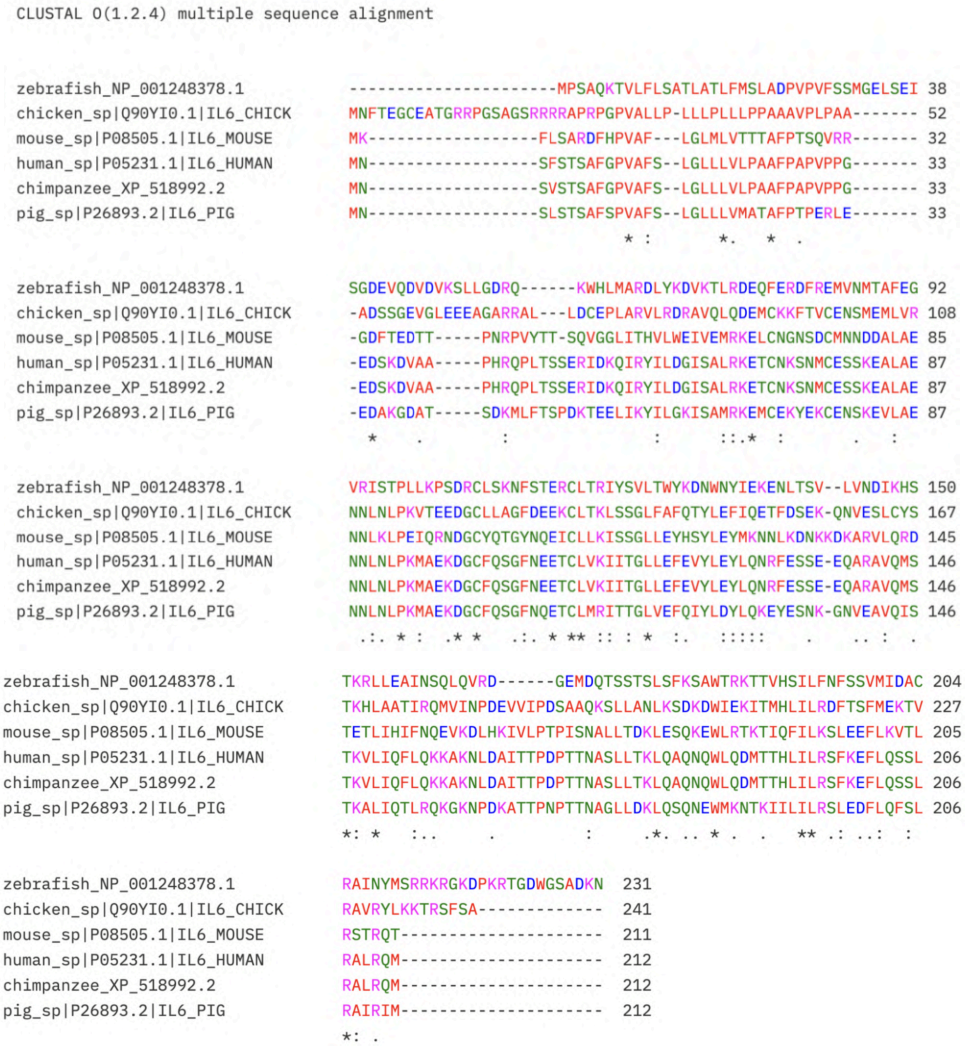


Figure 5. Multiple protein alignment of IL6 across different species. The alignment shows the species names on the left side and the amino acids aligned on the right side. There is a amino acid coloring scheme (ClustalX convention): which means RED: Small, hydrophobic amino acids (A, V, L, I, M, F, W), 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 the IL6

Protein-protein IL6 interaction 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, 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 a 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.

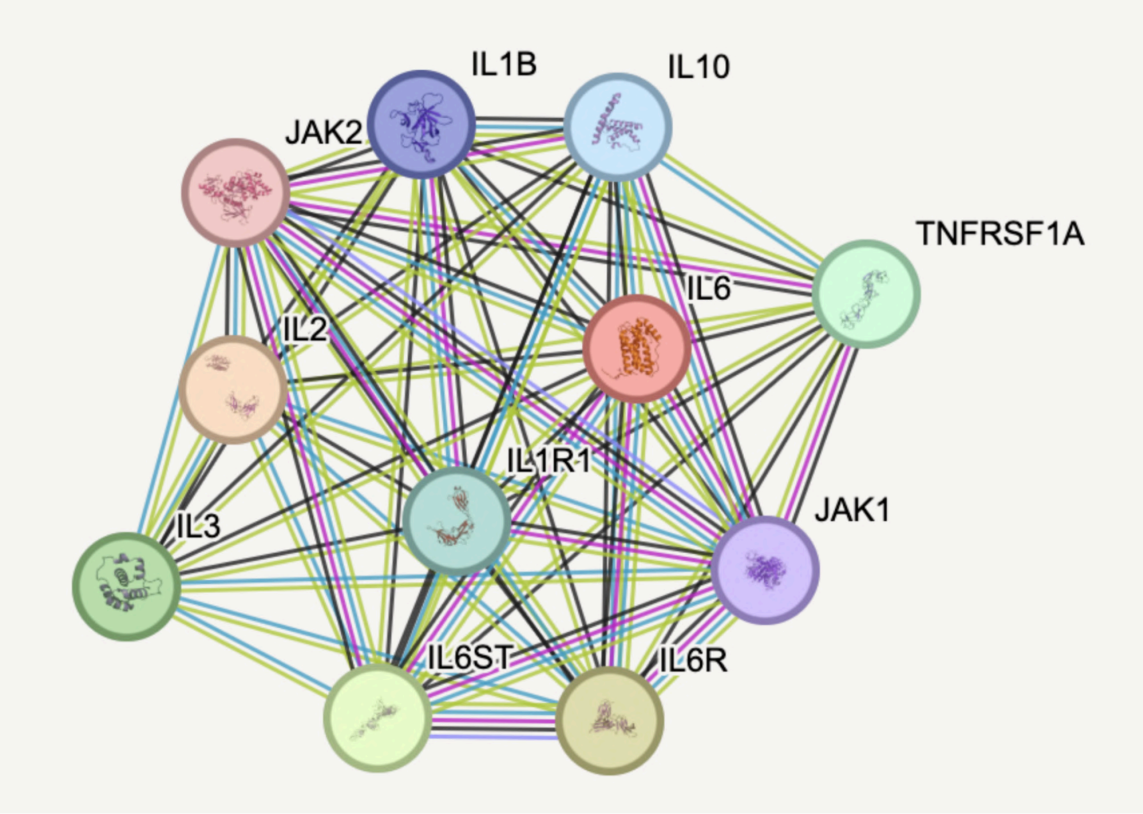


Figure 6. Protein-protein interaction map of IL-6. The map shows the proteins predicted to interact with IL-6, generated using the STRING database. Nodes represent proteins: IL6 (center), its receptor subunits IL6R and IL6ST, the associated kinases JAK1 and JAK2, and further inflammatory mediators IL1B, 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 IL-6 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 the actual laboratory practice.

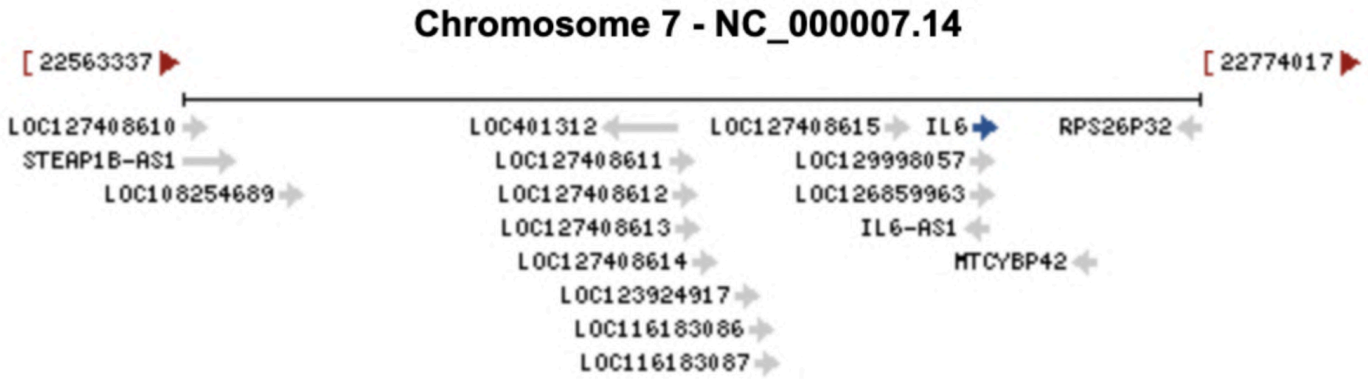


Figure 7. Chromosomal Location of IL-6. The genomic region spanning positions 22,563,337 to 22,774,017 on chromosome 7 is shown. The orientation of each gene is indicated by arrows. The purpose of this map is to identify the IL6 genomic location and guide the design of the CRISPR/Cas9 gene-editing targets within the IL6 locus.

