Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations | Graduate School

2016

Involvement of Signal Peptidase I in Streptococcus sanguinis Biofilm Formation

Jessica Aynapudi
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

 Part of the Oral Biology and Oral Pathology Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/4451

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

© Jessica Aynapudi, 2016

All Rights Reserved

INVOLVEMENT OF SIGNAL PEPTIDASE I IN *STREPTOCOCCUS SANGUINIS* BIOFILM FORMATION

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University School of Medicine

By

JESSICA AYNAPUDI

B.S., University of Texas at Austin, 2008

Director: PING XU, PH.D.

Professor, VCU School of Dentistry

Virginia Commonwealth University

Richmond Virginia

June 2016

Acknowledgements

I would like to sincerely thank my advisor Dr. Ping Xu who gave me the opportunity to work on this project and supported me throughout this past year in the lab. I would also like to thank all of my colleagues in the lab: Dr. Xiuchun Ge, Dr. Victoria Stone, and Dr. Weihua Chen for all their help. I would like to thank my committee and especially Dr. Todd Kitten for the stimulating conversations and tremendous guidance. I am very grateful to Dr. Scott Henderson and Ms. Frances White for their continuous support and assistance with the confocal laser scanning microscope. Lastly, I would like to thank Fadi El-Rami. His love for research is both inspiring and contagious. I am forever thankful to him for his mentorship, unwavering support, and especially the math problems.

Table of Contents

Abstract

INVOLVEMENT OF SIGNAL PEPTIDASE I IN STREPTOCOCCUS SANGUINIS BIOFILM FORMATION

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University School of Medicine

by

JESSICA AYNAPUDI

B.S., University of Texas at Austin, 2008

Major Director: Ping Xu, Ph.D.

Professor, VCU School of Dentistry

Virginia Commonwealth University

June 2016

Biofilm accounts for 65%-80% of microbial infections in humans. Considerable evidence links biofilm formation to oral disease and consequently systemic infections. Eradication of biofilm-associated infections is important. *Streptococcus sanguinis*, a Gram-positive bacterium, is one of the most abundant species in oral biofilm. It contributes to biofilm development in oral cavities and is one of the recognized causes of infective endocarditis. To study and identify biofilm genes in *S. sanguinis*, biofilm formation of 51 mutants was compared with the wild type SK36 strain using crystal violet (CV) staining in a microtiter plate. Confocal laser scanning microscopy (CLSM) and image analysis was done to compare biofilm formation by the mutant to the wild type SK36 strain. A biofilm mutant XG2_0351, encoding a type I signal peptidase (SPase I), was further investigated. SPase I cleaves proteins that are transported through secretory machinery and is necessary for the release of translocated preproteins from a cytoplasmic site of synthesis to extracytoplasmic/membrane destinations. *S. sanguinis*, like many Gram-positive bacteria, has multiple SPases I. The objective of this project is to investigate the distinctive role that SPase I plays in biofilm formation in *S. sanguinis*. Using a plate reader, the growth curves of the wild type strain SK36 and XG2_0351 were compared. The scanning electron microscope (SEM) was utilized to compare the cell surface morphologies. Coomassie staining was done to narrow the list of potential substrates of XG2_0351. CV staining and CLSM images indicated phenotypic differences between the SPase I mutant and SK36. The growth curves of XG2_0351 and SK36 showed no significant difference although SEM illustrated a difference in the cell surface morphologies. Coomassie staining illustrated a number of substrates that were present in SK36 but not XG2_0351. In addition bioinformatics was used to understand the gene function. In conclusion, XG2_0351 reduces biofilm formation in *S. sanguinis* but further research is necessary to elucidate the specific proteins that are involved. Clarifying the

role that SPase I plays in reduced biofilm formation in *S. sanguinis* will give a better understanding of the biofilm formation mechanism.

List of Tables

List of Figures

INTRODUCTION

The highly developed oral microbiome is a diverse environment that is largely influenced by oral streptococci (Kreth, Merritt, & Qi, 2009). *Streptococcus sanguinis*, a gram-positive bacterium, is a known pioneer colonizer on freshly cleaned tooth surfaces and one of the abundant species in oral biofilm (Ge et al., 2008; Jenkinson & Lamont, 2005). As most members of the viridans group, *S. sanguinis* oxidizes hemoglobin in erythrocytes by secretion of H_2O_2 and produces alpha-hemolysis on blood agar (Barnard & Stinson, 1996; Xu et al., 2007). Through a variety of mechanisms, normal inhabitant *streptococci* form dental plaque which is involved in the development of caries. To initially colonize the tooth surface and form dental plaque, streptococci serve as a tether for the attachment of other microorganisms to a salivary glycoprotein-coated surface (Kolenbrander & London, 1993; Xu et al., 2007). One of these microorganisms is *S. mutans*, whose overgrowth is often associated with the development of dental caries (Kreth, Merritt, Shi, & Qi, 2005; Loesche, 1986). The shift from a healthy to cariogenic streptococcal environment is caused by coexistence and competition of interspecies interactions in a microbial community. Steering the outcome of interactions between species are determining environment conditions, such as nutritional availability, cell density, and pH (Kreth et al., 2005). Significantly higher number of *S. sanguinis* is reported in healthy subjects whereas there is almost no detectable level in those subjects with caries. An inverse relationship between commensal and pathogenic streptococci exists where a high levels *S. sanguinis* correlated with subsequent delayed acquisition of *S. mutans* (Caufield et al., 2000; Kreth et al., 2005). Through a well-regulated production of chemicals, such as H_2O_2 by *S. sanguinis* and mutacins by *S. mutans* these interspecies interactions are mediated (Kreth et al., 2005). Thus we can presume that the ability of *S. sanguinis* to interfere with the colonization of *S. mutans* on a tooth may be

beneficial for oral health (Caufield et al., 2000; Xu et al., 2007). Although *S. sanguinis* is a member of the normal flora in the oral cavity and is considered benign, it has the potential to be pathogenic in patients through bacteremia (Ge et al., 2008; Turner, Das, Kanamoto, Munro, & Kitten, 2009; Xu et al., 2007). This opportunistic pathogen infection could lead to infective endocarditis or cause fatality in patients who are neutropenic (Bochud, Calandra, & Francioli, 1994; Ge et al., 2008; Xu et al., 2007).

Biofilm

Biofilm is an accumulation of microorganisms embedded in a protective extracellular polymeric matrix that adheres to biotic or abiotic surfaces in nature (Hall, McGillicuddy, & Kaplan, 2014). This accumulation of either a single or multiple species lives in a nutrient-sufficient ecosystem as a sessile microbial community. Biofilm exhibits a distinct physiologically altered pattern when compared to the gene expression and protein production of planktonic cells (Costerton, Stewart, & Greenberg, 1999; Donlan, 2002; Hall et al., 2014). Furthermore, an extracellular polymeric substance (EPS) matrix is produced to differentiate the biofilm-associated cells from the suspended planktonic cells (Donlan, 2000). The composition and structure of the polysaccharides in the EPS matrix determine their primary conformation. Usually the EPS in not uniform but varies spatially and temporally (Donlan & Costerton, 2002; Leriche, Sibille, & Carpentier, 2000). It is possible that the high level of hydration of the EPS prevents desiccation in some natural biofilms. The EPS can also impede the mass transport of antibiotics through the biofilm, which may promote their antimicrobial resistance properties (Donlan, 2000; Donlan, 2002). The formation of biofilm occurs in five major stages: initial attachment, irreversible attachment, maturation I, maturation II and dispersal.

