Annotated Bibliography

Wai-Tong (Louis) Fan (waifan@iu.edu)
Indiana University

May 26, 2022

References

[1] Amarjit Budhiraja and Wai-Tong Fan. Uniform in time interacting particle approximations for nonlinear equations of Patlak-Keller-Segel type. *Electron. J. Probab.*, 22:Paper No. 8, 37, 2017.

> We study a system of interacting diffusions that models chemotaxis of biological cells or microorganisms (referred to as particles) in a chemical field that is dynamically modified through the collective contributions from the particles. Such systems of reinforced diffusions have been widely studied and their hydrodynamic limits that are nonlinear nonlocal partial differential equations are usually referred to as Patlak-Keller-Segel (PKS) equations. "Solutions of the classical PKS equation may blow up in finite time and much of the PDE literature has been focused on understanding this blow-up phenomenon. In this work we study a modified form of the PKS equation that is natural for applications and for which global existence and uniqueness of solutions are easily seen to hold. Our focus here is instead on the study of the long time behavior through certain interacting particle systems. "Under the so-called 'quasi-stationary hypothesis' on the chemical field, the limit PDE reduces to a parabolic-elliptic system that is closely related to granular media equations whose time asymptotic properties have been extensively studied probabilistically through certain Lyapunov functions. The modified PKS equation studied in the current work is a parabolic-parabolic system for which analogous Lyapunov function constructions are not available. A key challenge in

the analysis is that the associated interacting particle system is not a Markov process as the interaction term depends on the whole history of the empirical measure. "We establish, under suitable conditions, uniform in time convergence of the empirical measure of particle states to the solution of the PDE. We also provide uniform in time exponential concentration bounds for rate of the above convergence under additional integrability conditions. Finally, we introduce an Euler discretization scheme for the simulation of the interacting particle system and give error bounds that show that the scheme converges uniformly in time and in the size of the particle system as the discretization parameter approaches zero.

[2] Amarjit Budhiraja, Wai-Tong Fan, and Ruoyu Wu. Large deviations for Brownian particle systems with killing. *J. Theoret. Probab.*, 31(3):1779–1818, 2018.

Particle approximations for certain nonlinear and nonlocal reaction-diffusion equations are studied using a system of Brownian motions with killing. The system is described by a collection of i.i.d. Brownian particles where each particle is killed independently at a rate determined by the empirical sub-probability measure of the states of the particles alive. A large deviation principle (LDP) for such sub-probability measure-valued processes is established. Along the way a convenient variational representation, which is of independent interest, for expectations of nonnegative functionals of Brownian motions together with an i.i.d. sequence of random variables is established. Proof of the LDP relies on this variational representation and weak convergence arguments.

[3] Kin Hang Chan, Wai-Tong Fan, and Ka Ho Law. *Pui Ching invitational mathematics competition 2007-2008 questions and answers*. Pui Ching Education Center, 2008.

This problem solving book contains the solutions for the Pui Ching invitational mathematics competition, a regional competition for more than 5000 high school and middle school students in Hong Kong. The authors (also problem poser of the competition) not only explain the solutions, but also

discuss about general problem solving strategies beyond the competition.

[4] Zhen-Qing Chen and Wai-Tong Fan. Scaling limits of interacting diffusions in domains. *Front. Math. China*, 9(4):717–736, 2014.

This survey article summarizes the PhD thesis work of Wai-Tong (Louis) Fan.

[5] Zhen-Qing Chen and Wai-Tong Fan. Functional central limit theorem for Brownian particles in domains with Robin boundary condition. *J. Funct. Anal.*, 269(12):3765–3811, 2015.