and gastroenteritis: a CRISPR/Cas9 framework

Exons for ENST00000404625

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

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

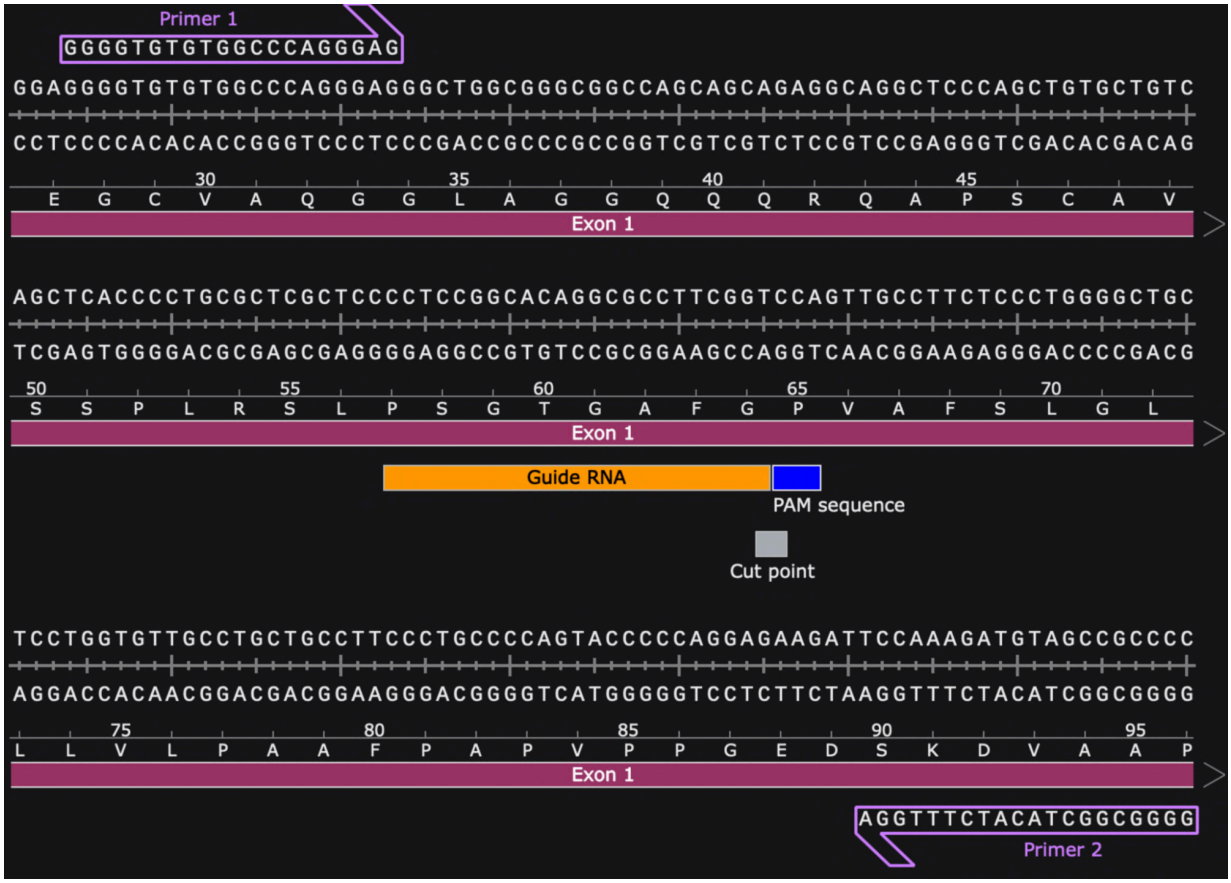


Figure 8. SnapGene view of CRISPR/Cas9 gene editing design for IL6. The upper image is the list of exons for IL6 in the CRISPR Finder database. The one highlighted is the selected exon to do gene editing. The middle image shows the selected Guide RNA (orange in Snapgene view below) and the off-target scores (the very 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.

The proposed strategy enables a specific follow-up experiment. Intestinal epithelial cells carrying the IL-6 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).

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.

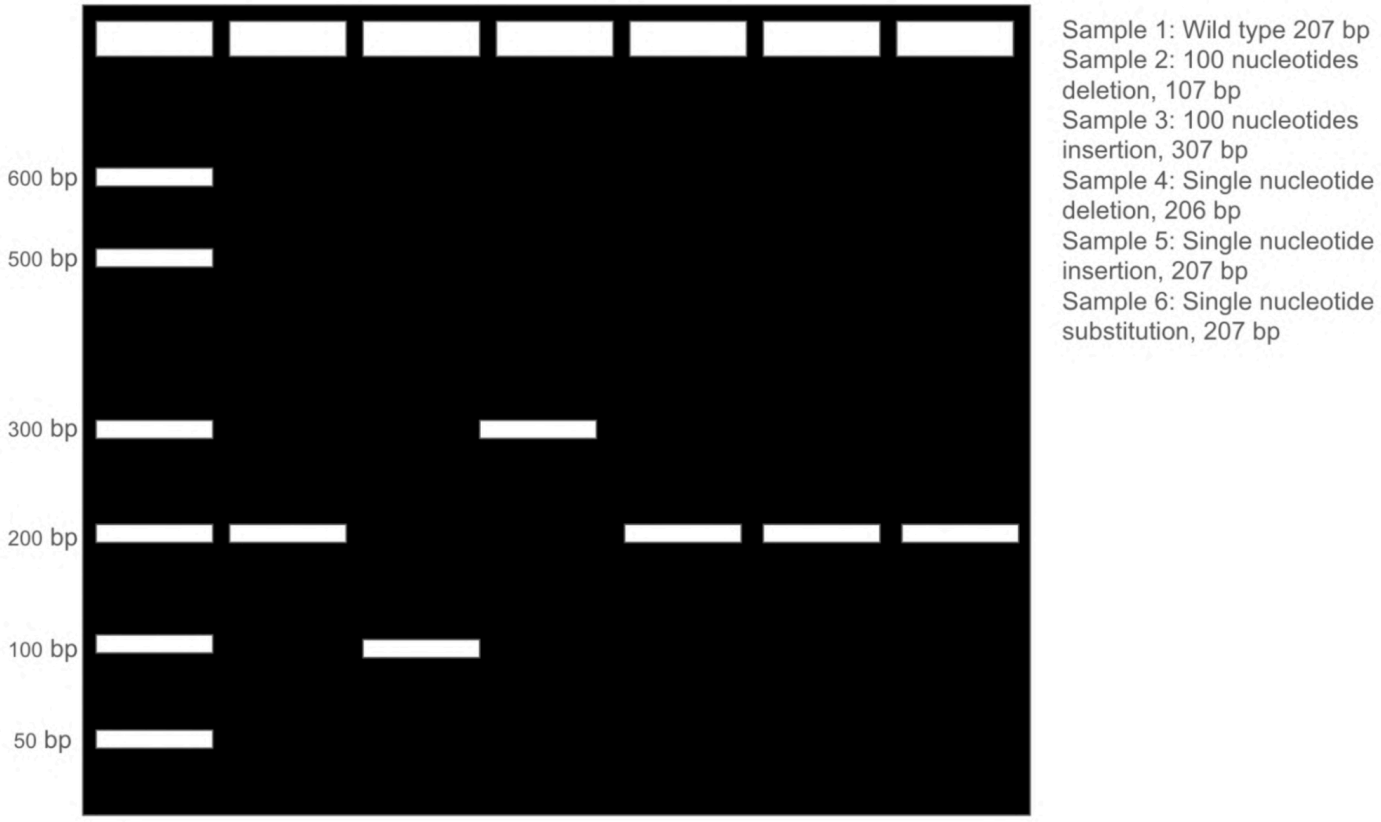


Figure 9. Simulated DNA gel electrophoresis to present potential outcomes of our CRISPR/Cas9 gene editing for training purposes. Lane 1: DNA molecular weight ladder /size marker (50-600bp). Lane 2 (Sample 1): Wild-type IL6 control showing a band at 207 bp. Lane 3 (Sample 2): Large-scale 100-nucleotide 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.

Ok donw

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, 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 IL6 signaling is connected to major inflammatory pathways that include: IL1B, TNFRSF1A, IL10, IL2, and IL3. The presence of IL10 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 a 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. NM., Antonius, Y., Susilo, R. JK., Hayaza, S., Kharisma, V. D., Parikesit, A. A., Zainul, R., Jakhmola, V., Saklani, T., Rebezov, M., Ullah, Md. 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., Eyras, 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

Diagnosis of Viral Gastroenteritis ("Stomach Flu")—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved April 3, 2026, from https://www.niddk.nih.gov/health-information/digestive-diseases/viral-gastroenteritis/diagnosis

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*, 40(3), 727. 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/

Staff, N. (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/

Viral gastroenteritis (stomach flu)—Diagnosis & treatment. (n.d.). Mayo Clinic. Retrieved December 30, 2025, from 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

Acknowledgements

[redacted]

Author Biography

[redacted]

Mentor Contribution Statement

[redacted]

Convergence
A Journal for Young Researchers

Received: [for official use only] ; Accepted: [for official use only];
Managing Editor: [for official use only];
Action Editor: [for official use only]

Bioinformatic characterization of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis~~Assessing the development of gastroenteritis via the IL-6-mediated pathway and designing CRISPR/Cas9 gene editing~~

Abstract

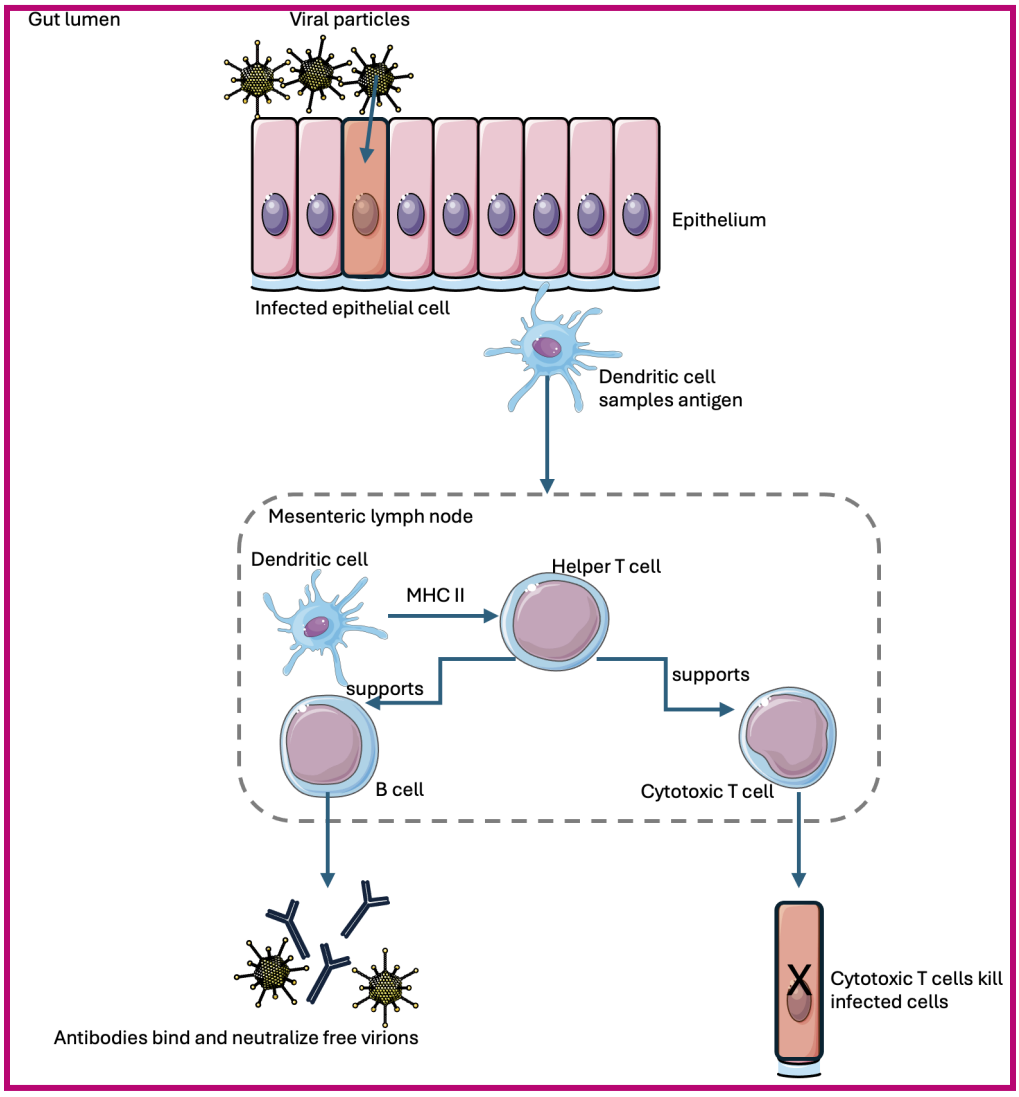
Gastroenteritis is a significant diarrheal disease worldwide, and its incidence is associated with environmental factors including temperature and rainfall. The molecular basis of this association is poorly understood. Elevated temperature has been shown to induce IL-6 release via the aryl hydrocarbon receptor pathway, and IL-6 levels are elevated in acute gastroenteritis and correlate with disease severity. This paper In silico analyses were conducted using TMHMM, SignalP, Clustal Omega, and the STRING database to investigate the properties of the IL-6 protein. TMHMM SignalP identified a 27-amino-acid signal peptide that is cleaved during secretion. Clustal Omega alignment across six species The STRING interactors, including IL6R, IL6ST/gp130, JAK1, and JAK2, IL1B, TNFRSF1A, IL10, IL2, and IL3 A CRISPR/Cas9 his study provides a

