Supplementary Material

Empirical and theoretical analysis of particle diffusion in mucus

Antonio Cobarrubia[1,2], Jarod Tall[1,2], Austin Crispin-Smith[1,2], and Antoni Luque[1,3,4]

[1]Viral Information Institute, San Diego State University, San Diego, CA 92182, USA,
[2]Department of Physics, San Diego State University, San Diego, CA 92182, USA, [3]Department of
Mathematics and Statistics, San Diego State University, San Diego, CA 92182, USA,
[4]Computational Science Research Center, San Diego State University, San Diego, CA 92182,
USA.

Source Data 1. File containing the list of all the references initially collected with information regarding the diffusion of particles in mucus. The references used in the meta-analysis are highlighted in bold.

Source Data 2. Full dataset used in the meta-analysis. The first row is a header with designated column names fitted to the physical properties collected. Particle name, particle type, zeta potential, particle size, effective diffusion constant at 1 s, anomalous exponent, diffusion in water, the ratio between effective diffusion constant at 1 s and diffusion in water, temperature, pH, dosing medium, the salt type used, salt concentration, mucus used in the experiment, mucus concentration, mucus purification, mucin gene expression level, and dominant mucin gene is denoted as Particle, Surface_Chemistry, Zeta, Diameter, D_w, Diffusion_constant, alpha, Ratio_Diffusion, Temperature, pH, Dosing Medium, Salt_type, Salt_Concentration, Mucus_Type, Muc_Con, Purification, 'mucin gene name'_EL, and Dominate_Mucin, respectively.

Source Data 3. This file contains the data analysis used to extract the dominant mucin for each mucus source. Mucin_expression_level. The first row is a header, and the first column designates mucus type. Represented expression levels are denoted as 'mucin gene name'_EL. Expression levels are classified as below cutoff, low, medium, high, or unavailable. The average median Transcripts per million (TPM) for each mucin gene is designated as 'mucin gene name'_TPM_median_avg.

Dependent variable	Independent variable	Slope	Intercept	R^2	p-value‖	rho‡	p-value‡
Effective Diffusion¶∧	Diameter∧○	-2.1 ± 0.3	3.5 ± 0.4	0.67	3.5E-7***	-0.9	1.7E-12***
Effective Diffusion¶∧	Negative†	0.024 ± 0.006	-1.6 ±0.2	0.30	0.0006***	0.6	0.0002***
Effective Diffusion¶∧	Positive†	0.01 ± 0.02	-2.8 ± 0.4	0.03	0.5	0.3	0.2
Effective Diffusion¶∧	Alpha	5.3 ± 0.3	-5.0 ± 0.2	0.89	<2.0E-16***	0.9	<2.2E-16***
Anomalous Exponent	Diameter∧○	-0.13 ± 0.03	1.11 ± 0.04	0.39	0.0007***	-0.6	0.001
Anomalous Exponent	Negative†	0.004 ± 0.001	0.68 ± 0.04	0.26	0.002***	0.5	0.0007***
Anomalous Exponent	Positive†	0.003 ± 0.003	0.44 ± 0.07	0.04	0.4	0.3	0.2

Table S.1: Supplementary material of linear Analysis of Nanoparticles' Mobility Through Mucus and Biohydrogels. ‖ Simple Linear Regression and Pearson's p-value for the slope. ‡ Spearman analysis. ¶ Effective Diffusion coefficient (μm^2/s). ∧ Logarithmic of base 10 (log_{10}). ○ Diameter less than 100 nm. *** Data with strong statistical significance. † Zeta potential (mV). Overview table of values associating with statistically significant simple linear regression along with spearman and pearson analysis.