Microbial surface components mediate attachment through surface proteins by recognizing adhesive matrix molecules (MSCRAMMs) (Foster, Geoghegan, Ganesh, & Hook, 2014; Lister & Horswill, 2014). Initially, bacteria adhered to the surface begin to multiply and differentiate, strengthening the attachment (Lister & Horswill, 2014). The maturation process occurs through the up-regulation of virulence factors, secretion of extra-cellular polymers, consumption of soluble nutrients, and recruitment of other bacterial species (Hall et al., 2014). When a biofilm is fully established it has a defined structure where the environment is conducive for the exchange of genetic material between cells (Donlan, 2002). Biofilm growth is an important advantage for bacteria because it provides a defense system against immune defenses such as macrophages. This can result in "frustrated phagocytosis" (Lister & Horswill, 2014; Scherr, Heim, Morrison, & Kielian, 2014). Biofilms demonstrate quorum sensing, which is cell-to-cell signaling that plays a role in cell attachment and detachment (Donlan, 2000). The cells of a biofilm may disperse by detachment caused by nutrient levels or quorum sensing, shearing of biofilm aggregates because of flow effects, or shedding of daughter cells from actively growing cells (Donlan, 2002). Eventually, individual cells from the original biofilm can disperse to start infection at new sights or mediate an acute infection (Costerton et al., 1999; Lister & Horswill, 2014).

According to estimates by the U.S. Centers for Disease Control and Prevention and the National Institutes of Health, biofilm accounts for 65%-80% of microbial infections in human beings (Donlan, 2002; Hall et al., 2014). Infections that are biofilm-based have been discovered in almost all tissues of the human body (Hall et al., 2014). Many studies have shown that there is considerable evidence linking biofilm formation in the oral cavity to oral disease and consequently systemic infections. These systemic conditions include cardiovascular disease, diabetes mellitus, preterm or low birth weights, rheumatoid arthritis, and infective endocarditis

(Hall et al., 2014; Seymour, Ford, Cullinan, Leishman, & Yamazaki, 2007). Biofilm growth has the ability to use a variety of defense mechanisms against infection so treatment and eradication of biofilm-associated infections are problematic and difficult (Hall et al., 2014). Cells in a biofilm show increased tolerance to antibiotics through different mechanisms. The biofilm matrix blocks access to actively growing cells by decreasing the antibiotic diffusion rates or physiologically dormant persister cells which are inherently resistant to antibiotics (de la Fuente-Nunez, Reffuveille, Fernandez, & Hancock, 2013; Lister & Horswill, 2014; Singh, Ray, Das, & Sharma, 2009)

Infective Endocarditis

Infective endocarditis (IE) is a potentially life-threatening microbial infection of the heart valves or endocardium (Cahill & Prendergast, 2015; Hoen & Duval, 2013; Selton-Suty et al., 2012). This disease has an estimated annual incidence of 3 to 9 cases per 100,000 people in industrialized countries (Hoen & Duval, 2013; Selton-Suty et al., 2012). In 2016, there are many emerging treatments and therapies for endocarditis but the 1-year mortality rate remains at 30% (Cahill & Prendergast, 2015). If not treated, IE can result in severe complications such as congestive heart failure and can become fatal.

Patients with prosthetic valves, intracardiac devices, unrepaired cyanotic congenital heart diseases, or a family history of IE have the highest rates of this illness. However, 50% of incidences of IE develop in patients with no known history of valve disease. There are several other risk factors for IE such as chronic rheumatic heart disease, hemodialysis, diabetes, HIV, and intravenous drug use (Hoen & Duval, 2013). In the United States, more than one third of the cases of IE are reported to be health care-associated. When several predisposing factors are

associated with age, the increased number of cases of IE among persons 65 years of age or older is clear (Hoen & Duval, 2013; Selton-Suty et al., 2012).

Together, streptococci and staphylococci account for 80% of cases of IE. The proportions vary with regards to the source of infection, patient age, coexisting conditions, and valve (native vs. prosthetic) (Hoen & Duval, 2013). Recent studies have shown that staphylococci, specifically *S. aureus*, have surpassed streptococci as the most common cause of IE. IE by the streptococcal oral viridians group remain the most common in low-income countries (Cahill & Prendergast, 2016; Yew & Murdoch, 2012). Although *S. sanguinis* has no direct role in oral disease, it is often implicated as the most common streptococci isolated from patients (Mylonakis & Calderwood, 2001).

Pathogenesis of Infective Endocarditis

Conventional IE results when there is colonization of damaged valvular endothelium by circulating bacteria with specific adherence properties. Lesions that cause endothelial damage may be the result of turbulent blood, catheters, electrodes, or by repeated intravenous-drug use (Hoen & Duval, 2013). Certain types of congenital or acquired heart disease cause turbulent blood flow, traumatizing the endothelium and causing the deposition of fibrin and platelets on the damaged endocardium or endothelial surface. IE results when microbes invade the bloodstream and colonize this damaged site. Microbial organisms cause IE when they disseminate into the bloodstream and the intricately composed biofilm within the gingival tissue niche is disrupted. A disturbance of the delicate barrier between the oral biofilm and host tissues may cause periodontitis, gingivitis, pulpal or root canal infections. There is a heightened risk of bacteremia when an increase of inflammation results from oral procedures and even routine oral

procedures such as brushing, flossing, and chewing (Parahitiyawa, Jin, Leung, Yam, & Samaranayake, 2009).

IE clinical manifestations can vary dramatically depending on the specific case. In 80% of cases fever is a prevalent symptom (Hoen & Duval, 2013; Selton-Suty et al., 2012). Also, a new murmur and worsening of a known murmur are reported. Less common signs are splenomegaly, splinter hemorrhages, Janeway's lesions, Roth's spots, conjunctival hemorrhage, sepsis, meningitis, unexplained heart failure, septic pulmonary emboli, stroke, acute peripheral arterial occlusion, and renal failure (Hoen & Duval, 2013; Richet et al., 2008). The most severe extracardiac problems of infective endocarditis are cerebral complications (Hoen & Duval, 2013; Sonneville et al., 2011; Thuny et al., 2007).

Diagnostic methods for IE generally rely on clinical, microbiologic, and echocardiographic findings. To identify the causative microorganism, three sets of blood cultures are performed and the pathogen is identified in 90% of cases. Transthoracic echocardiography is performed to diagnose valvular lesions (Hoen & Duval, 2013). The polymerase chain reaction can be utilized to identify unculturable organisms in excised vegetations or systemic emboli (Beynon, Bahl, & Prendergast, 2006; Mylonakis & Calderwood, 2001). Appropriate treatment for patients with IE revolves around prolonged bactericidal antibiotic treatment to eradicate the causative pathogen and possible surgery to remove the infected material (Hoen & Duval, 2013).