Keywords: gastroenteritis, IL6, environmental factors, temperature, CRISPR/Cas9, inflammation

1. Introduction

Gastroenteritis is the inflammation of the stomach and large or small 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, this is required for a diagnosis (*Diagnosis of Viral Gastroenteritis ("Stomach Flu")* - *NIDDK*, n.d.). ~~This disease occurs when the immune system activates to defend the stomach and intestinal tract against infection but fails.~~ 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. ~~Therefore, this paper investigates the role of IL-6-mediated inflammatory signaling in the development and progression of gastroenteritis under different environmental influences.~~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, 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. ~~Infection is closely linked to the environment; therefore, I look at IL-6-mediated gastroenteritis developments together with environmental factors.~~


Gut lumen
Viral particles
Epithelium
Infected epithelial cell
Dendritic cell
samples antigen
Mesenteric lymph node
Dendritic cell
MHC II
Helper T cell
supports
supports
B cell
Cytotoxic T cell
Antibodies bind and neutralize free virions
Cytotoxic T cells kill
infected cells

71

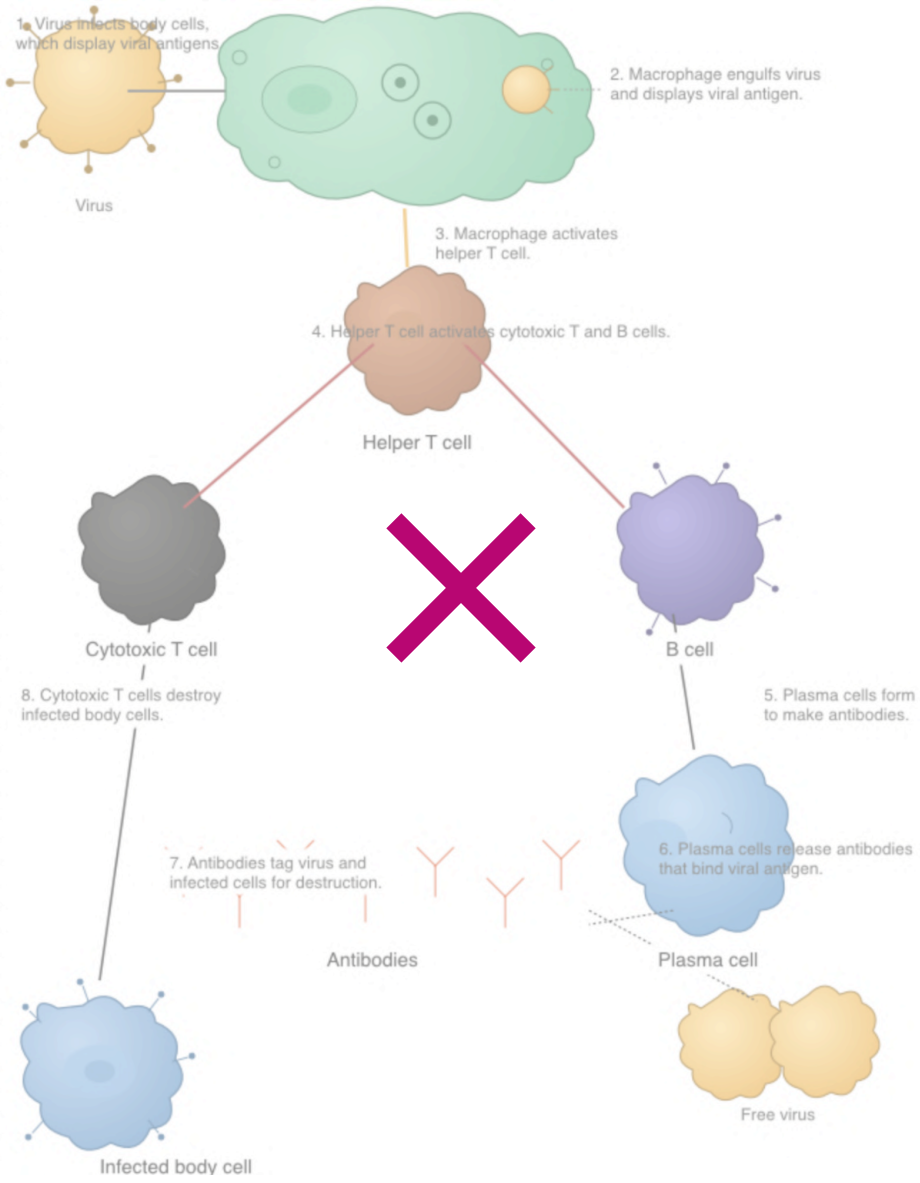


Figure 1. Immune response to enteric viral infection. Viral particles in the gut lumen infect intestinal epithelial cells. Dendritic cells sample viral antigen and migrate to the mesenteric lymph node, where antigen is presented on MHC II to helper T cells. Activated helper T cells support the responses 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 1. Steps that show the viral infection cycle linked to gastroenteritis.** The diagram describes the steps in how the immune system tackles infections. Each of these has specialized responsibilities and activities. The immune system fights pathogens (bacteria and viruses) and removes them from the body, recognizes and neutralizes toxins, and fights disease-causing changes in the body. It is activated by antigens, and when antigens attach to immune system cells (steps 2-8), a series of processes are triggered in the body. Once the body has encountered a pathogen for the first time, it stores information about how to fight it, so that a subsequent encounter with the same pathogen can be recognized and fought more quickly ("In Brief," 2023).¶

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.

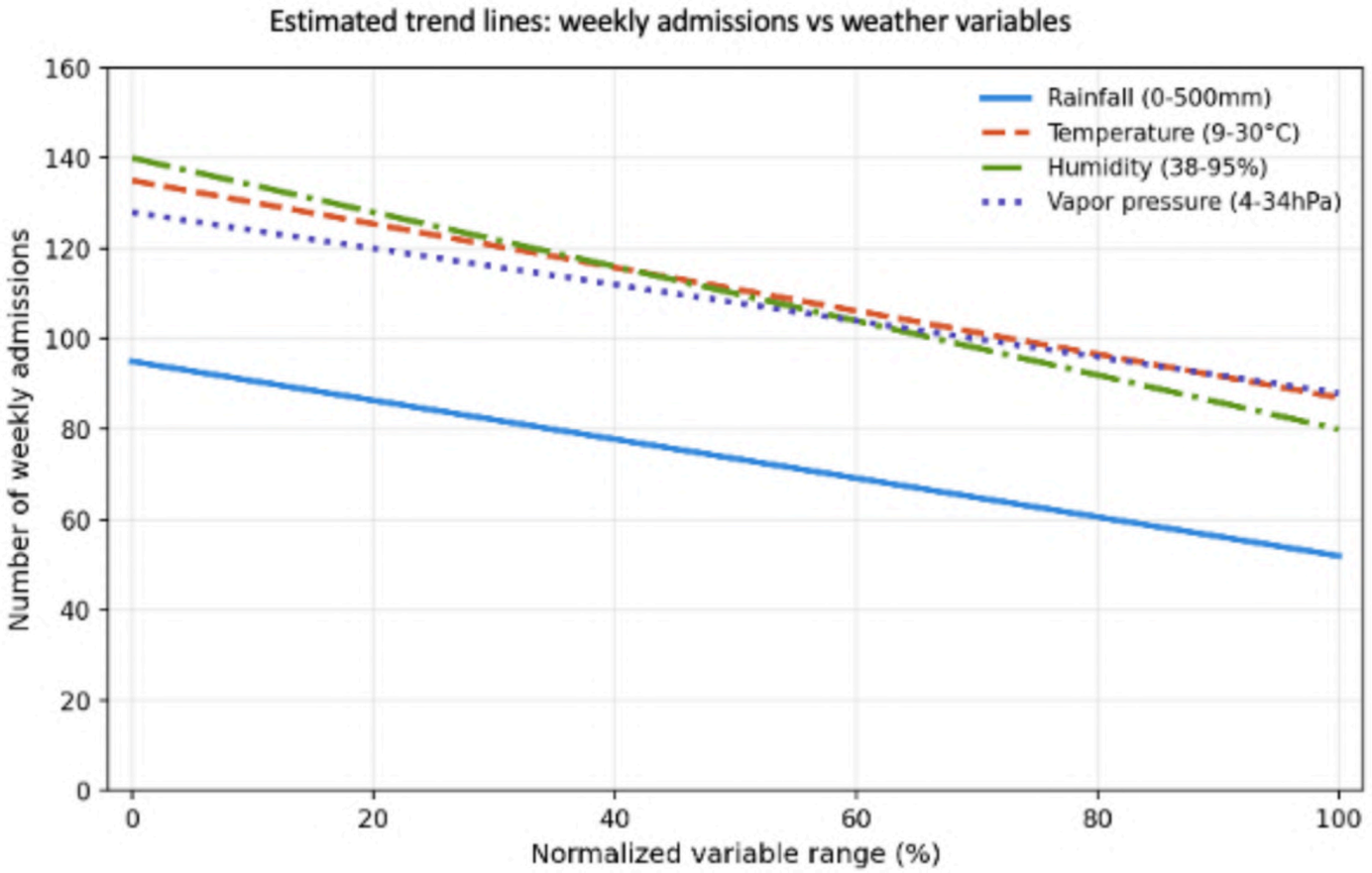


Figure 2. Statistics of pediatric patients admitted to hospital for gastrointestinal infection due to environmental factors (weather variables). A graph showing the weekly number of pediatric gastrointestinal infection admissions against different meteorological factors such as rainfall, temperature, 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 variable 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 ~~generated~~ produced by the author in Microsoft Excel using the underlying data reported by ~~with data provided 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~~they support the central hypothesis of this paper~~, 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 named 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 norivirus 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. ~~I selected to focus on the IL-6 gene for the scope of this paper in order to understand and develop a research framework to study gastroenteritis in the future.~~ The IL6 gene ~~produces~~ 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 ~~The encoded protein~~ capable of ~~has proven to be an endogenous pyrogen that can~~ induc~~ing~~e fever in people with infections or autoimmune diseases ~~autoimmune infections or diseases~~, and i~~. IL-6 contributes to the pathogenesis of acute gastroenteritis caused by rotavirus and norovirus, and i~~s 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~~n the context of acute gastroenteritis~~, both serum and fecal levels of IL-6 show significant elevation compared to healthy ~~individuals~~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 ~~. Critically, IL-6 levels are significantly higher in bacterial gastroenteritis than in viral cases, giving it strong diagnostic value as a biomarker to help differentiate between etiologies. IL-6 demonstrates superior sensitivity, specificity, and positive~~