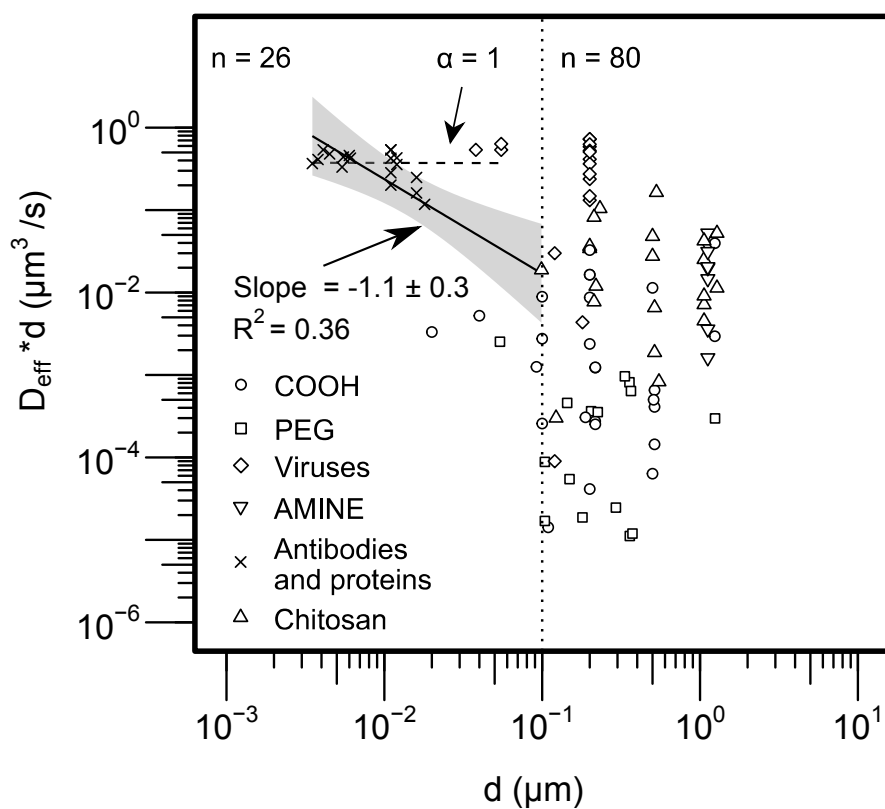
Figure S.1: Effective diffusion scaled by particle size. The effective diffusion times the particle size plotted as a function of the anomalous exponent. The solid line represents the regression model. The grey area represents the 95% confidence interval. Statistically significant slope and R^2 of linear regression are displayed. The dashed line shows the trend of particles predicted to display regular diffusion. The legend indicates the symbols associated with each particle type.
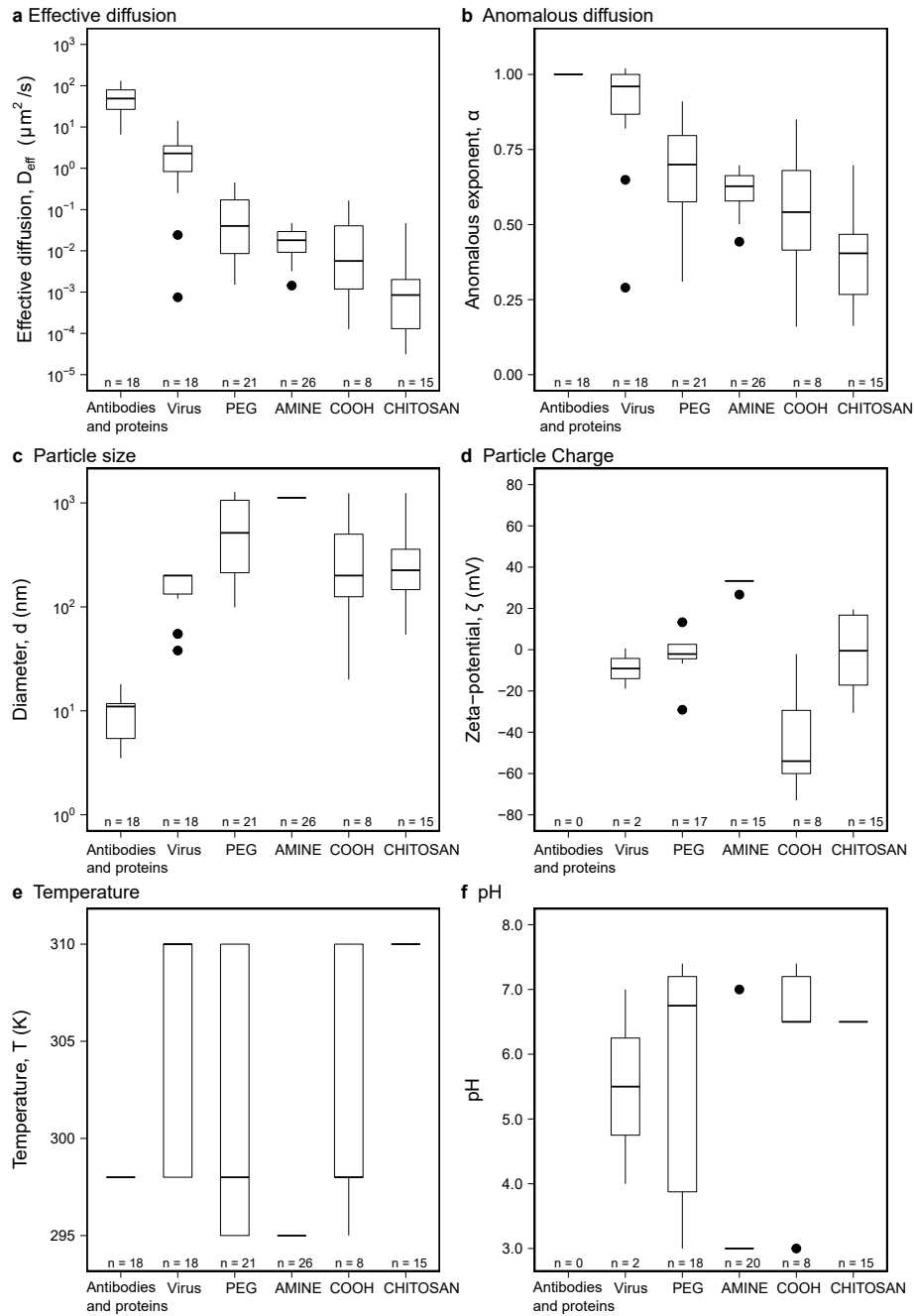