We rigorously derive non-equilibrium space–time fluctuation for the particle density of a system of reflected diffusions in bounded Lipschitz domains in . The particles are independent and are killed by a time-dependent potential which is asymptotically proportional to the boundary local time. We generalize the functional analytic framework introduced by Kotelenez [20], [21] to deal with time-dependent perturbations. Our proof relies on Dirichlet form method rather than the machineries derived from Kotelenez's sub-martingale inequality. Our result holds for any symmetric reflected diffusion, for any bounded Lipschitz domain and for any dimension.

[6] Zhen-Qing Chen and Wai-Tong Fan. Hydrodynamic limits and propagation of chaos for interacting random walks in domains. *Ann. Appl. Probab.*, 27(3):1299–1371, 2017.

A new non-conservative stochastic reaction–diffusion system in which two families of random walks in two adjacent domains interact near the interface is introduced and studied in this paper. Such a system can be used to model the transport of positive and negative charges in a solar cell or the population dynamics of two segregated species under competition. We show that in the macroscopic limit, the particle densities converge to the solution of a coupled nonlinear heat equations. For this, we first prove that propagation of chaos holds by establishing the uniqueness of a new BBGKY hierarchy. A local central limit theorem for reflected diffusions in bounded Lipschitz domains is also established as a crucial tool.

[7] Zhen-Qing Chen and Wai-Tong Fan. Systems of interacting diffusions with partial annihilation through membranes. *Ann. Probab.*, 45(1):100–146, 2017.

We introduce an interacting particle system in which two families of reflected diffusions interact in a singular manner near a deterministic interface I . This system can be used to model the transport of positive and negative charges in a solar cell or the population dynamics of two segregated species under competition. A related interacting random walk model with discrete state spaces has recently been introduced and studied in Chen and Fan (2014). In this paper, we establish the functional law of large numbers for this new system, thereby extending the hydrodynamic limit in Chen and Fan (2014) to reflected diffusions in domains with mixed-type boundary conditions, which include absorption (harvest of electric charges). We employ a new and direct approach that avoids going through the delicate BBGKY hierarchy.

[8] Zhen-Qing Chen and Wai-Tong Louis Fan. Fluctuation limit for interacting diffusions with partial annihilations through membranes. *J. Stat. Phys.*, 164(4):890–936, 2016.

We study fluctuations of the empirical processes of a non-equilibrium interacting particle system consisting of two species over a domain that is recently introduced in Chen and Fan (Ann Probab, to appear) and establish its functional central limit theorem. This fluctuation limit is a distribution-valued Gaussian Markov process which can be represented as a mild solution of a stochastic partial differential equation. The drift of our fluctuation limit involves a new partial differential equation with nonlinear coupled term on the interface that characterized the hydrodynamic limit of the system. The covariance structure of the Gaussian part consists two parts, one involving the spatial motion of the particles inside the domain and other involving a boundary integral term that captures the boundary interactions between two species. The key is to show that the Boltzmann–Gibbs principle holds for our non-equilibrium system. Our proof relies on generalizing the usual correlation functions to the join correlations at two different times.

[9] Rick Durrett and Wai-Tong Fan. Genealogies in expanding populations. *Ann. Appl. Probab.*, 26(6):3456–3490, 2016.

The goal of this paper is to prove rigorous results for the behavior of genealogies in a one-dimensional long range biased voter model introduced by Hallatschek and Nelson (2008). The first step, which is easily accomplished using results of Mueller and Tribe(1995), is to show that when space and time are rescaled correctly, our biased voter model converges to a Wright– Fisher SPDE. A simple extension of a result of Durrett and Restrepo (2008) then shows that the dual branching coalescing random walk converges to a branching Brownian motion in which particles coalesce after an exponentially distributed amount of intersection local time. Brunet et al. [Phys. Rev. E (3) 76 (2007) 041104, 20] have conjectured that genealogies in models of this type are described by the Bolthausen– Sznitman coalescent. However, in the model we study there are no simultaneous coalescences. Our third and most significant result concerns "tracer dynamics" in which some of the initial particles in the biased voter model are labeled. We show that the joint distribution of the labeled and unlabeled particles converges to the solution of a system of stochastic partial differential equations. A new duality equation that generalizes the one Shiga developed for the Wright–Fisher SPDE is the key to the proof of that result.