If there is no need for cardiac surgery, effective treatment for IE using antimicrobial agents begins in the hospital. Often this treatment is completed on an outpatient basis once the fever has resolved and follow-up blood cultures are negative. For common causes of IE, prolonged administration of a bactericidal antimicrobial agent or combination is currently recommended (Mylonakis & Calderwood, 2001). Granted, it is advised to use combination therapy over

monotherapy to reduce the potential for resistance development and to provide a concerted antimicrobial effect (Cahill & Prendergast, 2016). There is a high frequency of adverse events in patients who are being treated for IE so therapy revision is important (Mylonakis & Calderwood, 2001). The antimicrobial regimen can be modified depending on resistance patterns, severity of infection, presence or absence of prosthetic material, and culture results (Cahill & Prendergast, 2016). The minimal inhibitory concentration (MIC) of penicillin is necessary to determine the optimal therapy for streptococcal infection. Antibiotic therapy for IE caused by some microorganisms is frequently unsuccessful, and surgery is generally recommended (Mylonakis & Calderwood, 2001).

IE presents many challenges because it is heterogeneous and complex by nature. Even though there have been many advances in diagnostic procedures, antimicrobial treatments, and cardiovascular imaging it remains a serious threat to many lives (Cahill & Prendergast, 2015). As of 2015, IE has an incidence of 3 to 10 per 100,000 and has an in-hospitality of 20% (Cahill & Prendergast, 2015; Cahill & Prendergast, 2016). Therefore, it is imperative to study the causative agents and virulence factors of IE.

Importance of Signal Peptidase I

Almost one-third of all proteins need to be translocated through or into the cytoplasmic membrane because they function outside of the cytosol (Auclair, Bhanu, & Kendall, 2012). These preproteins are directed to the Sec- or Tat-translocation pathway by the signal sequence, an amino-terminal extension. In prokaryotes, signal peptidases (SPases) are classified into three groups: SPase I, II, and IV. SPase II and IV are necessary for cleaving signal peptides from lipoproteins and prepilin proteins, respectively. SPase I produces mature non-lipoproteins that

are then transported in bacteria by the general secretion (Sec) pathway. However there is recent evidence that indicates that SPase I can also be transported by means of the twin arginine translocation (Tat) pathway (Auclair et al., 2012; Luke, Handford, Palmer, & Sargent, 2009). SPase I is essential to the cell for the release of translocated preproteins from the membrane when they are transported from a site of cytoplasmic synthesis to extracytoplasmic locations (Auclair et al., 2012; du Plessis, Nouwen, & Driessen, 2011). The protein extracytoplasmic location and specific secretion pathway destination are determined by the signal peptide which marks the protein with a zipcode. The signal peptidase enzyme has the responsibility of cleaving the signal peptide from the preprotein once the majority of it is translocated. This enzymatic action allows the protein to release from the membrane and correctly fold into a mature protein. SPase enzymes are critical for cell survival because without them, accumulation of preproteins at the membrane would occur and have a deleterious effect on the growth of the cell (Dalbey & Wickner, 1985; Auclair et al., 2012). Bacterial species such as *E. coli* have only one essential SPase I enzyme whiles others such as *B. subtilis* have multiple enzymes (Auclair et al., 2012; Meijer et al., 1995; Tjalsma et al., 1997; Tjalsma et al., 1998). *S. sanguinis* possesses two SPase I enzymes: XG2_0351 AND XG2_0849. The bacterial SPase I belongs to a unique group of serine endoproteases, which use a Ser-Lys catalytic dyad instead of the standard Ser-His-Asp triad utilized by eukaryotes. This distinctive feature makes SPase I a desirable antimicrobial target (Rawlings & Barrett, 1993; Auclair et al., 2012).

Purpose of Study

Identifying which genes in *S. sanguinis* can cause biofilm formation and their mechanisms is an important step in being able to effectively prevent and treat systemic infections that originate in the oral cavity. In this study we attempted to elucidate the role that SPases I plays in biofilm formation in *S. sanguinis* when a paralogue is available. In the future, *S. sanguinis* will be used as a model for further study of biofilm genes in other types of bacteria.

Genome-wide deletion mutants of *S. sangunis* strain SK36 have been constructed in our laboratory. The comprehensive library of deletion mutants of SK36 provided the unique opportunity to apply a systems biology approach to investigate the effect of genetic mutations on biofilm formation.

This study involved the observation of phenotypic characteristics of biofilm formation by the non-essential deletion mutants and the wild type of *S. sanguinis* SK36 using crystal violet staining and confocal laser scanning microscopy. Cell morphologies for selected mutant XG2_0351 and SK36 were compared through visualization by scanning electron microscopy. To compare the growth between SK36 and XG2_0351 the growth curves were examined using a plate reader. Computational prediction data was used to gain more knowledge about gene interactions and functional similarities to other species. Finally, coomassie staining was done to narrow the list of potential substrates of XG2_0351 involved in biofilm formation.

METHODS AND MATERIALS

Bacterial strain and mutants

S. sanguinis strain SK36 that was used in this study (Table 1) was isolated from human dental plaque (Kilian, Mikkelsen, & Henrichsen, 1989). Single gene deletion mutants for the SK36 strain were previously constructed using a recombinant PCR method (Figure 1). Utilizing the complete *S. sanguinis* SK36 genome sequence, three sets of primers (F1/R1, F2/R2, and F3/R3) were designed. Primers were constructed to amplify a linear DNA fragment containing a kanamycin resistance cassette flanked by upstream and downstream sequences of the targeted gene. The 5' ends of the F2 and R2 of primers of the kanamycin cassette were created to complement the sequences of DNA that flank the target gene (Xu et al., 2011). The linear recombinant PCR amplicons containing the kanamycin cassette (Turner, Das, Kanamoto, Munro, & Kitten, 2009) flanked by *S. sanguinis* DNA were transformed into competent *S. sanguinis* cells and integrated into the *S. sanguinis* genome via double crossover recombination (Ge et al., 2008). A genome-wide mutant library containing 2,048 deletion mutants was constructed using a 96 well high-throughput format (Chen et al., 2011; Xu et al., 2007). To determine putative gene functions proteins were searched against the previously annotated genome (Ge et al., 2008; Xu et al., 2007). Genes of SK36 are referred to as "SSA" followed by the corresponding gene number while mutants are indicated by "XG2" followed by the gene number.