Figure S.2: Transport capabilities based on particle type. **a**, effective diffusion constant at one second based on particle type **b**, anomalous exponent based on particle type. **c**, .particle size based on particle type. **d**, particle net charge based on particle type. **e**, mucus temperature based on particle type. **f**, mucus pH based on particle type. **a-f**, box plots are ranked by effective diffusion from high to low. The total amount of data points for each particle type is designated as n aligned with their respected particle type for each panel.
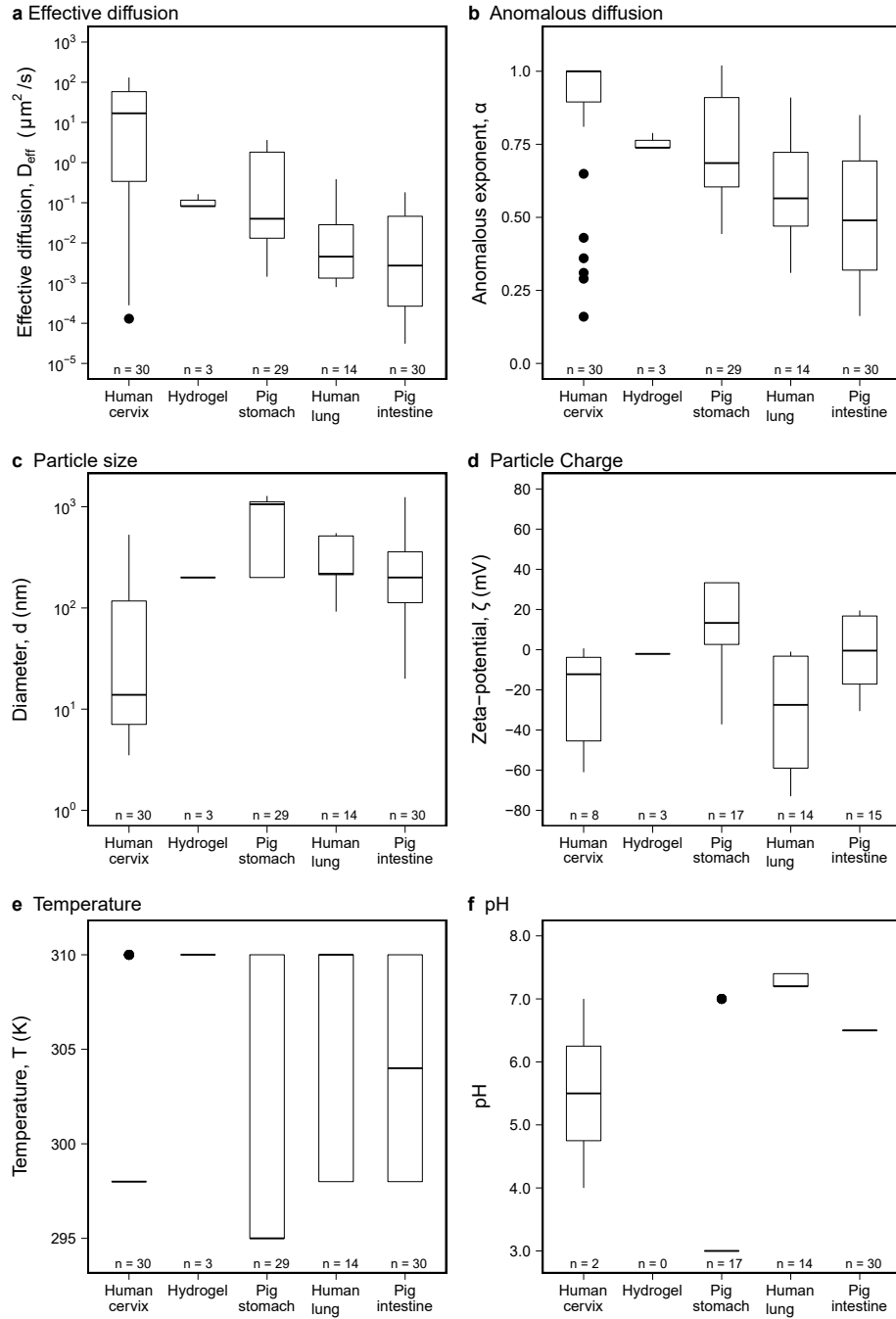