[10] Wai Tong Fan. *Interacting particle systems with partial annihilation through membranes*. ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–University of Washington.

This thesis studies the hydrodynamic limit and thefluctuation limit for a class of interacting particle systems in domains. These systems are introduced to model the transport of positive and negative charges in solar cells. However, they are general microscopic models that can describe a variety of macroscopic phenomena with coupled boundary conditions, such as the population dynamics of two segregated species under competition. Proving these two types of limits represents establishing the functional law of large numbers and the functional central limit theorem, respectively, for the time-trajectory of the particle densities. This also corresponds to

the study of the behavior of the system at two different scales. We show that the hydrodynamic limit is a pair of deterministic measures whose densities solve a coupled nonlinear heat equations, while the fluctuation limit can be described by a Gaussian Markov process that solves a stochastic partial differential equation.

[11] Wai-Tong Fan. Discrete approximations to local times for reflected diffusions. *Electron. Commun. Probab.*, 21:Paper No. 16, 12, 2016.

We propose a discrete analogue for the boundary local time of reflected diffusions in bounded Lipschitz domains. This discrete analogue, called the discrete local time, can be effectively simulated in practice and is obtained pathwise from random walks on lattices. We establish weak convergence of the joint law of the discrete local time and the associated random walks as the lattice size decreases to zero. A cornerstone of the proof is the local central limit theorem for reflected diffusions developed in [7]. Applications of the join convergence result to PDE problems are illustrated.

[12] Wai-Tong Fan, Chanh Kieu, Dimitrios Sakellariou, and Mahashweta Patra. Hitting time of rapid intensification onset in hurricane-like vortices. *Physics of Fluids*, 33(9):096603, 2021.

Predicting tropical cyclone (TC) rapid intensification (RI) is an important yet challenging task in current weather forecast due to our incomplete understanding of TC nonlinear processes. This study examines the variability of RI onset, including the probability of RI occurrence and the timing of RI onset, using a low-order stochastic model for TC development. Defining RI onset as the first hitting time for a given subset in the TC-scale state space, we quantify the probability of the occurrence of RI onset and the distribution of the timing of RI onset for a range of initial conditions and model parameters. Based on asymptotic analysis for stochastic differential equations, our results show that RI onset occurs later, along with a larger variance of RI onset timing, for weaker vortex initial condition and stronger noise amplitude. In the small noise limit, RI onset probability approaches one and the RI onset timing has less uncertainty (i.e., a smaller

variance), consistent with observation of TC development under idealized environment. Our theoretical results are verified against Monte-Carlo simulations and compared with explicit results for a general 1-dimensional system, thus providing new insights into the variability of RI onset and helping better quantify the uncertainties of RI variability for practical applications.

[13] Wai-Tong Fan, On Shun Pak, and Mario Sandoval. Ellipsoidal brownian self-driven particles in a magnetic field. *Physical Review E*, 95(3):1–6, 2017.

We study the two-dimensional Brownian dynamics of an ellipsoidal paramagnetic microswimmer moving at a low Reynolds number and subject to a magnetic field. Its corresponding mean-square displacement, showing the effect of a particles's shape, activity, and magnetic field on the microswimmer's diffusion, is analytically obtained. Comparison between analytical and computational results shows good agreement. In addition, the effect of self-propulsion on the transition time from anisotropic to isotropic diffusion of the ellipse is investigated.

[14] Wai-Tong Fan and Sebastien Roch. Necessary and sufficient conditions for consistent root reconstruction in Markov models on trees. *Electron. J. Probab.*, 23:Paper No. 47, 24, 2018.