predictive value compared to C-reactive protein (CRP) and IL-8 in distinguishing bacterial from viral causes. Furthermore, increased IL-6 levels correlate with disease severity and complications, including benign seizures, complications from Shigella infection, and increased liver enzymes (ALT/AST) in rotavirus gastroenteritis. Hence, IL-6 serves not only as a diagnostic marker for etiological differentiation but also as a prognostic indicator of disease severity in pediatric acute gastroenteritis (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.. Therefore, IL-6 is a good candidate to focus on for the investigation related to contributions of genetics and environmental factors to gastroenteritis.

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 IL6 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 IL6 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.The aim of this study is to evaluate the role of IL-6 in the development and progression of gastroenteritis in the context of environmental and genetic influences. This is achieved by investigating the structural and functional properties of the IL6 protein and proposing a CRISPR/Cas9 gene-editing strategy. Together, these analyses establish a framework for future studies investigating how environmental factors and genetic variation may influence IL-6-mediated inflammatory pathways in gastroenteritis.¶

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. ~~I used the NCBI>protein and NCBI>gene sections to retrieve protein and gene sequences.~~

Ensembl: A bioinformatics project with a purpose to organize biological information around the sequences of large genomes (Birney et al., 2004). Ensembl was used to annotate the exon and intron regions of IL6. ~~I used Ensembl 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. This analysis was performed using STRING v12.0 to identify predicted interactors of IL-6 ~~I used this analysis to search whether IL-6 has any transmembrane region.~~

SignalP: A widely used computational tool for predicting signal peptides from amino acid sequences (Teufel et al., 2022). The ~~I retrieved the~~ IL6 protein sequence was retrived from the NCBI protein database in FASTA format and submitted ~~it~~ to the SignalP algorithm for analysis.

Clustal Omega: An algorithm for multiple sequence alignment that can deal with very large numbers of DNA/RNA 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. ~~I used Clustal Omega analysis to look for the degree of conservation of the IL6 protein across different 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: textmining, 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. ~~I used this analysis to look for the interactors of IL-6.~~

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. ~~(Ansori et al., 2023).~~ 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 IL6 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 ~~I used it for DNA sequence alignment, annotation, editing, and visualizing the IL6 gene to present the location of the guide RNA and PCR primers for the gene editing design of IL-6~~ (*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 on the basis of 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~~I searched the gene of interest and selected an exon, and then the algorithm produced the potential guide RNAs in a table. Then I looked for the off-target scores for 0, 1, 2, 3, and 4 mismatches. I aimed to make 0=1, 1=0, 2=0, and for 3,4 the lowest possible scores.~~

3. Results

3.1 IL-6 does not contain any 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 about their function/s 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
TCNKSNMCESSKEALAENNLNLPKMAEKDGCFQSGFNEETCLVKIITGLLEFEVYLEYLQNRFESSEEQA
RAVQMSTKVLIQFLQKKAKNLDAITTPDPTTNASLLTKLQAQNQWLQDMTTHLILRSFKEFLQSSLRALR
```