Figure S.3: Mucus source influence on particle transport. **a**, effective diffusion constant at one second based on mucus source. **b**, anomalous exponent based on mucus source. **c**, particle size based on mucus source. **d**, particle net charge based on mucus source. **e**, mucus temperature based on mucus source. **f**, mucus pH based on mucus source. **a-f**, box plots are ranked by effective diffusion from high to low. The total amount of data points for each mucus source is designated as n aligned with their respected mucus source for each panel.
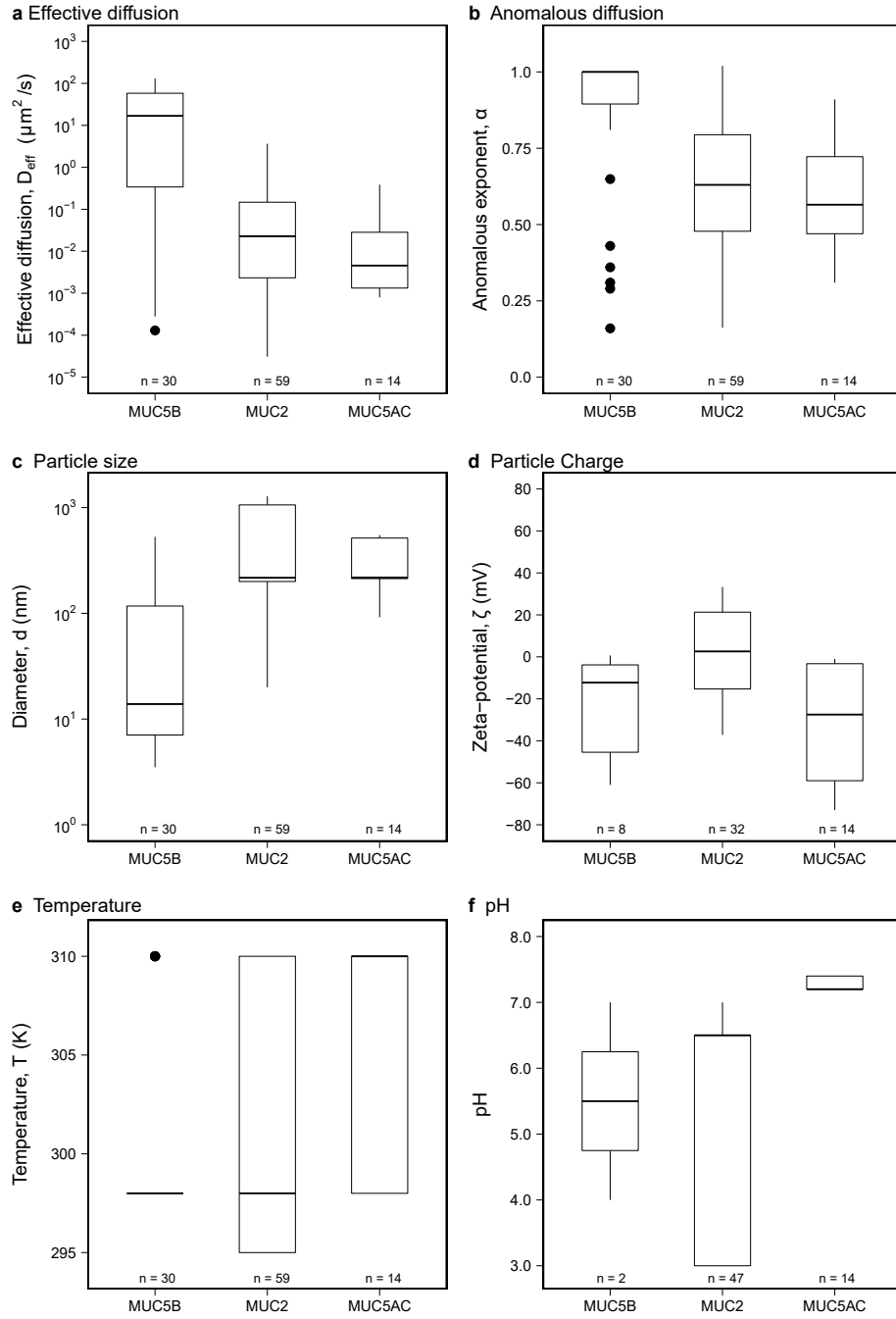
Figure S.4: Impact of dominant mucin genes on particle transport. **a**, effective diffusion constant at one second based on dominate mucin. **b**, anomalous exponent based on dominate mucin. **c**, particle size based on dominate mucin. **d**, particle net charge based on dominate mucin. **e**, mucus temperature based on dominate mucin. **f**, mucus pH based on dominate mucin. **a-f**, box plots are ranked by effective diffusion from high to low. The total amount of data points for each dominate mucin is designated as n aligned with their respected dominate mucin for each panel.