We establish necessary and sufficient conditions for consistent root reconstruction in continuous-time Markov models with countable state space on bounded-height trees. Here a root state estimator is said to be consistent if the probability that it returns to the true root state converges to 1 as the number of leaves tends to infinity. We also derive quantitative bounds on the error of reconstruction. Our results answer a question of Gascuel and Steel [GS10] and have implications for ancestral sequence reconstruction in a classical evolutionary model of nucleotide insertion and deletion [TKF91].

[15] Wai-Tong Fan and Sebastien Roch. Statistically consistent and computationally efficient inference of ancestral DNA sequences in the TKF91 model under dense taxon sampling. *Bull. Math. Biol.*, 82(2):Paper No. 21, 32, 2020.

In evolutionary biology, the speciation history of living organisms is represented graphically by a phylogeny, that is, a rooted tree whose leaves correspond to current species and whose branchings indicate past speciation events. Phylogenetic analyses often rely on molecular sequences, such as DNA sequences, collected from the species of interest, and it is common in this context to employ statistical approaches based on stochastic models of sequence evolution on a tree. For tractability, such models necessarily make simplifying assumptions about the evolutionary mechanisms involved. In particular, commonly omitted are insertions and deletions of nucleotides—also known as indels. Properly accounting for indels in statistical phylogenetic analyses remains a major challenge in computational evolutionary biology. Here, we consider the problem of reconstructing ancestral sequences on a known phylogeny in a model of sequence evolution incorporating nucleotide substitutions, insertions and deletions, specifically the classical TKF91 process. We focus on the case of dense phylogenies of bounded height, which we refer to as the taxon-rich setting, where statistical consistency is achievable. We give the first explicit reconstruction algorithm with provable guarantees under constant rates of mutation. Our algorithm succeeds when the phylogeny satisfies the 'big bang' condition, a necessary and sufficient condition for statistical consistency in this setting.

[16] Wai-Tong Louis Fan. Stochastic PDEs on graphs as scaling limits of discrete interacting systems. *Bernoulli*, 27(3):1899–1941, 2021.

Stochastic partial differential equations (SPDE) on graphs were recently introduced by Cerrai and Freidlin (Ann. Inst. Henri Poincaré Probab. Stat. 53 (2017) 865–899). This class of stochastic equations in infinite dimensions provides a minimal framework for the study of the effective dynamics of much more complex systems. However, how they emerge from microscopic individual-based models is still poorly understood, partly due to complications near vertex singularities. In this work, motivated by the study of the dynamics and the genealogies of expanding populations in spatially structured environments, we obtain a new class of SPDE on graphs of

Wright-Fisher type which have nontrivial boundary conditions on the vertex set. We show that these SPDE arise as scaling limits of suitably defined biased voter models (BVM), which extends the scaling limits of Durrett and Fan (Ann. Appl. Probab. 26 (2016) 3456–3490). We further obtain a convergent simulation scheme for each of these SPDE in terms of a system of Itô SDEs, which is useful when the size of the BVM is too large for stochastic simulations. These give the first rigorous connection between SPDE on graphs and more discrete models, specifically, interacting particle systems and interacting SDEs. Uniform heat kernel estimates for symmetric random walks approximating diffusions on graphs are the keys to our proofs. Some open problems are provided as further motivations of our study.

[17] Wai-Tong Louis Fan, Wenqing Hu, and Grigory Terlov. Wave propagation for reaction-diffusion equations on infinite random trees. *Communications in Mathematical Physics*, 384(1):109–163, 2021.