```
# 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
```

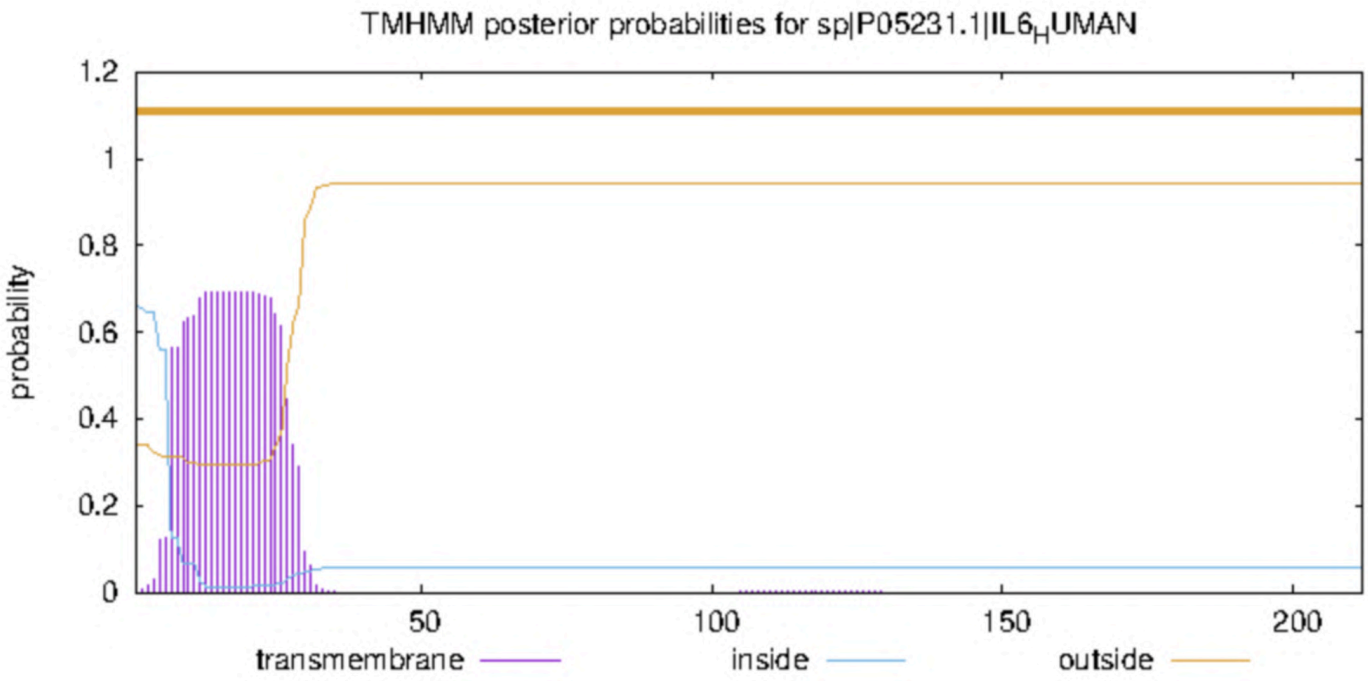


Figure 3. TMHMM transmembrane domain prediction result for the IL-6 protein. The upper text is the protein amino acid sequence retrieved from the NCBI Protein database. The bottom text and plot are the result of TMHMM analysis for IL-6. The text of the TMHMM result shows the amino acid length and the 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 for it 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.

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

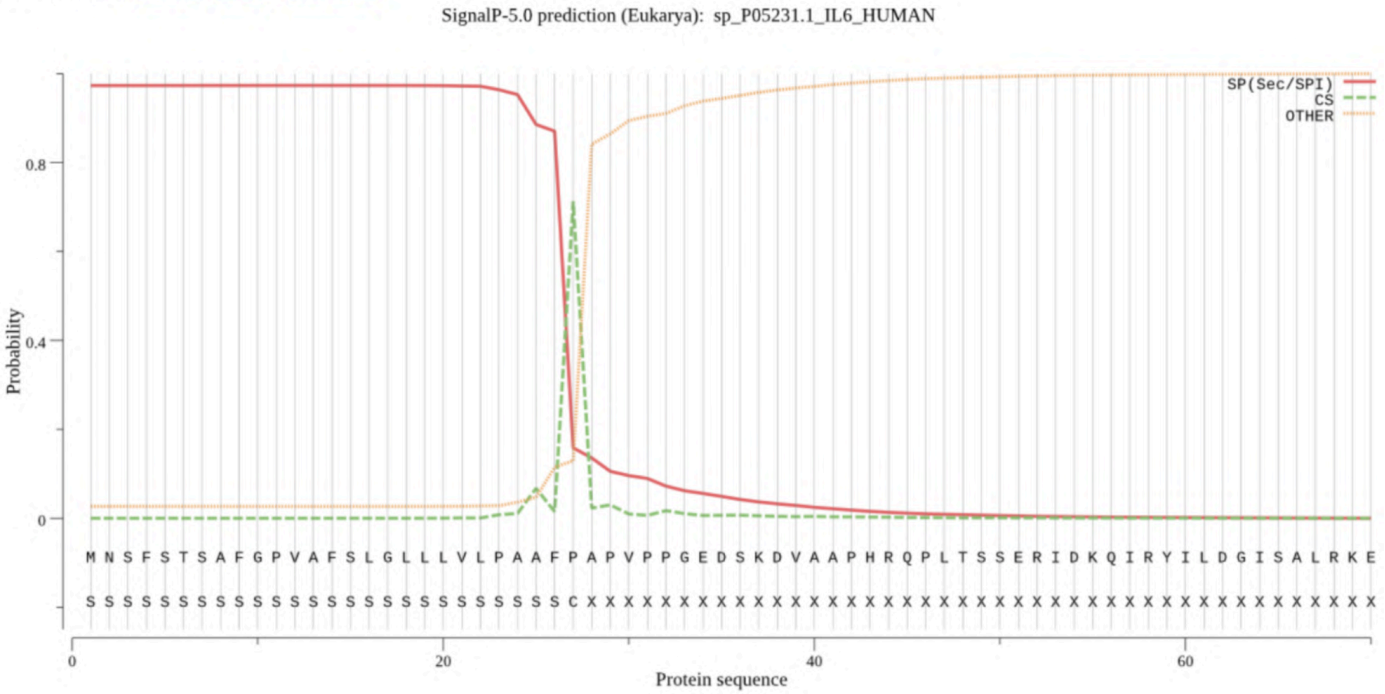


Figure 4. SignalP 5.0 analysis of IL-6. The plot shows the prediction of the presence of a signal peptide in IL6. The x-axis shows the amino acid sequence of IL-6; the y-axis shows the probability calculation between 0 and 1. The red line represents the amino acids that are signal peptides, the green 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.

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 IL6 protein, with the purpose of finding out

which parts of the IL6 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 IL6-related immune mechanisms is not conserved across the species put in this analysis.

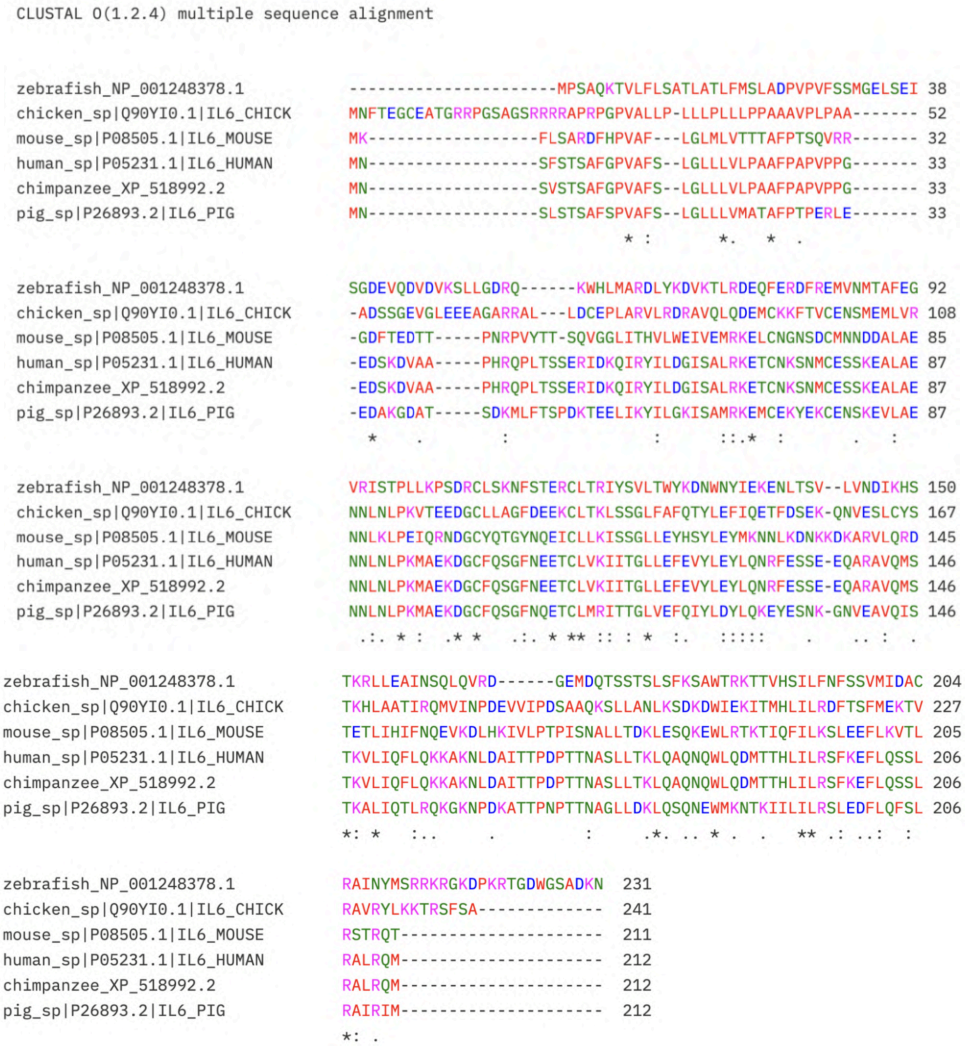


Figure 5. Multiple protein alignment of IL6 across different species. The alignment shows the species names on the left side and the amino acids aligned on the right side. There is a amino acid coloring scheme (ClustalX convention): which means RED: Small, hydrophobic amino acids (A, V, L, I, M, F, W), 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 the IL6

Protein-protein IL6 interaction 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, 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 a 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. ~~This analysis means that IL-6 is not merely a component of the inflammatory module, but rather a central node within it.~~

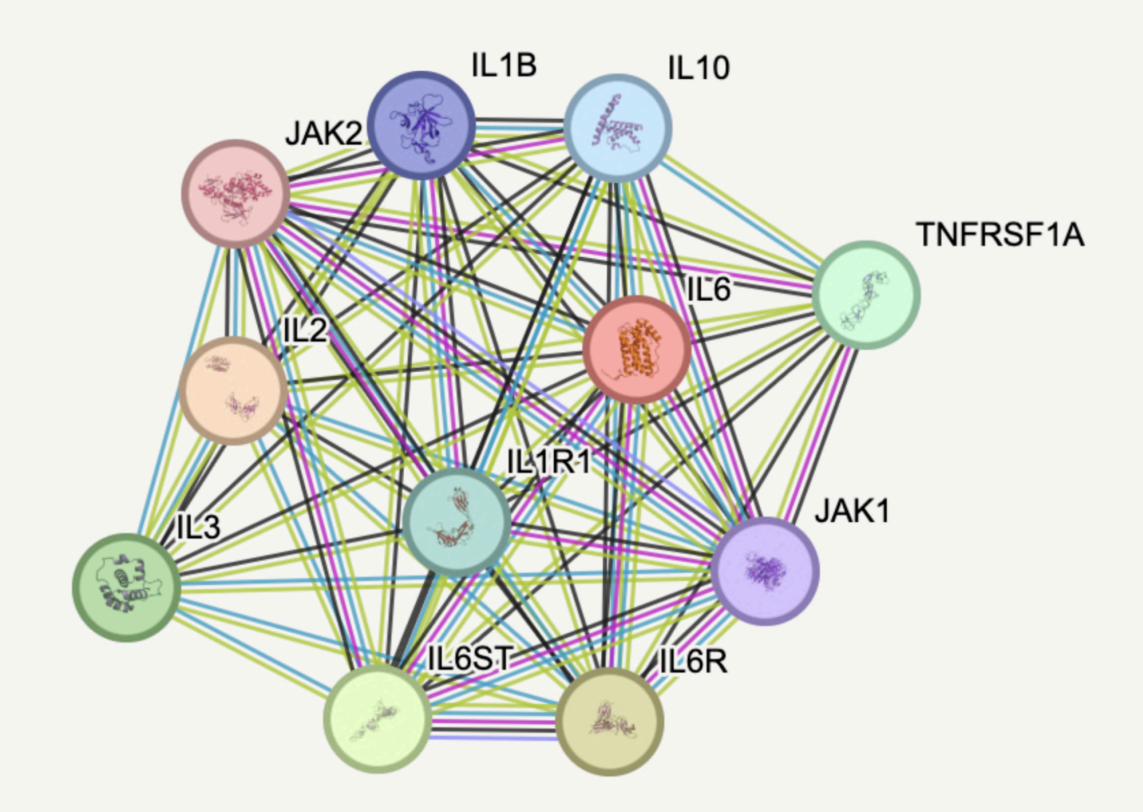


Figure 6. Protein-protein interaction map of IL-6. The map shows the proteins predicted to interact with IL-6, generated using the STRING database. Nodes represent proteins: IL6 (center), its receptor subunits IL6R and IL6ST, the associated kinases JAK1 and JAK2, and further inflammatory mediators IL1B, 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: ~~turquoise~~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 IL-6

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.~~Here, I proposed a gene editing strategy for generating mutants of IL-6 to study the effects of environmental factors combined with different genetic mutations on developing gastroenteritis in the future.~~

For the gene editing design, NCBI>GENE, Ensembl, CRISPR Finder (Sanger), and SnapGene were used. The gene sequence of IL-6 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 IL-6 on the NCBI gene database was checked on the NCBI database to ~~be sure~~confirm that the target region does not overlap ~~IL-6 does not collide~~ with ~~different~~ 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. ~~Later, the guide RNA was found through the CRISPR Finder database, and the guide RNA that was selected had the best possible off-target scores: 0=1, 1=0, 2=0, 3 and 4= as low as possible.~~ The guide and the PAM sequences were ~~also~~ marked on SnapGene Viewer (Figure 8). T~~Then t~~he PCR primers flanking the cut site were then ~~were~~ designed for hypothetical follow-up experiments. ~~for training purposes.~~ The precise outcomes would need to be verified ~~via~~ by sequencing in the actual laboratory practice.

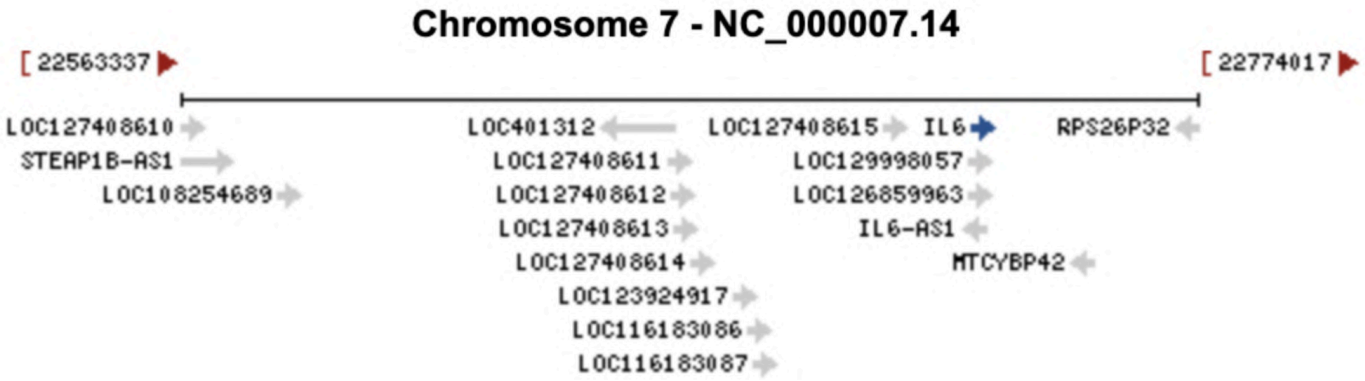
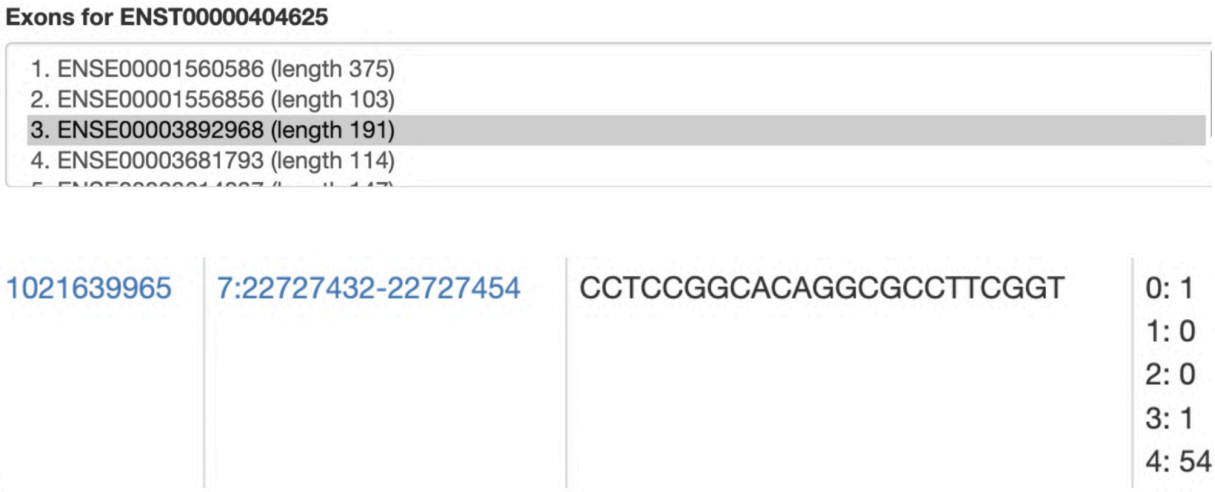


Figure 7. Chromosomal Location of IL-6. The genomic region spanning positions 22,563,337 to 22,774,017 on chromosome 7 is shown. The orientation of each gene is indicated by arrows. The purpose of this map is to identify the IL6 genomic location and guide the design of the CRISPR/Cas9 gene-editing targets within the IL6 locus.



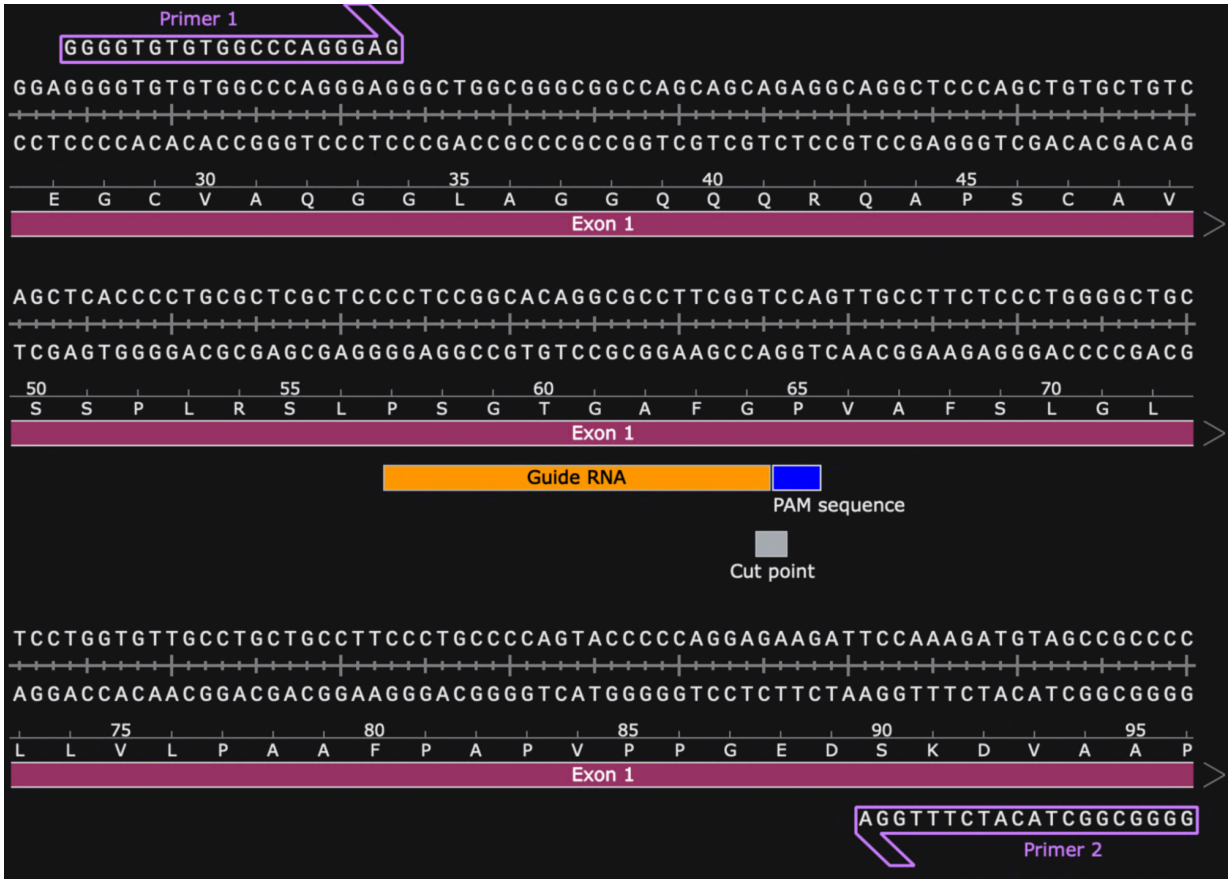


Figure 8. SnapGene view of CRISPR/Cas9 gene editing design for IL6. The upper image is the list of exons for IL6 in the CRISPR Finder database. The one highlighted is the selected exon to do gene editing. The middle image shows the selected Guide RNA (orange in Snapgene view below) and the off-target scores (the very 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 cut point for the PAM sequence (grey),~~ 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.

The proposed strategy enables a specific follow-up experiment. Intestinal epithelial cells carrying the IL-6 ~~edits~~ 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).~~IL-6 induction would be quantified at the transcript level by qPCR and at the protein level by ELISA and compared across genotypes. If IL-6 genetic variation modifies the effect of thermal stress, the magnitude of induction should differ between edited and wild-type cells. Equivalent induction across genotypes would argue against IL-6 variation as a modifier of the environmental effect and would redirect attention to the other interactors identified in the STRING network (Figure 6).~~

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. ~~The CRISPR/Cas9 gene-editing strategy targeting the IL-6 gene to study its role in the development of gastroenteritis was hypothetically designed, and PCR followed by DNA gel electrophoresis experiments were proposed and simulated as follow-up experiments. Different mutations in IL-6 could affect disease development in different ways, and these effects can vary when combined with environmental factors.~~ 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.

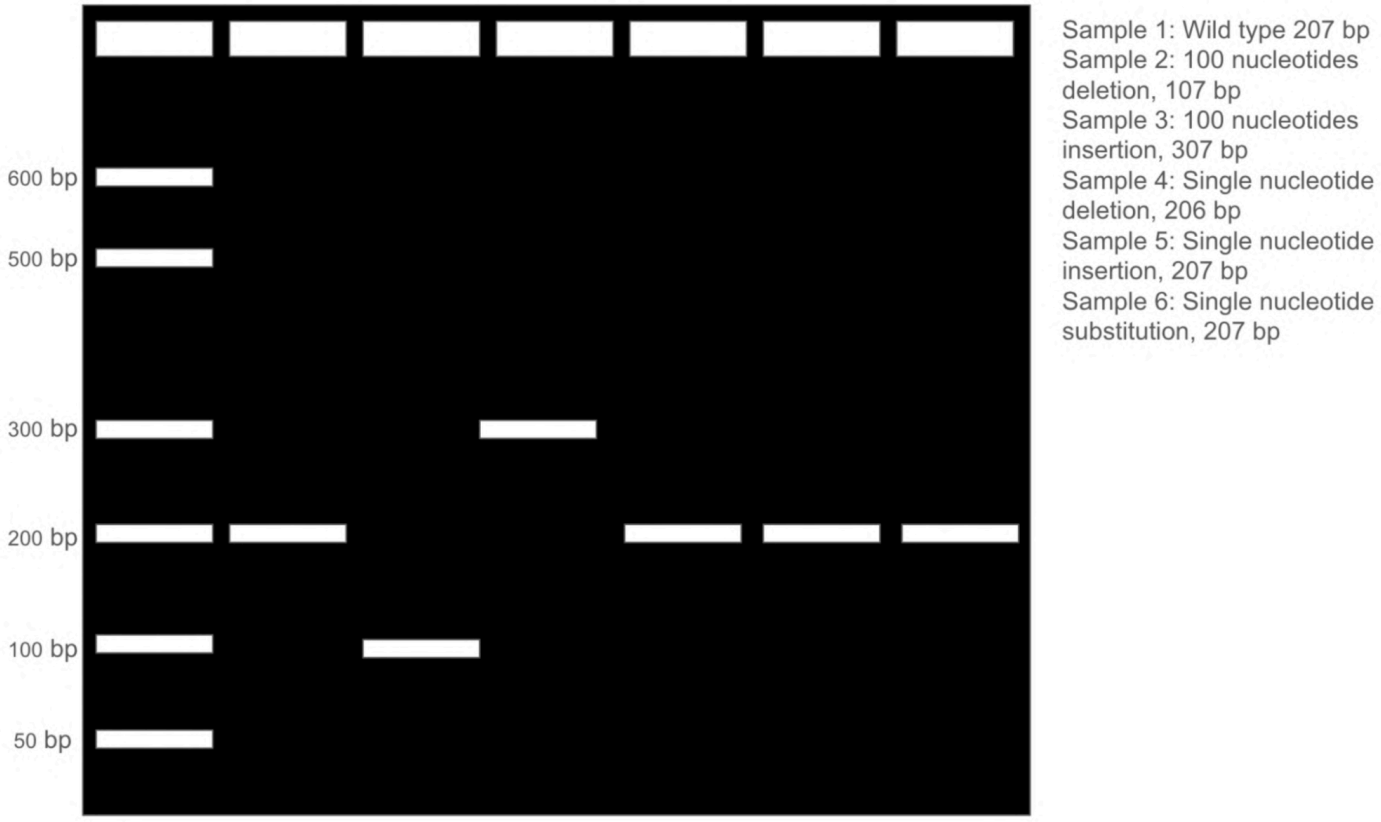


Figure 9. Simulated DNA gel electrophoresis to present potential outcomes of our CRISPR/Cas9 gene editing for training purposes. Lane 1: DNA molecular weight ladder /size marker (50-600bp). Lane 2 (Sample 1): Wild-type IL6 control showing a band at 207 bp. Lane 3 (Sample 2): Large-scale 100-nucleotide 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.

Ok donw

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.~~In this study, I established a framework by analyzing the protein properties of IL-6 and designing a gene-editing strategy to investigate the combined effects of environmental factors and IL-6 mutations in gastroenteritis.~~ 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, 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.~~Based on this interaction network, a biological model can be proposed to explain how these molecular components may influence disease development.~~ 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 IL6 signaling is connected to major inflammatory pathways that include: IL1B, TNFRSF1A, IL10, IL2, and IL3. The presence of IL10 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 a 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.~~There are underlying genetic factors, particularly involving the IL-6 gene, that could explain the diversity seen among populations despite the shared environmental conditions for developing gastroenteritis. I wanted to see whether the IL-6 protein is a conserved protein across animals and which amino acids are conserved evolutionarily.~~ 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 ~~could~~ would enable researchers to generate ~~to introduce targeted mutations into the~~ IL6 knockout and ~~6 gene 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. ~~The IL-6-mediated inflammatory pathway plays an important role in the inflammatory response associated with gastroenteritis and may be influenced by environmental factors.~~ 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 ~~genetic variation modifies~~ 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. NM., Antonius, Y., Susilo, R. JK., Hayaza, S., Kharisma, V. D., Parikesit, A. A., Zainul, R., Jakhmola, V., Saklani, T., Rebezov, M., Ullah, Md. 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., Eyras, 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

Diagnosis of Viral Gastroenteritis ("Stomach Flu")—NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved April 3, 2026, from https://www.niddk.nih.gov/health-information/digestive-diseases/viral-gastroenteritis/diagnosis

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*, 40(3), 727. 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/

Staff, N. (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/