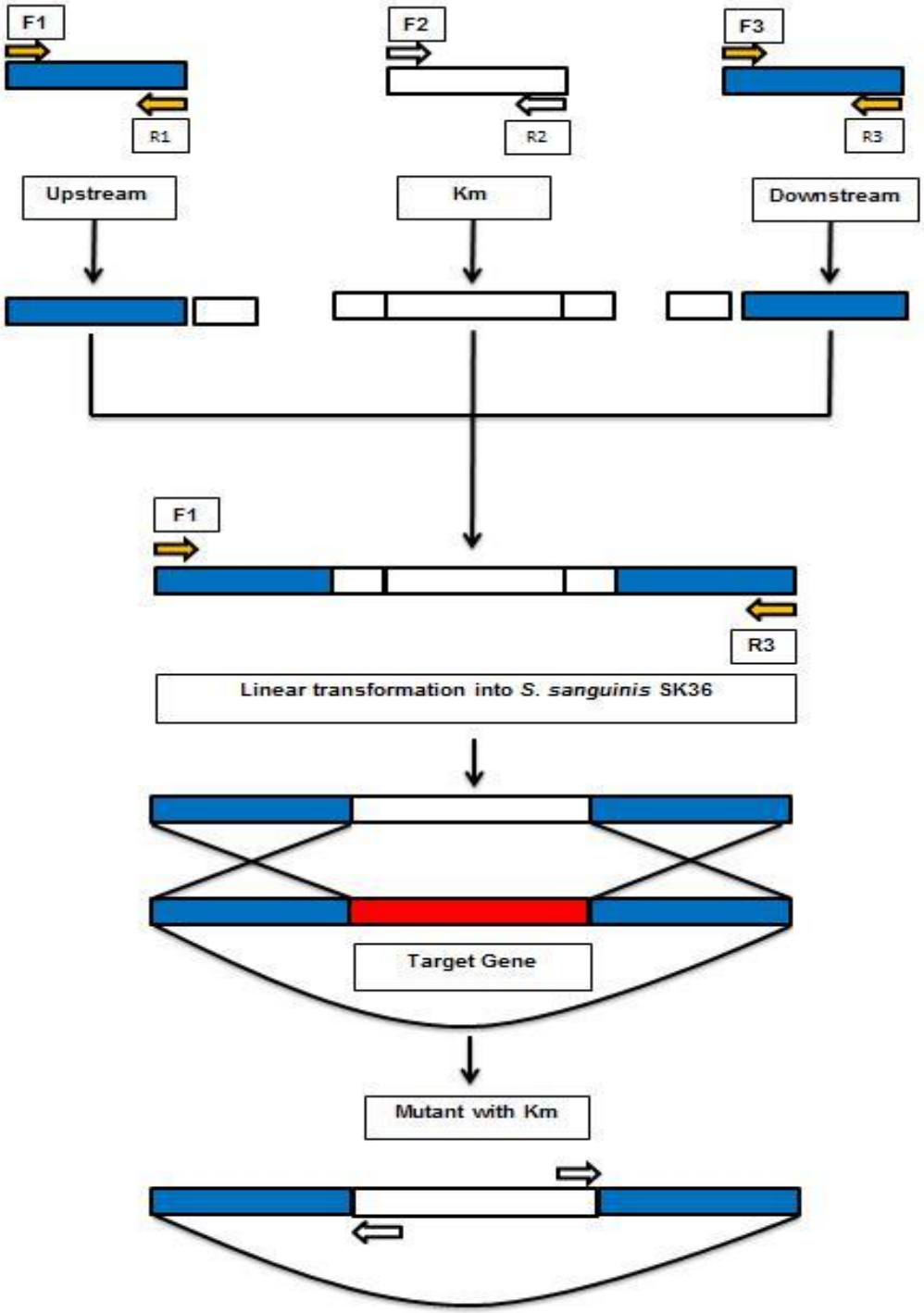
Table 1. Bacterial strains used in study.

Table 1.

Strain	Phenotype	Source
S. sanguinis		
SK36	Human dental plaque isolate	[Kilian 1989, Xu 2007]
SK36 mutants	All non-essential gene deletion mutants of SK36 from genome-wide library	[Xu 2007, Chen 2011]

Figure 1. Schematic representation of the insertion of kanamycin (Km) resistance gene cassette into SK36 chromosome to construct single gene deletion mutants.

Figure 1.


F1
R1
F2
R2
F3
R3
Upstream
Km
Downstream
F1
R3
Linear transformation into S. sanguinis SK36
Target Gene
Mutant with Km

Crystal Violet Assay

To study and identify biofilm genes in *S. sanguinis*, 51 mutants were compared with the wild type SK36 strain using the microtiter dish biofilm formation assay. Using preliminary data from our lab (Table 2), the set of 51 mutants were tested for biofilm formation using the O'Toole method to observe bacterial adherence to an abiotic surface (O'Toole, 2011). This method was also used to test the downstream gene of XG2_0351, as well as its paralogue: XG2_0849. SK36 and mutants were grown anaerobically in BHI and BHI supplemented with kanamycin respectively, and incubated at 37°C for 16 hours. A multichannel pipette was used to inoculate a 96-well plate with 99 µl of biofilm medium supplemented with 1% sucrose and 1 µl of overnight culture of SK36 of each mutant for 8 repeats. One column of 8 wells was loaded with medium alone as a negative control. The total bacterial growth was determined by measuring the absorbance at 450 nm with a Synergy H1 Hybrid Reactor (BioTek, VT, USA) microplate reader. We decanted the media and then removed the remaining planktonic cells by gently rinsing with 200 µl of distilled H_2O. 50 µl of 0.4% (wt/vol) crystal violet (CV) solution was added to each well and was left for 15 minutes to dry. Wells were rinsed three times with 200 µl of distilled H_2O and air-dried. The CV was solubilized by 200 µl of 33% acetic acid. After 30 minutes, 100 µl from each well was transferred to a new plate and the absorbance at 600 nm was measured by a Synergy H1 Hybrid Reactor (BioTek, VT, USA) microplate reader. The results from microtiter staining were statistically analyzed by ANOVA. The significance was set as P-value <0.05 (Ge et al., 2008).

Confocal Laser Scanning Microscopy (CLSM) and Image Analysis

Confocal laser scanning microscopy and image analysis was used to further study and identify biofilm genes in *S. sanguinis*. Six wells in a 12 well-plate were filled with 1000 µl of biofilm medium supplemented with 1% sucrose and inoculated with 10 µl of overnight culture. The first column contained SK36 and the second contained a mutant for 3 repeats each. Overnight biofilms were rinsed 1 time with 1000 µl of PBS to remove the unattached bacteria. For 15 minutes, biofilms were labeled using live staining. This was done with 1.5 µM SYTO9 (a green fluorescent dye that can cross intact membranes). Afterwards the wells were rinsed with 1000 µl of PBS to remove the remaining dye. The biofilms were viewed through a 10x dry lens with a Zeiss LSM 710 confocal laser scanning microscope. Green fluorescence was imaged and an image stack of 1 randomly chosen spot was collected for each sample. The computer program ImageJ analyzed CLSM images. Image stacks were converted to individual Tiff images for the front, middle, and side of each sample. The image stacks of biofilm grown by the mutant were compared to growth by the wild type by using the T-test. This method was also used to test the downstream gene of XG2_0351, as well as its paralogue: XG2_0849.

Confocal Laser Scanning Microscopy and Scanning Electronic Microscopy were performed at the VCU Department of Anatomy and Neurobiology Microscopy Facility, supported in part, by funding from NIH-NCI Cancer Center Support Grant P30 CA016059.

Examination of Mutant Growth in vitro

This experiment was done to elucidate and compare growth rates of SK36 and XG2_0351. The wild type and mutant were inoculated individually and then measured for three trials concurrently using a Synergy H1 Hybrid Reactor (BioTek, VT, USA) microplate reader for 14-16 hours.

Elucidated Gene Functions using Clusters of Orthologous Groups

Using PubMed (http://www.ncbi.nlm.nih.gov/pubmed), we searched for mutant genes that are organized into operons (Table 3). Uniprot (http://www.uniprot.org/) was utilized to search clusters of orthologous groups (COGs) (Table 3).

STRING Analysis and Gene Co-occurrence Network

The STRING database (http://string-db.org/) provided a way to visualize and predict protein-protein interactions of XG2_0351 through bioinformatics data (Figure 7A). The co-occurrence network shows the relationship between XG2_0351 with other bacterial species (Figure 7B).

