A structurally unique *Fusobacterium nucleatum* tannase provides detoxicant activity against gallotannins and pathogen resistance

José Miguel Mancheño¹, Estíbaliz Atondo², Julen Tomás-Cortázar²,³, José Luis Lavín⁴&, Laura Plaza-Vinuesa⁵, Itziar Martín-Ruiz², Diego Barriales², Ainhoa Palacios², Claudio Daniel Navo⁶, Leticia Sampedro², Ainize Peña-Cearra²,⁷, Miguel Ángel Pascual-Itoiz², Janire Castelo², Ana Carreras-González², Donatello Castellana⁸, Aize Pellón², Susana Delgado⁹, Patricia Ruas-Madiedo⁹, Blanca de las Rivas⁵, Leticia Abecia²,⁷, Rosario Muñoz⁵, Gonzalo Jiménez-Osés⁶, Juan Anguita*²,¹⁰, Héctor Rodríguez*².

**Supplementary Figures**
Fig S1. Genetic context of the tannase gene (FN_0616) within the *F. nucleatum subsp. nucleatum 25586* genome. The arrows indicate the predicted direction of transcription. Predicted functions of the proteins encoded by each gene are also indicated. The numbers below indicate the base pair location in the chromosome.
**Fig. S2.** TanB\textsubscript{Fan} is monomeric in solution. (A) Elution profile of TanB\textsubscript{Fan} from a HiLoad® 16/600 Superdex® 75 pg Size Exclusion Column and (B) the corresponding SDS-PAGE gel representing fractions from C5 to C12. The monomeric protein eluted around 60 ml between Canalbumin (C, 75 kDa) and Ovalbumin (O, 43 kDa) in concordance with the denatured SDS-PAGE.
Fig S3. **TanB\textsubscript{Fnn} substrate specificity.** Activity of TanB\textsubscript{Fnn} on gallic acid esters (A), protocatechuic acid esters (B) and complex tannins (C). The substrates used were: MG-methyl gallate, EG-ethyl gallate, PG-propyl gallate, LG-lauryl gallate, EPGCG-epigallocatechin gallate, GCQG-gallocatechin gallate, E3,4DB-ethyl 3,4-dihydroxybenzoate, E3,5DB-ethyl 3,5-dihydroxybenzoate, TA-tannic acid. The products generated were: GA-gallic acid, PA-protocatechuic acid. See also Fig. 2.
Figure S4. Biochemical properties of TanB<sub>Fnn</sub>. (A) Relative enzymatic activity of TanB<sub>Fnn</sub> over a range of temperatures. The observed maximum activity was defined as 100%. (B) Relative enzymatic activity of TanB<sub>Fnn</sub> using methyl gallate as substrate over a range of pH. The observed maximum activity was defined as 100%. (C) Relative enzymatic activity of TanB<sub>Fnn</sub> after preincubation of the purified protein at 22, 30, 37, 45, 55 and 65 °C for the indicated lengths of time. Enzyme activity without preincubation under standard conditions was considered as control (100%). See also Fig. 2.
Figure S5. Amino acid sequence and active site comparisons between TanB<sub>Fnn</sub> and TanA<sub>Lp</sub>. (A) Amino acid sequence alignment showing the elements of secondary structure of TanB<sub>Fnn</sub> and TanA<sub>Lp</sub>. (B) The superimposition of the structures of TanB<sub>Fnn</sub> (blue sticks) and TanA<sub>Lp</sub> (orange sticks) reveals the conservation of the catalytically indispensable residues. Numbering corresponds to TanB<sub>Fnn</sub>. See also Fig. 3.
Figure S6. Representative snapshots of three different 1 microsecond MD simulations of ethyl gallate (EG, in red) bound to TanBFnn. Despite the flap being in a closed conformation throughout the entire simulations, many orientations of EG inside of the active were observed, in which different phenolic hydroxyl groups alternatively form hydrogen bonds with negatively charged residues Glu$^{365}$ and Asp$^{446}$. In one of the trajectories, EG abandoned the active site and migrated to the bulk solvent after around 700 ns. Residues involved in binding, catalysis and located at the flexible flap are shown as green, blue and yellow sticks, respectively. Non-polar hydrogens, solvent molecules and buffer ions have been omitted for clarity. See also Fig. 4.
Figure S7. Representation of MD simulations of methyl digallate (MDG, in red) bound to TanB_{Fnn}. Representative snapshot of a 1 microsecond MD simulation of methyl digallate (MDG, in red) bound to Tan Fn. a) the flexible lid (in light brown) was found to close over the substrate correctly positioning it inside the active site. b) Highly predominant binding mode of MDG to the TanB_{Fnn} active site though a combination of hydrogen bond and van der Waals interactions, as determined through MD simulations. Residues involved in binding, catalysis and located at the flexible flap are shown as green, blue and yellow sticks, respectively. Non-polar hydrogens, solvent molecules and buffer ions have been omitted for clarity. See also Fig. 4.
**Figure S8. SPD interacts with TanB<sub>Fnn</sub>.** Thermal denaturation of TanB<sub>Fnn</sub> in the absence (A, B) or presence (C, D) of 50 mM of spermidine. Panels A and C represent the raw data. Panels B and D represent the first derivative. Two independent determinations are shown.
Figure S9. Structure of spermidine and its N1 and N8-acetylated derivatives in their charged form at neutral pH.
Figure S10. Multiple sequence alignment of TanB_{Fnn} with homologs found within the order Fusobacteriales, particularly, from the genus *Leptotrichia*: *Leptotrichia sp. OH3620_COT-345* (WP_125237392.1); *L. wadei* (WP_146961278.1); *L. massiliensis* (WP_071123737.1); *L. buccalis* (WP_157859547.1). The NCBI reference is indicated in parenthesis for each protein sequence. The secondary structural elements of TanB_{Fnn} are displayed. The figure was prepared with ESPript 3.0 (Robert and Gouet, 2014).
Figure S11. Multiple sequence alignment of TanB_{Fnn} with homologs from bacterial genera that are prevalent in the oral cavity: *Aggregatibacter* (WP_006718097.1); *Selenomonas* (WP_051586482.1); *Veillonella* (WP_127058572.1); *Mitsuokella* (WP_118475533.1). The NCBI reference is indicated in parenthesis for each protein sequence. The secondary structural elements of TanB_{Fnn} are displayed. The figure was prepared with ESPript 3.0 (Robert and Gouet, 2014).
Figure S12. Multiple sequence alignment of TanB<sub>Fmin</sub> with homologs from bacterial genera present in the human gut: *Leptotrichia* (WP_157859547.1); *Veillonella* (WP_127058572.1); *Megasphaera* (WP_044503365.1); *Roseburia* (WP_173884414.1); *Bacteroides* (WP_164733466.1); *Parabacteroides* (WP_172733427.1); *Lachnoclostridium* (WP_087219107.1); *Fusicatenibacter* (WP_173816499.1); *Dorea* (WP_161159564.1); *Blautia* (WP_095175021.1); *Lactobacillaceae* (WP_054397147.1); *Sutterella* (GenBank: RGU74189.1). The NCBI reference is indicated in parenthesis for each protein sequence except that from Sutterella. The secondary structural elements of TanB are displayed. The figure was prepared with ESPript 3.0 (Robert and Gouet, 2014).

REFERENCES
Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014 42(Web Server issue):W320-4.