Viral gastroenteritis (stomach flu)—Diagnosis & treatment. (n.d.). Mayo Clinic. Retrieved December 30, 2025, from 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

Acknowledgements

[redacted]

Author Biography

[redacted]

Mentor Contribution Statement

[redacted]

Response to Reviewer 1

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

I thank Reviewer 1 for a careful and constructive assessment of the manuscript. The central criticism — that the scope and interpretation extended beyond the analyses actually performed, was well founded, and addressing it has substantially improved the paper. The manuscript has been reframed throughout as a bioinformatic characterization of IL-6 together with a proposed gene-editing framework, rather than as an investigation of gastroenteritis development. Each comment is addressed in turn below.

1. Overall scope and interpretation

Comment:

> *The overall scope and interpretation of the manuscript should be revised to better match the analyses actually performed. IL-6 is occasionally presented as a demonstrated mediator of gastroenteritis development and progression under environmental stress, whereas the current study primarily consists of bioinformatic characterization and a hypothetical CRISPR/Cas9 gene-editing strategy. Statements regarding environmental factors, IL-6 signaling, and gastroenteritis pathogenesis should be appropriately toned down throughout the Introduction, Results, Discussion, and Conclusion.*

Response: Agreed, and this has been addressed as a global revision rather than a series of local edits. The title has been changed to "Bioinformatic characterization of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis." The abstract now states explicitly that the paper does not test the environmental–IL-6–gastroenteritis relationship, and adds the sentence "No editing, expression, or infection experiments were performed." Predictive language has replaced demonstrative language throughout the Results, so TMHMM and SignalP outputs are now described as predicted rather than revealed or identified. In the Introduction, the closing sentence of the first paragraph no longer states that the paper investigates the role of IL-6 in disease development. In the Discussion, the framework sentence has been rewritten, and the biological model is now explicitly identified as hypothetical and untested in this work. The Conclusion no longer asserts that the IL-6 pathway plays an important role in gastroenteritis.

2. Figure 2 — rationale and permission

Comment:

> *The rationale for including this figure should be clarified. The connection between the environmental factors discussed in the manuscript and IL-6 signaling is currently weak. Also, is this figure adapted or reproduced from a previously published paper? If so, please clarify whether permission to use the figure has been obtained.*

Response: On provenance: the figure is my own work and no permission is required. It was produced in Microsoft Excel using the underlying data reported by Chong et al. (2021); it is not reproduced or adapted from any published figure. The caption now states this explicitly and also reports the normalization method, which was previously omitted. On rationale: a sentence has been added immediately before the figure explaining that it illustrates the seasonal relationship between weather

variables and pediatric gastroenteritis admissions, which motivates the focus on environmental influences developed in the following sections.

3. Line 153 — stated purpose of gene selection

Comment:

> *The statement that the IL-6 gene was selected "to understand and develop a research framework to study gastroenteritis in the future" is problematic. The study did not experimentally investigate gastroenteritis or establish such a framework. Future applications should not be presented as the primary objective or main thesis of the current study.*

Response: Agreed. That clause has been removed. The paragraph now gives a substantive scientific rationale for selecting IL6: that it sits at the convergence of the two influences the paper is concerned with, being inducible by environmental exposures such as thermal stress and particulate air pollution while also being elevated during enteric infection.

4. Line 177 — stated aim

Comment:

> *The stated aim of the study — to evaluate the role of IL-6 in the development and progression of gastroenteritis under environmental and genetic influences — does not appear to be supported by the actual analyses presented. The aim should be revised to more accurately reflect the scope and results of the current study.*

Response: The aim has been narrowed to match what was done. It now reads that the study characterizes the structural and functional properties of IL-6 using publicly available bioinformatic tools and designs a defined CRISPR/Cas9 strategy for generating an IL6 knockout in intestinal epithelial cells, followed by two explicit sentences stating that the study does not investigate gastroenteritis directly and does not test whether IL-6 mediates the relationship between environmental exposure and disease outcome.

5. Methods — passive voice

Comment:

> *The Methods section should be revised to use passive voice consistently.*

Response: All first-person constructions have been converted throughout sections 2.1 to 2.3.

6. Multiple protein alignment — rationale

Comment:

> *The rationale for including this analysis is unclear. Please explain how the interspecies conservation of IL-6 is directly relevant to the research question and its relationship to gastroenteritis.*

Response: The rationale has been restated, and in doing so the original justification has been corrected rather than merely expanded. The alignment was previously introduced as evidence bearing on variation among human populations, which cross-species comparison cannot support. The Discussion now states that the alignment was performed to assess which residues of IL-6 are conserved across species, that highly conserved residues are often functionally important, and that this information is relevant both to

interpreting the consequences of disruption and to selecting appropriate model organisms for future experimental work.

7. Section 3.5 — the intended CRISPR/Cas9 modification

Comment:

> *The intended CRISPR/Cas9 editing strategy is unclear. It is not specified whether the authors aim to generate an IL-6 knockout, a specific point mutation, or an insertion/deletion mutation. These are fundamentally different experimental objectives and would require different CRISPR/Cas9 designs and validation strategies. The authors should clearly define the intended genetic modification and provide a specific rationale for why that particular modification is relevant to the research question.*

Response: This was the most substantive gap in the original manuscript and it has been closed. The intended modification is now stated as a loss-of-function IL6 knockout, defined explicitly in Methods 2.3: a single guide RNA directs Cas9 to a coding exon, the resulting double-strand break is resolved by non-homologous end joining, and frameshifting indels are expected to disrupt the reading frame and abolish production of functional IL-6 protein. The section also now states that no donor template was designed and that no HDR, base-editing, or prime-editing strategy was used, because the objective is gene disruption rather than reconstruction of a defined naturally occurring variant. Section 3.5 and the abstract have been revised to use this definition consistently.

On the related point that the section read as a research proposal rather than a completed study: I have retained the proposed workflow but reframed it. The manuscript now presents the in-silico characterization as its own contribution and the editing strategy explicitly as a framework for future experimental work, rather than presenting the future experiment as the paper's central component.

8. Section 3.6 — simulated gel electrophoresis

Comment:

> *Please clarify what type of IL-6 mutation you are intending to generate. At present, the manuscript appears to describe several possible outcomes without clearly specifying which mutation is actually being generated. Please revise this section so that the proposed CRISPR/Cas9 strategy leads to one clearly defined genetic modification rather than presenting multiple hypothetical possibilities.*