The asymptotic wave speed for FKPP type reaction-diffusion equations on a class of infinite random metric trees are considered. We show that a travelling wavefront emerges, provided that the reaction rate is large enough. The wavefront travels at a speed that can be quantified via a variational formula involving the random branching degrees $\vec{d}$ and the random branch lengths $\vec{\ell}$ of the tree $T_{\vec{d},\vec{\ell}}$. This speed is *slower than* that of the same equation on the real line R, and we estimate this slow down in terms of $\vec{d}$ and $\vec{\ell}$. the Galton-Watson tree whose offspring distribution is supported on a finite set of positive integers. The key idea is to project the Brownian motion on the tree onto a one-dimensional axis along the direction of the wave propagation. The projected process is a multi-skewed Brownian motion, introduced by Ramirez [**?**], with skewness and interface sets that encode the metric structure $(\vec{d},\vec{\ell})$ of the tree. Combined with analytic arguments based on the Feynman-Kac formula, this idea connects our analysis of the wavefront propagation to the large deviations principle (LDP) of the multi-skewed Brownian motion with random skewness and random interface set. Our LDP analysis involves delicate estimates for an infinite product of 2×2

random matrices parametrized by $\vec{d}$ and $\vec{\ell}$ and for hitting times of a random walk in random environment.

[18] Wai-Tong Louis Fan, Michael Jolly, and Ali Pakzad. Three-dimensional shear driven turbulence with noise at the boundary. *Nonlinearity*, 34(7):4764–4786, jun 2021.

We consider the incompressible 3D Navier–Stokes equations subject to a shear induced by noisy movement of part of the boundary. The effect of the noise is quantified by upper bounds on the first two moments of the dissipation rate. The expected value estimate is consistent with the Kolmogorov dissipation law, recovering an upper bound as in (Doering and Constantin 1992 Phys. Rev. Lett. 69 1648) for the deterministic case. The movement of the boundary is given by an Ornstein–Uhlenbeck process; a potential for over-dissipation is noted if the Ornstein–Uhlenbeck process were replaced by the Wiener process.

[19] Wai-Tong Louis Fan, Brandon Legried, and Sebastien Roch. Impossibility of consistent distance estimation from sequence lengths under the TKF91 model. *Bull. Math. Biol.*, 82(9):Paper No. 123, 12, 2020.

We consider the problem of distance estimation under the TKF91 model of sequence evolution by insertions, deletions and substitutions on a phylogeny. In an asymptotic regime where the expected sequence lengths tend to infinity, we show that no consistent distance estimation is possible from sequence lengths alone. More formally, we establish that the distributions of pairs of sequence lengths at different distances cannot be distinguished with probability going to one.

[20] Wai-Tong Louis Fan, Brandon Legried, and Sebastien Roch. Impossibility of phylogeny reconstruction from k-mer counts. *Annals of Applied Probability*, to appear.

We consider phylogeny estimation under a two-state model of sequence evolution by site substitution on a tree. In the asymptotic regime where the sequence lengths tend to infinity, we show that for any fixed k no statistically consistent phylogeny estimation is possible from k-mer counts over the full leaf sequences alone. Formally, we establish that the joint

distribution of k-mer counts over the entire leaf sequences on two distinct trees have total variation distance bounded away from 1 as the sequence length tends to infinity. Our impossibility result implies that statistical consistency requires more sophisticated use of k-mer count information, such as block techniques developed in previous theoretical work.

[21] Wai-Tong (Louis) Fan and Timo Seppäläinen. Joint distribution of Busemann functions in the exactly solvable corner growth model. *Probab. Math. Phys.*, 1(1):55–100, 2020.

The 1+1-dimensional corner growth model with exponential weights is a centrally important exactly solvable model in the Kardar–Parisi–Zhang class of statistical mechanical models. While significant progress has been made on the fluctuations of the growing random shape, understanding of the optimal paths, or geodesics, is less developed. The Busemann function is a useful analytical tool for studying geodesics. We describe the joint distribution of the Busemann functions, simultaneously in all directions of growth. As applications of this description we derive a marked point process representation for the Busemann function across a single lattice edge and calculate some marginal distributions of Busemann functions and semi-infinite geodesics.

[22] Chanh Kieu, Weiran Cai, and Wai-Tong Louis Fan. On the existence of low-dimensional chaos of tropical cyclone intensity. *arXiv preprint arXiv:2110.05190*, 2021.