Scanning Electron Microscopy (SEM) Analysis of *S. sanguinis* WT and Mutant Cells

The cell morphologies of *S. sanguinis* SK36 and selected mutant XG2_0351 were examined using SEM (Figure 5). The two samples were deposited onto a 0.1 μm disposable Millipore filter to remove medium. Samples were fixed using 2% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) for 30 min, followed by 1% osmium tetroxide in 0.1 M sodium cacodylate buffer (pH 7.4). The samples embedded in the filters were then dehydrated in ethanol followed by phosphate-buffered saline (PBS) and allowed to air-dry. The filters were sectioned and mounted

onto stubs and coated with gold for three minutes (EMS– 550 Automated Sputter Coater, Electron Microscopy Sciences, Hatfield, PA). Micrographs were taken at 15,000× total magnification using a Zeiss EVO 50 XVP scanning electron microscope (Carl Zeiss, Peabody, MA).

Coomassie Staining

This staining assay was done to elucidate which proteins XG2_0351 is targeting (Figure 11). We inoculated the stocked strain into 5 ml medium (e.g. BHI) in a 15 ml round-bottom tube. After this we prepared 3 round-bottom tubes of 15 ml containing 5 ml medium, and then incubated them in an anaerobic jar at 37 °C for overnight. Then we transferred 50 µl overnight bacterial culture into each of 3 pre-incubated conical tubes and incubated at the same condition as previously stated to mid log-phase (about 4 hours). All of the following steps were done at 4 °C. Bacterial cells were centrifuged for 10 min at 4,000 rpm using Sorvall Legend RT centrifuge (MN, USA). Following this we added 10 µl protease inhibitor cocktail (Sigma P8430) and DNase (1 µg/ml) to 1 ml non-denaturing lysis buffer (PBS, EDTA=5 mM, pH = 7.4) to 10^6 cells in mid log-phase bacterial culture. We re-suspended the pellets by pipetting up and down several times. The suspension was transferred into the 2 ml Lysing Matrix B Beads. Cells were disrupted in the Fast Prep 24 for 30s at level 4. Then we sonicated for short pulses (5-10 sec) with pauses (10-30 sec) to re-establish a low temperature. We centrifuged for 10s at maximum speed and transferred supernatant into a new tube. Then we determined the volume of the supernatant and quantitated the protein in sample using Pierce BCA Protein Assay kit (Cat #23227, IL, USA). 15 µg of protein were loaded into each well and stained with Coomassie

Brilliant Blue R-250 solution (Bio-Rad, USA) for 4 hours. Finally, we washed with washing buffer (50 % methanol, 10% acetic acid, 40% distilled water) until bands were visible.

RESULTS

We compared 51 mutants to SK36 using crystal violet staining to determine if there is a statistically significant difference in biofilm formation (Figure 2). Using the ANOVA statistical test, we determined t-test with p-values <0.05 were statistically different. We found that 41 mutants formed biofilms that were statistically significant when compared to SK36. Using a multiple comparison method known as Dunnett's test, we found that XG2_0351 has a p value <.0001 when compared to SK36. 25 mutants that showed less biofilm formation than SK36 were viewed via CLSM. Using CLSM, z-stacks were created by stacking successive slices, which were then processed into a 3D image using ImageJ software (Figure 3). The z-stack for XG2_0351 showed effectively no biofilm formation and therefore, was much thinner than SK36. The downstream mutant, XG2_0350 was tested to rule out the possibility of the polar effect (Figure 8). XG2_0351 was selected for further investigation based on phenotypic differences with SK36 that were determined by CV staining and CLSM.

SEM was used to visualize the morphological differences between SK36 and type I signal peptidase mutant XG2_0351 (Figure 5). Biofilm did not form during CV staining or CLSM, XG2_0351 showed growth in SEM images. Although, the mutant chains were much shorter in length and had a different shape than SK36. To gain more insight into the growth differences we looked at the growth curves by using a plate reader (Figure 6). Three separate repetitive trials showed that the growth rate of SK36 and XG2_0351 are not significantly different.

Through computational prediction methods, the STRING database predicted that this signal peptidase I protein interacts with signal recognition particles and several ribosomal proteins (Figure 7A). Gene co-occurrence showed that SPase I shares many similarities with the

Firmicutes phylum (Figure 7B). CLSM was utilized to visualize XG2_0849, the paralogue of XG2_0351, for biofilm formation (Figure 9). The mutant showed a slight difference in biofilm formation when compared to SK36 but lacked the dramatically decreased effect seen by XG2_0351. Finally, the coomassie assay was done to see the concentrations of proteins that are in SK36 but missing in XG2_0351 (Figure 11). Nine substrates of SPase I mutant, XG2_0351, that were previously shown through CV staining and confocal imaging to exhibit reduced biofilm formation, were absent in XG2_0351 in comparison to wild type.

Table 2. Gene annotation of biofilm mutants

Table 2.

Biofilm Mutant	Gene Annotation
XG2_0004	Lipoprotein, putative
XG2_0036	Secreted protein, possible function in cell-wall metabolism (amidase), putative (N-acetylmuramoyl-L-alanine amidase)
XG2_0048	Transcriptional regulator, TetR/AcrR family, putative
XG2_0115	50S ribosomal protein L29, putative
XG2_0144	Transcriptional regulator, TetR family, putative
XG2_0299	Integral membrane protein / nodulin 21-like protein
XG2_0305	Heat-inducible transcription repressor/Conserved hypothetical protein/p-hydroxybenzoic acid efflux subunit AaeB
XG2_0351	Signal peptidase I, putative
XG2_0363	D-alanine/glycine/Na permease, putative
XG2_0364	Serine/threonine:Na+ symporter, putative
XG2_0387	Transcriptional regulator, GntR family, putative
XG2_0460	Multiple antibiotic resistance operon transcription repressor (MarR), putative
XG2_0613	Glucosyltransferase, putative
XG2_0708	Hypothetical protein (isoprenylcysteine carboxyl methyltransferase (ICMT) family protein)
XG2_0758	Ornithine acetyltransferase / amino-acid acetyltransferase, putative
XG2_0767	Diacylglycerol kinase catalytic domain protein, putative
XG2_0805	Collagen-binding surface protein, putative
XG2_0816	Copper transport operon or penicillinase transcription repressor, putative
XG2_0879	Cell division regulator, negative regulator of FtsZ septation ring formation, putative
XG2_0960	Sensor protein ciaH, putative
XG2_1030	Transcriptional regulator, TetR family, putative
XG2_1064	Hypothetical protein (dextransucrase/glucansucrase, N-acetylmuramoyl-L-alanine amidase, family 2)
XG2_1068	Transcriptional regulator, LysR family (capsular polysaccharide biosynthesis regulation), putative
XG2_1103	Hypothetical protein/ABC transporter?/capsular polysaccharide biosynthesis protein?/PAS/PAC sensor signal transduction histidine kinase
XG2_1171	Tyrosine recombinase xerC
XG2_1219	Sortase, putative
XG2_1271	DHH subfamily 1 protein/Exopolyphosphatase-related proteins
XG2_1301	Conserved uncharacterized protein, possible surface protein
XG2_1309	Predicted RNA-binding protein (contains KH domain), very conserved, putative
XG2_1363	FmtA-like protein, putative
XG2_1368	Hypothetical protein (beta-lactamase)
XG2_1371	FmtA-like protein, putative (beta-lactamase family protein)