Response: Section 3.6 has been rewritten to follow from the single defined objective. It no longer presents the gel as a survey of possible mutations, but as a genotyping screen for the indels that NHEJ repair is expected to produce, noting that edited cells would need to be genotyped before any downstream comparison could be made. The Figure 9 caption now clarifies that Samples 4 to 6 are included to illustrate the resolution limits of agarose gel electrophoresis rather than to represent expected NHEJ outcomes, and states that a single-nucleotide substitution in particular would not arise from NHEJ repair and would require a donor template or base editing. It also notes that sequencing would be required to confirm the precise edit in any real application.

9. Figure 8 caption

Response: Separately, the Figure 8 caption has been corrected: what was previously described as "the cut point for the PAM sequence" is now described as the predicted Cas9 cleavage site, approximately 3 nt upstream of the PAM.

Closing

I am grateful to Reviewer 1 for identifying the mismatch between the manuscript's claims and its analyses. The revision has made the paper both narrower and more defensible, and I believe it is now a clearer account of what was actually done.

Response to Reviewer 2

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

I thank Reviewer 2 for an exceptionally detailed and generous review. The comments identified several places where the manuscript's language outran its evidence, and one place where it conflated two genuinely different kinds of genetic variation. I have accepted all of the substantive recommendations, including the suggested title and the suggested framing of the central claim. Each comment is addressed in turn below.

Major comment 1 — Title, aims and conclusions should reflect the study performed

Comment:

> *The manuscript does not directly assess the development or progression of gastroenteritis, nor does it perform CRISPR editing. Statements such as "this paper investigates the role of IL-6" and "IL-6 genetic variation modifies the effect of environmental stressors" sound more definitive than the evidence allows. A more accurate title would be something like: Bioinformatic characterisation of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis.*

Response: I have adopted the suggested title, with spelling adjusted to match the manuscript's American English convention: "Bioinformatic characterization of IL-6 and a proposed CRISPR/Cas9 framework for studying environmental influences on gastroenteritis." Both specific sentences quoted have been removed. The abstract now states that the paper does not test the relationship directly and adds an explicit sentence that no editing, expression, or infection experiments were performed. The Introduction's opening paragraph no longer states that the paper investigates the role of IL-6-mediated signaling in disease development, and the sentence about looking at IL-6-mediated gastroenteritis developments has been reworded to remove the causal implication. The aim, Discussion framework sentence, and Conclusion have all been narrowed correspondingly.

Major comment 2 — Association, causation and mediation require clearer separation

Comment:

> *Elevated IL-6 does not necessarily cause the disease to become more severe. More severe infection and tissue injury may instead produce a greater inflammatory response and consequently higher IL-6 levels. The central claim should be framed as: 'Environmental exposures can induce IL-6 and may consequently modify gastroenteritis-associated inflammation.' It should not yet be framed as: 'Environmental exposures modify gastroenteritis outcomes through IL-6.'*

Response: I have adopted the suggested formulation verbatim in the Introduction, where the paragraph on AhR-mediated IL-6 induction now closes by stating that these findings motivate the hypothesis that environmental exposures can induce IL-6 and may consequently modify gastroenteritis-associated inflammation. The previous wording, which asserted modification of outcomes through IL-6, has been removed. Separately, the paragraph introducing IL-6 now acknowledges the reverse-direction explanation directly: it states that these are associations reported within specific study populations rather than established clinical performance, and that the direction of the relationship is not resolved, since more severe infection and tissue injury may produce a larger inflammatory response and

consequently higher IL-6 rather than IL-6 driving severity. The Conclusion has been revised on the same basis.

Major comment 3 — Rationale for selecting IL-6, and interpretation of genetic variation

Comment:

> *The Introduction lists 17 genes and then describes each individually, but the transition to IL6 remains abrupt. The author should add a concise explanation of why IL6 was selected over genes with established pathogen-specific relevance, such as FUT2. Regarding the cross-species alignment, this does not demonstrate that human IL6 genetic variation contributes to differences in gastroenteritis susceptibility. Sequence divergence between zebrafish, chickens, rodents and humans is not equivalent to variation among human individuals.*

Response: Both points have been accepted. On gene selection, the paragraph following the 17-gene list now explains the choice explicitly: it notes that several of the other genes have well-characterized pathogen-specific roles, giving FUT2 as the example, since secretor status is an established determinant of norovirus susceptibility and its relationship to gastroenteritis is therefore already comparatively well defined. It then states that IL6 was chosen instead because it sits at the convergence of environmental exposure and infection-associated inflammation, being inducible by thermal stress and particulate air pollution as well as elevated during enteric infection.

The abstract previously stated that low conservation suggested genetic variation may contribute to differences in disease susceptibility between individuals, it now states that the alignment identifies conserved and divergent residues and may inform the selection of appropriate model organisms for studying human IL-6 biology, which follows the interpretation suggested in the review. The corresponding Discussion paragraph previously opened by framing the analysis as explaining diversity among human populations despite shared environmental conditions. That framing has been removed entirely and replaced with a statement that the alignment was performed to assess cross-species conservation, relevant to interpreting the consequences of disruption and to model organism selection. I have not added human variant analysis, as SNP, allele frequency and eQTL analysis is beyond the scope of the present study, instead the claim has been withdrawn to what the data support.

Major comment 4 — STRING network wording is circular; parameters should be reported

Comment:

> *The claim that IL-6 is a "central node" is partly circular because IL-6 was entered as the query protein and its first-shell associations were displayed. Network centrality would require a defined network and a quantitative centrality analysis. For the sake of transparency, the STRING version, organism, confidence threshold, number of interactors and evidence settings should be reported.*

Response: The circularity is accepted and the claim has been withdrawn. The closing sentence of section 3.4, which stated that IL-6 is not merely a component of the inflammatory module but a central node within it, has been replaced. The text now states that 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, and that 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. The same correction has been applied to the abstract.

All requested parameters are now reported in Methods 2.2: STRING v12.0, query submitted as IL-6 and resolved to the Homo sapiens IL6 entry, full STRING network with edges indicating both functional and physical associations, evidence mode display, all seven interaction sources active (textmining, experiments, databases, co-expression, neighborhood, gene fusion, co-occurrence), minimum required interaction score of medium confidence (0.400), first shell limited to a maximum of 10 interactors, and no second shell requested.

Major comment 5 — Figure 6 legend appears incorrect

Comment:

> *The Figure 6 colour legend appears incorrect. In STRING evidence mode, yellow represents text-mining, black represents co-expression, blue represents gene co-occurrence, purple represents experimental evidence and light blue represents curated database evidence.*

Response: The Figure 6 caption now reports the evidence colors as they are actually assigned by STRING, following the scheme given in the review.

Major comment 6 — Positioning of IL-6 as harmful; dual role should be discussed