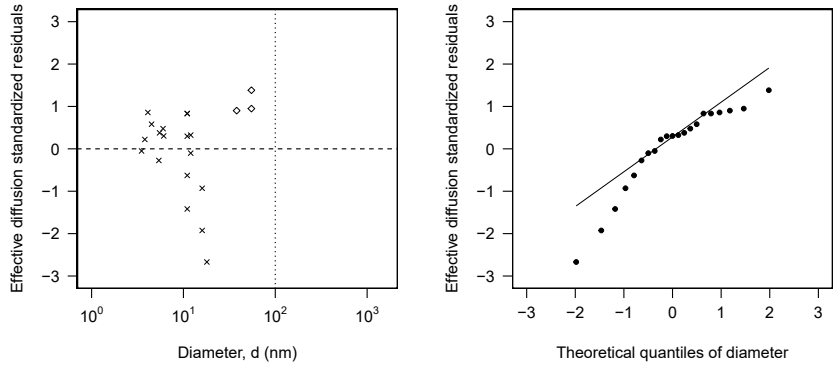
a Overview of the linear regression model for effective diffusion against anomaous exponent
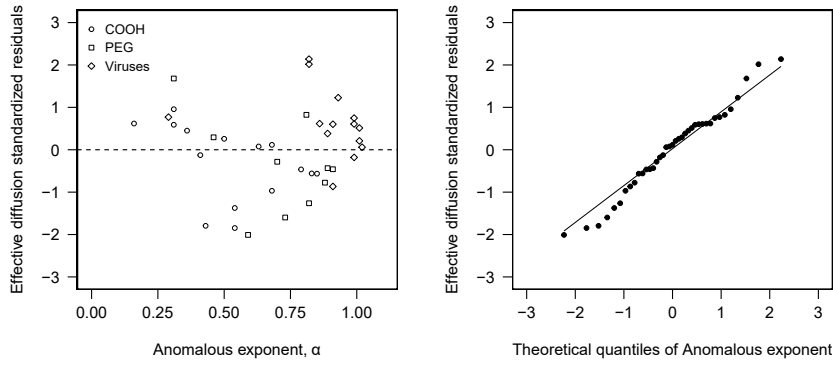
b Overview of the linear regression model for particle size smaller than 100 nm, d < 100 nm
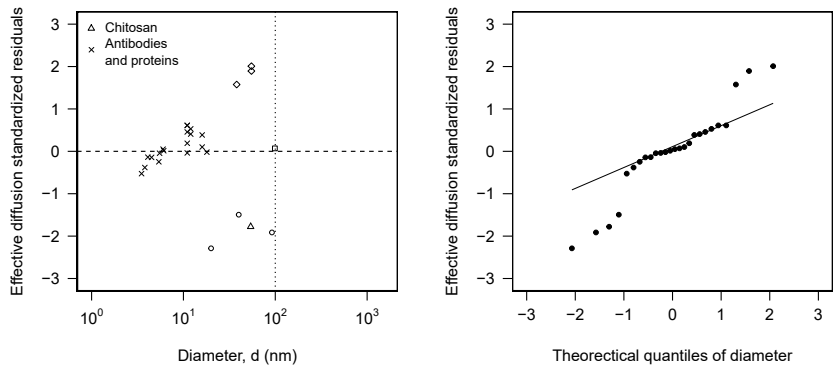
c Overview of the linear regression model for particles displaying normal brownian motion, α = 1

d Overview of the linear regression model for negative charged correlation
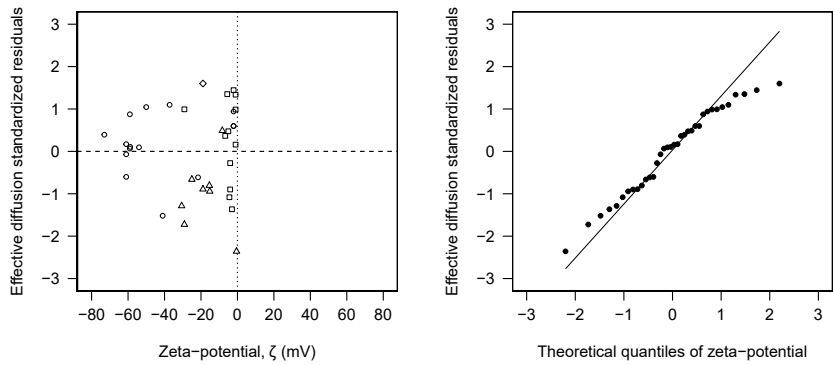
Figure S.5: In-depth overview of significant linear regression models. **a** standardized residuals, and the normal probability of standardized residuals for effective diffusion as a function of the anomalous exponent. **b** standardized residuals, and the normal probability of standardized residuals for effective diffusion as a function of particle size for sizes less than 100 nm. The dotted line is a visualization marker for particles smaller than 100 nm. **c** standardized residuals, and the normal probability of standardized residuals for effective diffusion as a function of particle size for particles displaying normal Brownian motion. The dotted line is a visualization marker for particles smaller than 100 nm. **d** standardized residuals, and the normal probability of standardized residuals for effective diffusion as a function of zeta potential for negatively charged particles. The dotted line is a visualization marker for negatively charged particles. **a-d**, different particle types with corresponding symbols are designated in **a** and **b**'s legend.

—