XG2_1451	Exonuclease RexA, putative
XG2_1476	Phosphoglycerol transferase and related proteins, alkaline phosphatase superfamily, putative
XG2_1741	ABC-type Fe3+-siderophores transporter, ATPase component, putative
XG2_1744	Iron compound ABC transporter, permease protein, putative
XG2_1792	Preprotein translocase subunit YidC, putative
XG2_1845	Serine/threonine protein kinase, putative
XG2_1984	Cell surface SD repeat antigen precursor, putative
XG2_2094	Spermidine synthase/hemolysin-type calcium-binding region
XG2_2128	Transporter, major facilitator family protein/ transport of a lantibiotic
XG2_2138	RNA-binding protein, Jag family, putative
XG2_2150	Transglycosylase associated protein/Predicted membrane protein, putative
XG2_2170	Peptidase, S54 (membrane-associated serine protease)
XG2_2205	Transcription antitermination factor NusG, putative
XG2_2234	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthases and related enzymes, phospholipase D family, putative
XG2_2320	Uncharacterized protein
XG2_2331	Protein involved in D-alanine esterification of lipoteichoic acid and wall teichoic acid (D-alanine transfer protein), putative
XG2_2333	Integral membrane protein, putative
XG2_2335	D-Ala-teichoic acid biosynthesis protein, putative
XG2_2346	3-oxoacyl-[acyl-carrier-protein] synthase III, putative
XG2_2364	Immunodominant staphylococcal antigen A precursor, putative
XG2_2386	Conserved hypothetical protein

Table 3. Biofilm mutants with COG function, operon, and paralogue.

Table 3.

SSA	COG Function	Operon/Synteny	Paralogues
SSA_0004	Predicted small periplasmic lipoprotein YifL (function Unknown	SSA_0004	
SSA_0036	Surface antigen	SSA_0036	
SSA_0048	DNA-binding transcriptional regulator, AcrR family	SSA_0048-0047	
SSA_0115	Ribosomal protein L29	SSA_0143-0104	
SSA_0144	DNA-binding transcriptional regulator, AcrR family	SSA_0144-0145	
SSA_0299	Predicted Fe2+/Mn2+ transporter, VIT1/CCC1 family	SSA_0311-0297	
SSA_0305	Uncharacterized membrane protein YgaE, UPF0421/DUF939 Family	SSA_0311-0297	
SSA_0351	Signal peptidase I	SSA_0350-0352	
SSA_0363	Na+/alanine symporter	SSA_0363-0359	
SSA_0364	Na+/serine symporter	SSA_0364	
SSA_0387	DNA-binding transcriptional regulator, GntR family	SSA_0387-0385	
SSA_0460	DNA-binding transcriptional regulator, MarR family	SSA_0505-0456	
SSA_0613	Glucan-binding domain (YG repeat)	SSA_0613-0610	
SSA_0708	Uncharacterized protein YpbQ, isoprenylcysteine carboxyl Methy	SSA_0706-0708	
SSA_0758	N/A	SSA_0767-0746	
SSA_0767	Diacylglycerol kinase family enzyme	SSA_0767-0746	
SSA_0805	Uncharacterized surface anchored protein	SSA_0802-0812	
SSA_0816	Predicted transcriptional regulator	SSA_0815-0820	
SSA_0879	Septation ring formation regulator EzrA	SSA_0873-0783	
SSA_0960	Signal transduction histidine kinase	SSA_0923-0963	
SSA_1030	DNA-binding transcriptional regulator, AcrR family	SSA_1030	
SSA_1064	N/A	SSA_1064	
SSA_1068	DNA-binding transcriptional regulator, LysR family	SSA_1065-1122	
SSA_1103	N/A	SSA_1065-1122	
SSA_1171	Integrase	SSA_1171	
SSA_1219	Sortase (surface protein transpeptidase)	SSA_1218-1220	
SSA_1271	nanoRNase/pAp phosphatase, hydrolyzes c-di-AMP and oligoRNAs	SSA_1271-1272	
SSA_1301	N/A	SSA_1301-1318	
SSA_1309	Predicted RNA-binding protein YlqC, contains KH domain, UPF010	SSA_1301-1318	
SSA_1363	CubicO group peptidase, beta-lactamase class C family	SSA_1363-1366	1363 vs 1368= 20% coverage, E-53,ID 77%; 1363 vs 1371= 100% cov, E 0, ID 58%

SSA_1368	CubicO group peptidase, beta-lactamase class C family	SSA_1368-1372	1368 vs 1363: cov= 99%, e 52, 77% ID
SSA_1371	CubicO group peptidase, beta-lactamase class C family	SSA_1368-1372	1371 vs 1363=100% cov, E 0, ID 58%
SSA_1451	ATP-dependent exoDNAse (exonuclease V) beta subunit (contains	SSA_1402-1475	
SSA_1476	Phosphoglycerol transferase MdoB or a related enzyme of AlkP s	SSA_1476-1479	
SSA_1741	ABC-type cobalamin/Fe3+-siderophores transport system, ATPase	SSA_1741-1744	
SSA_1744	ABC-type Fe3+-siderophore transport system, permease Component	SSA_1744-1741	
SSA_1792	Membrane protein insertase Oxa1/YidC/SpoIIIJ, required for The	SSA_1792-1791	
SSA_1845	Serine/threonine protein kinase	SSA_1840-1852	
SSA_1984	N/A	SSA_1972-2019	
SSA_2094	N/A	SSA_2082-2121	
SSA_2128	MFS family permease	SSA_2122-2129	
SSA_2138	Predicted RNA-binding protein Jag, conains KH and R3H Domains	SSA_2136-2150	
SSA_2150	Uncharacterized membrane protein YeaQ/YmgE, transglycosylase-a	SSA_2150-2136	
SSA_2170	Membrane associated serine protease, rhomboid family	SSA_2170-2185	
SSA_2205	Transcription antitermination factor NusG	SSA_2205	
SSA_2234	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin Sy	SSA_2211-2246	
SSA_2320	N/A	SSA_2321-2296	
SSA_2331	Poly D-alanine transfer protein DltD, involved inesterificatio	SSA_2330-2335	Not a paralog
SSA_2333	D-alanyl-lipoteichoic acid acyltransferase DltB, MBOAT Superfamily	SSA_2330-2335	Not a paralog
SSA_2335	N/A	SSA_2330-2335	Not a paralog
SSA_2346	3-oxoacyl-[acyl-carrier-protein] synthase III	SSA_2346-2360	
SSA_2364	N/A	SSA_2364-2371	
SSA_2386	Uncharacterized membrane protein	SSA_2150-2136	

Table 4. Potential targets for SPase I.

Table 4.