Comment:

> *The manuscript must recognise that IL-6 is not simply harmful. It participates in host defence, epithelial survival and intestinal repair. Experimental IL-6 deficiency has worsened pathogen control and epithelial injury in some intestinal infection models. This dual role should be discussed before presenting IL6 editing as a potentially beneficial intervention.*

Response: A new paragraph has been added to the Discussion, placed before the discussion of the editing strategy as recommended. It states that IL-6 is not straightforwardly pathogenic in the gut, that it contributes to epithelial repair, supports the acute-phase response and participates in controlling enteric pathogens, and that its elevation during gastroenteritis therefore reflects an active host response as well as a potential contributor to tissue damage. It notes that removing IL-6 might worsen outcomes rather than improve them, and concludes that 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.

Major comment 7 — CRISPR strategy needs a clearly defined biological purpose

Comment:

> *The manuscript alternates between discussing "genetic variation," "mutations" and IL6 disruption, but these are not equivalent experimental goals. Cas9 cleavage followed by non-homologous end joining generally creates a mixture of insertions and deletions and is most suitable for disrupting or knocking out a gene. It would not accurately reproduce a specific naturally occurring IL6 variant. The author should state whether the intended experiment is an IL6 knockout experiment, introduction of a particular human variant, or generation of nonspecific mutants as a preliminary functional screen.*

Response: The intended experiment is an IL6 knockout, and this is now stated explicitly. Methods 2.3 defines the modification as a loss-of-function IL6 knockout produced by NHEJ-derived frameshifting indels, and states that 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 IL6 variant. The point that NHEJ cannot reproduce a specific human variant is therefore now acknowledged in the manuscript as the reason for the design choice rather than left as an unexamined inconsistency. Language referring to "genetic variation" and "mutants" in the context of the editing strategy has been removed from the abstract, section 3.5, section 3.6, the Discussion and the Conclusion, so that the manuscript now describes one objective consistently.

Major comment 8 — Proposed future experiment does not test gastroenteritis outcomes

Comment:

> *The proposed experiment compares IL-6 transcript and protein levels between edited and wild-type intestinal epithelial cells at 37°C and 38.5°C. This could test whether the edit changes temperature-associated IL-6 expression, but it would not test whether IL-6 modifies gastroenteritis. A future experiment addressing the central hypothesis would need an infectious challenge and relevant endpoints, for example edited and control intestinal organoids exposed to an enteric pathogen, measuring pathogen replication or burden, epithelial barrier integrity, cell death and viability, and inflammatory cytokine production.*

Response: The proposed experiment has been rewritten. Defining the edit as a knockout also exposed a problem with the original readout, since IL-6 transcript and protein cannot be measured in cells from which functional IL-6 has been removed. The revised paragraph states this explicitly and changes the endpoints to downstream measures: STAT3 phosphorylation by Western blot, transcript levels of other first-shell cytokines such as IL1B and IL10 by qPCR, and epithelial barrier integrity by transepithelial electrical resistance. Following the recommendation, an infectious challenge has been added: both genotypes would additionally be challenged with an enteric pathogen or viral surrogate at each temperature, with barrier disruption and cytokine output as endpoints. The interpretation has been reframed accordingly, so that the experiment now tests whether IL-6 is required for the thermal component of the epithelial inflammatory response rather than whether IL-6 variation modifies environmental effects.

Major comment 9 — Figure 1 needs revision

Comment:

> *Dendritic cells, rather than macrophages alone, should be shown as the main cells initiating T-cell activation. The figure should also clarify that helper T cells support, rather than directly activate, antigen-specific cytotoxic T cells and B cells. Antibodies mainly bind and neutralise free virus, whereas cytotoxic T cells destroy infected body cells. Some of the lines have been out of order as well as the Y shapes of the antibodies.*

Response: Figure 1 has been redrawn in full, and all five points are addressed. Dendritic cells are now shown as the cells that sample viral antigen at the epithelium and present it on MHC II in the mesenteric lymph node, initiating helper T cell activation. The two branching arrows from the helper T cell to the B cell and the cytotoxic T cell are both labeled "supports" rather than indicating direct activation. The effector arms are now separated correctly: antibodies are shown binding and neutralizing free virions, while cytotoxic T cells are shown killing an infected epithelial cell. The sequence now runs in a single direction from lumen to lymph node to effector response, which resolves the line ordering, and the antibodies are drawn with correct Y morphology. The figure was rebuilt using Servier Medical Art assets, licensed under CC BY 4.0, and this is credited in the caption.

Minor comments

1. Section numbering:

> *The Results heading is numbered "2," while its subsections begin with 3.1; the Discussion is then numbered 3.*

Response: Corrected. The cause was that the Materials and Methods heading had not been included in the automatic numbered list. All six section headings are now numbered consistently: Introduction 1, Materials and Methods 2, Results 3, Discussion 4, Conclusion 5, References 6, so that the 2.x and 3.x subsections align with their parent sections.

2. Gene and protein nomenclature:

> *IL6 should refer to the gene, whereas IL-6 should refer to the protein.*

Response: Applied throughout. IL6 is now used for the gene, locus, sequence, transcript, knockout and editing target; IL-6 is used for the protein, its signal peptide, its interactors and measured levels. This affected the abstract, the 3.3 and 3.4 headings and captions, the 3.5 heading and body, the Figure 4, 6 and 7 captions, the Discussion and the Conclusion.

3. Terminology:

> *"Autoimmune infections or diseases" should be corrected to "infections or autoimmune diseases."*

Response: Corrected.

4. Mechanism of disease:

> *The statement that gastroenteritis occurs because the immune system "activates but fails" is inaccurate. Infection and symptoms result from pathogen-associated injury together with the host response; the presence of disease does not necessarily mean that the immune system has failed.*

Response: Corrected. The sentence has been replaced with a statement that symptoms result from a combination of pathogen-associated injury to the gut lining and the host inflammatory response mounted against it. The following sentence, which stated that inflammatory cells are sometimes not enough, has been deleted.

5. Diagnostic claims:

> *Statements describing IL-6 as having "strong diagnostic value" should be qualified as findings reported in particular cohorts, rather than accepted clinical performance across all gastroenteritis populations.*

Response: Qualified. The relevant sentences now attribute these findings to pediatric cohorts specifically and describe them as reported sensitivity and specificity within those study populations rather than as established clinical performance across all gastroenteritis presentations.

6. Repeated definitions:

> *The manuscript overall would benefit from removal of repeated definitions.*

Response: Addressed. Tool descriptions for TMHMM, SignalP, Clustal Omega and STRING are now given once, in the Methods, and the duplicate explanations that previously appeared again in Results 3.1 to 3.4 have been removed so that the Results report findings only.

7. Figure 6 legend:

Colour legend corrected — see Major comment 5 above.

Closing

I am particularly grateful for the care taken over the distinction between cross-species divergence and human individual variation, and over the mismatch between NHEJ-derived editing and the modeling of specific variants. Both were errors I would not have identified myself, and correcting them has improved the manuscript considerably.