This study examines the potential limit in the reliability of tropical cyclone (TC) intensity prediction. Using the phase-space reconstruction method for TC intensity time series, it is found that TC dynamics contains low-dimensional chaos at the maximum intensity equilibrium. Examination of several attractor invariants including the largest Lyapunov exponent, the Sugihara-May correlation, and the correlation dimension captures a consistent range of the chaotic attractor dimension between 4-5 for TC intensity. In addition, the error doubling time estimated from the largest Lyapunov exponent for TC intensity is roughly 1-3 hours, which accords with the decay time obtained from the Sugihara-May correlation at the

maximum intensity equilibrium. Furthermore, the findings in this study reveal a relatively short limit for TC intensity predictability based on the traditional maximum surface wind, which is ∼3-9 hours after reaching the mature stage, but noticeably longer for the minimum central pressure (∼12-18 hours). So long as the traditional metrics for TC intensity such as the maximum surface wind or the minimum central pressure is used for intensity forecast, our results support that TC intensity forecast errors will not be reduced indefinitely in any operational model, even in the absence of all model and observational errors. As such, the future improvement of TC intensity forecast should be based on different metrics beyond the absolute intensity errors that are currently used in real-time intensity verification.

[23] Wang-Chun Kwok, Ka-Chun Wong, Ting-Fung Ma, Ka-Wai Ho, Louis Wai-Tong Fan, King-Pui Florence Chan, Samuel Shung-Kay Chan, Terence Chi-Chun Tam, and Pak-Leung Ho. Modelling the impact of travel restrictions on covid-19 cases in hong kong in early 2020. *BMC public health*, 21(1):1–8, 2021.

Coronavirus Disease 2019 (COVID-19) led to pandemic that affected almost all countries in the world. Many countries have implemented border restriction as a public health measure to limit local outbreak. However, there is inadequate scientific data to support such a practice, especially in the presence of an established local transmission of the disease. Here we develop a new metapopulation Susceptible-Exposed-Infectious-Recovered (SEIR) model with inspected migration to investigate the effect of border restriction as a public health measure to limit outbreak of coronavirus disease 2019.

[24] Brian McLoone, Wai-Tong Fan, Adam Pham, Smead Rory, and Laurence Loewe. Stochasticity, selection, and the evolution of cooperation in a two-level moran model of the snowdrift game. *Complexity*, 2018:1–14, 2018.

The Snowdrift Game, also known as the Hawk-Dove Game, is a social dilemma in which an individual can participate (cooperate) or not (defect) in producing a public good. It is relevant to a number of collective action problems in biology. In a

population of individuals playing this game, traditional evolutionary models, in which the dynamics are continuous and deterministic, predict a stable, interior equilibrium frequency of cooperators. Here, we examine how finite population size and multilevel selection affect the evolution of cooperation in this game using a two-level Moran process, which involves discrete, stochastic dynamics. Our analysis has two main results. First, we find that multilevel selection in this model can yield significantly higher levels of cooperation than one finds in traditional models. Second, we identify a threshold effect for the payoff matrix in the Snowdrift Game, such that below (above) a determinate cost-to-benefit ratio, cooperation will almost surely fix (go extinct) in the population. This second result calls into question the explanatory reach of traditional continuous models and suggests a possible alternative explanation for high levels of cooperative behavior in nature.

[25] Mahashweta Patra, Wai-Tong Louis Louis, and Chanh Kieu. Sensitivity of tropical cyclone intensity variability to different stochastic parameterization methods. *Frontiers in Earth Science*, page 820.