Gene	Description	Molecular Weight (Da)	Paralogue
SSA_0012	Beta-lactamase class A, putative	48396.6418	No
SSA_0015	Membrane ATPase FtsH, degrades sigma32 (integral membrane cell-division Zn metallo-peptidase), putative	72536.7981	No
SSA_0017	Cell shape determining protein MreC, putative	29544.234	No
SSA_0019	Secreted antigen GbpB/SagA; peptidoglycan hydrolase; PcsB protein precursor, putative	42469.4498	Yes
SSA_0021	Hypothetical protein	19560.3138	No
SSA_0036	Secreted protein, possible function in cell-wall metabolism (amidase), putative	72153.7086	Yes
SSA_0094	Cell wall metabolism, LysM type protein, putative	38190.3647	Yes
SSA_0140	Copper-translocating P-type ATPase, putative	79654.5004	Yes
SSA_0146	DNA repair ATPase, putative	84740.3229	Yes
SSA_0157	Hypothetical protein	9830.5367	No
SSA_0165	Conserved hypothetical protein	22562.2951	No
SSA_0167	Hypothetical protein (Asparagine/proline-rich)	33290.0671	No
SSA_0173	23S rRNA m1G745 methyltransferase, putative	32071.4906	No
SSA_0175	Penicillin-binding protein 1B, putative	80849.4257	Yes
SSA_0181	Glycosyltransferase (vectorial glycosyl polymerization (VGP) family), putative	49942.7224	No
SSA_0186	Competence protein ComYC, putative	11434.1479	No
SSA_0187	Competence protein ComYD, putative	15922.1403	No
SSA_0210	Conserved hypothetical protein	10675.0381	No
SSA_0215	Periplasmic sugar-binding protein (ribose porter), putative	36734.0881	Yes
SSA_0218	Sugar-binding periplasmic protein, putative	48332.3799	Yes

SSA_0227	Collagen-binding surface protein, putative	66956.5729	Yes
SSA_0243	Cyclo-nucleotide phosphodiesterase, putative	86679.2483	Yes
SSA_0257	N-acetylmuramidase/lysin, putative	104901.3981	Yes
SSA_0273	Hypothetical protein	50478.7114	No
SSA_0291	Oxidoreductase, putative	29913.2519	Yes
SSA_0301	Conserved hypothetical protein	21121.8259	Yes
SSA_0303	Surface protein C	162864.2939	Yes
SSA_0304	Bacterial cell wall degradation (CHAP/LysM domains), putative	23670.7359	Yes
SSA_0396	Conserved hypothetical protein	35467.6105	No
SSA_0398	Conserved hypothetical protein, beta-lactamase family	67219.477	Yes
SSA_0400	Conserved hypothetical protein, beta-lactamase family	67173.1453	Yes
SSA_0424	Exopolysaccharide biosynthesis protein, putative	32497.4588	No
SSA_0453	Type II secretory pathway, pullulanase PulA glycosidase, putative	136563.4945	Yes
SSA_0477	Cobalamin biosynthesis protein CbiM (ABC-type cobalt transporter), putative	27984.33	No
SSA_0478	Cobalt transport protein cbiN, putative	11108.6369	No
SSA_0498	ABC-type dipeptide/oligopeptide/nickel transport systems, permease components, putative	29272.4342	Yes
SSA_0500	Peptide ABC transporter, permease protein, putative	35018.2904	Yes
SSA_0521	Ethanolamine utilization protein EutL, putative	22395.7496	Yes
SSA_0565	Conserved hypothetical protein	92304.059	No
SSA_0607	ABC transporter, permease component, putative	87572.3394	No
SSA_0610	LemA-like protein, putative	20939.6228	No
SSA_0613	Glucosyltransferase, putative	175399.2	Yes

SSA_0684	Fibril-like structure subunit FibA, putative	139324.4303	Yes
SSA_0689	Penicillin-binding protein 2B, putative	75285.4986	Yes
SSA_0723	Hypothetical protein	5304.2581	No
SSA_0726	FmtA-like protein, putative	67437.3138	Yes
SSA_0747	DD-carboxypeptidase, putative	50393.737	Yes
SSA_0793	DNA-entry nuclease, putative	31237.3587	No
SSA_0801	Mur ligase family protein, putative	49483.2476	No
SSA_0803	Conserved hypothetical protein	27132.1818	No
SSA_0805	Collagen-binding surface protein, putative	60556.5712	Yes
SSA_0871	Cell division protein FtsX, putative	34500.8682	No
SSA_0881	Lipoprotein, putative	25788.1976	Yes
SSA_0897	Two component system histidine kinase, putative	52198.3589	Yes
SSA_0904	CshA-like fibrillar surface protein A	316344.8238	Yes
SSA_0905	CshA-like fibrillar surface protein B	207893.0251	Yes
SSA_0906	CshA-like fibrillar surface protein C	283050.8434	Yes
SSA_0908	ABC-type uncharacterized transport system, periplasmic component, putative	35242.882	No
SSA_0947	Hypothetical protein	22254.0843	Yes
SSA_0956	Surface protein D	153644.0111	Yes
SSA_0963	Peptidoglycan N-acetylglucosamine deacetylase A, putative	52559.0496	No
SSA_0967	Conserved hypothetical protein	24999.4402	No
SSA_0970	Conserved hypothetical protein	17131.3139	Yes
SSA_0991	Deoxyribonuclease, putative	28746.0106	No
SSA_1018	Zinc metalloprotease zmpC precursor, putative	339189.1463	Yes
SSA_1019	Collagen-binding surface protein, putative	87795.544	Yes
SSA_1023	Von Willebrand factor-binding protein precursor, putative	100844.6464	No

SSA_1042	Xylanase/chitin deacetylase, putative	35656.9693	No
SSA_1051	Spermidine/putrescine ABC transporter, spermidine/putrescine-binding protein, putative	40929.175	No
SSA_1052	Hypothetical protein	14191.2411	No
SSA_1063	Peptidoglycan-binding domain-containing protein, putative	48445.9462	No
SSA_1064	Conserved hypothetical protein (contains glucan-binding domain)	28533.3142	Yes
SSA_1065	Beta-hexosamidase A, putative	99533.6984	No
SSA_1095	Peptidoglycan hydrolase, putative	25795.2184	No
SSA_1106	IgA-specific metalloendopeptidase	208437.007	Yes
SSA_1112	Cell wall surface anchor family protein, putative	56954.5939	No
SSA_1118	Peptide methionine sulfoxide reductase, putative	41843.5912	Yes
SSA_1130	Iron-dependent peroxidase, putative	45410.802	No
SSA_1132	TatC, sec-independent protein translocase, putative	28740.0497	No
SSA_1148	Beta-glucosides PTS, EIIC, putative	47963.4949	Yes
SSA_1158	Conserved hypothetical protein	31692.5521	Yes
SSA_1161	Conserved hypothetical protein, possibly secreted	35214.2297	No
SSA_1219	Sortase, putative	27918.7002	Yes
SSA_1221	L-lactate dehydrogenase, putative	35328.6664	No
SSA_1234	5'-nucleotidase, putative	76340.7606	Yes
SSA_1274	Hypothetical protein	74589.9418	No
SSA_1301	Conserved uncharacterized protein, possible surface protein	90577.9472	Yes
SSA_1339	Pneumococcal histidine triad protein D precursor, putative	131547.081	Yes
SSA_1359	Arginine/histidine ABC transporter, permease component, putative	78675.5391	Yes
SSA_1363	FmtA-like protein, putative	66457.5498	Yes
SSA_1365	FmtA-like protein, putative	35566.1972	Yes
SSA_1368	Hypothetical protein	12234.0033	Yes
SSA_1369	FmtA-like protein, putative	67107.2338	Yes
SSA_1371	FmtA-like protein, putative	66339.2022	Yes