Proper representations of stochastic processes in tropical cyclone (TC) models are critical for capturing TC intensity variability in real-time applications. In this study, three different stochastic parameterization methods including random initial conditions, random parameters, and random forcing are used to study TC intensity variation and uncertainties. Using a fidelity-reduced dynamical model and a cloud-resolving model (CM1), it is shown that random forcing produces the largest variability of TC intensity at the maximum intensity equilibrium and the fastest intensity error growth during TC rapid intensification. In contrast, random initial condition tends to be more effective during the early stage of TC development but becomes less significant at the mature stage. For the random parameter method, it is found that this approach depends sensitively on how the model parameters are randomized. Specifically, randomizing model parameters at the initial time appears to produce much larger effects on TC intensity variability and error growth as compared to randomizing model parameters every model time step, re-

gardless of how large the random noise amplitude is. These results suggest the importance of choosing a random representation scheme to capture proper TC intensity variability in practical applications.

[26] Javier Rubio-Herrero, Carlos Ortiz Marrero, and Wai-Tong Louis Fan. Modeling atmospheric data and identifying dynamics temporal data-driven modeling of air pollutants. *Journal of Cleaner Production*, 333:129863, 2022.

Atmospheric modeling has recently experienced a surge with the advent of deep learning. Most of these models, however, predict concentrations of pollutants following a data-driven approach in which the physical laws that govern their behaviors and relationships remain hidden. With the aid of real-world air quality data collected hourly in different stations throughout Madrid, we present an empirical approach using data-driven techniques with the following goals: (1) Find parsimonious systems of ordinary differential equations via sparse identification of nonlinear dynamics (SINDy) that model the concentration of pollutants and their changes over time; (2) assess the performance and limitations of our models using stability analysis; (3) reconstruct the time series of chemical pollutants not measured in certain stations using delay coordinate embedding results. Our results show that Akaike's Information Criterion can work well in conjunction with best subset regression as to find an equilibrium between sparsity and goodness of fit. We also find that, due to the complexity of the chemical system under study, identifying the dynamics of this system over longer periods of time require higher levels of data filtering and smoothing. Stability analysis for the reconstructed ordinary differential equations (ODEs) reveals that more than half of the physically relevant critical points are saddle points, suggesting that the system is unstable even under the idealized assumption that all environmental conditions are constant over time.

[27] Krishn SR Kaur S, Smith LM, Johansson SL, Jain M, Patel A, Gautam SK, Hollingsworth MA, Mandel U, Clausen H, Lo WC, Fan WT, Manne U, and Batra SK. Mucins and associated glycan signatures in colon

adenomacarcinoma sequence: prospective pathological implications for early diagnosis of colon cancer. *Cancer Letters*, 374(2):304–314, 2017.

Development of biomarkers that detect early stage resectable premalignant lesions of colon can provide critical aid in the prevention of colorectal cancer. Recent lines of evidence suggest the utility of mucin expression to predict malignant transformation of colon pre-neoplastic lesions. In this study, we investigated the combined expression of multiple mucins and mucin-associated glycans during the adenoma-carcinoma sequence of colon cancer progression. Further, we evaluated their applicability as markers for differentiating adenomas/adenocarcinomas from hyperplastic polyps. Immunohistochemical analyses performed on colon disease tissue microarrays revealed downregulation of MUC2 and MUC4 expression (p ¡ 0.0001) while MUC1 and MUC5AC expressions were upregulated (p = 0.01) during adenoma-adenocarcinoma progression. Expression of MUC17 was downregulated in inflamed tissues compared to normal tissues, but its increased expression differentiated adenomas (p = 0.0028) and adenocarcinomas (p = 0.025) from inflammation. Glycan epitope-Tn/STn on MUC1 showed higher expression in hyperplastic polyps (p = 0.023), adenomas (p = 0.042) and adenocarcinomas (p = 0.0096) compared to normal tissues. Multivariate regression analyses indicated that a combination of MUC2, MUC5AC, and MUC17 could effectively discriminate adenoma-adenocarcinoma from hyperplastic polyps. Altogether, a combined analysis of altered mucins and mucin-associated glycans is a useful approach to distinguish premalignant/malignant lesions of colon from benign polyps.