SSA_1372	Hypothetical protein	6968.0981	No
SSA_1390	Hypothetical protein	18387.7033	No
SSA_1391	Hypothetical protein	24811.6397	No
SSA_1408	Conserved hypothetical protein	21550.7326	No
SSA_1415	Oxidoreductase, putative	39914.9737	No
SSA_1434	Conserved uncharacterized Firmicutes protein	5894.894	No
SSA_1481	FmtA-like protein, putative	67460.5781	Yes
SSA_1489	Hypothetical protein	39767.8167	Yes
SSA_1525	Lyzozyme M1 (1,4-beta-N-acetylmuramidase), putative	32425.9415	No
SSA_1532	Membrane-fusion protein / periplasmic component of efflux system, putative	40852.7964	Yes
SSA_1544	Conserved uncharacterized protein	17995.4248	No
SSA_1567	Polar amino acid ABC transporter, amino acid-binding protein, putative	29898.7625	Yes
SSA_1588	Conserved ABC-type antimicrobial permease-like protein, putative	101024.5143	No
SSA_1591	Dipeptidase, putative	75007.1993	Yes
SSA_1593	Dipeptidase, putative	61779.1816	Yes
SSA_1594	Metalloendopeptidase, putative	79277.3956	Yes
SSA_1596	Hypothetical protein	36580.2214	Yes
SSA_1597	Hypothetical protein	37837.2035	Yes
SSA_1598	Hypothetical protein	36877.3853	Yes
SSA_1599	Hypothetical protein	38230.0062	Yes
SSA_1626	DNA translocase ftsK, putative	84255.6297	Yes
SSA_1631	Sortase-like protein, putative	33177.1937	Yes
SSA_1632	Surface protein, putative	51822.5553	Yes
SSA_1633	FimA fimbrial subunit-like protein, putative	51569.2038	Yes
SSA_1634	Heme utilization/adhesion exoprotein, putative	53218.4441	Yes
SSA_1635	Hypothetical protein	77849.1805	Yes
SSA_1649	Conserved hypothetical transmembrane protein	37430.4913	Yes
SSA_1650	3-Ketoacyl-ACP reductase, putative	25126.5123	Yes
SSA_1653	Hypothetical protein	37671.6974	Yes

SSA_1663	Collagen-binding protein A	163017.8908	Yes
SSA_1671	Conserved hypothetical protein	35911.9174	No
SSA_1673	Hypothetical protein	6245.3806	No
SSA_1680	ABC-type bacitracin resistance protein A, permease component, putative	74398.8545	Yes
SSA_1692	Phospho-B-galactosidase LacG, putative	54094.9495	Yes
SSA_1744	Iron compound ABC transporter, permease protein, putative	36333.7876	Yes
SSA_1750	Extracellular nuclease, putative	80467.6835	No
SSA_1793	Histidine kinase (sensor protein), putative	47218.5222	Yes
SSA_1871	Penicillin-binding protein 2X, putative	84338.3141	Yes
SSA_1882	Subtilisin-like serine proteases, putative	162877.6003	Yes
SSA_1909	Transcriptional attenuator LytR, putative	44849.8077	Yes
SSA_1951	Penicillin-binding protein 3, putative	45867.2977	Yes
SSA_1960	Conserved hypothetical protein	70290.9424	No
SSA_1961	Amino acid ABC transporter, amino acid-binding protein/permease protein, putative	57437.3438	Yes
SSA_1984	Cell surface SD repeat antigen precursor, putative	99884.0631	Yes
SSA_1985	Conserved hypothetical protein	72476.2762	Yes
SSA_1991	Pneumococcal histidine triad protein A, putative	90478.1281	Yes
SSA_2004	Zinc metalloprotease zmpB precursor, putative	209552.9886	Yes
SSA_2014	D-alanyl-D-alanine carboxypeptidase, putative	28870.2489	Yes
SSA_2020	Conserved hypothetical protein	114163.2144	No
SSA_2023	Fructan beta-fructosidase precursor, putative	155910.6148	Yes
SSA_2056	Cinnamoyl ester hydrolase, putative	34354.6188	Yes
SSA_2060	Arabinose efflux permease, putative	42316.0242	No
SSA_2074	Preprotein translocase subunit YajC, putative	12123.8761	No

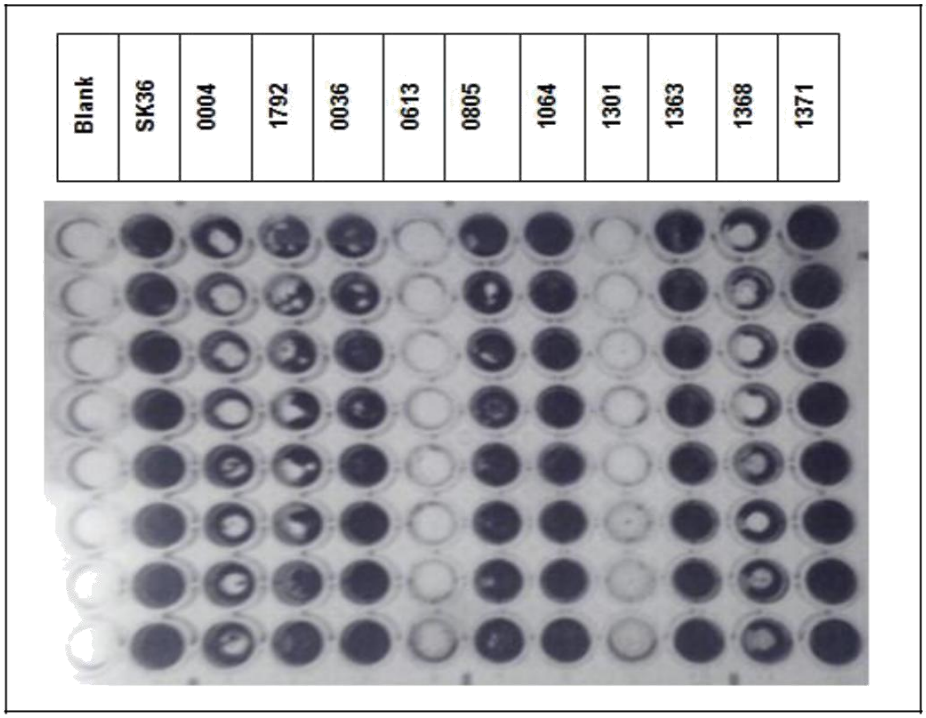
SSA_2101	Amino acid ABC transporter, periplasmic amino acid-binding protein, putative	32773.4118	Yes
SSA_2103	Hypothetical protein	7403.5321	No
SSA_2121	Cell wall surface anchor family protein, putative	171211.2214	Yes
SSA_2169	Glucose-1-phosphate uridylyltransferase, putative	35605.5304	Yes
SSA_2248	Conserved hypothetical protein	18481.9426	No
SSA_2250	ABC-type antimicrobial peptide transporter, permease component, putative	74446.4569	No
SSA_2264	Conserved hypothetical protein	18976.0874	No
SSA_2269	Conserved hypothetical protein	34777.2793	Yes
SSA_2281	Conserved hypothetical protein	17118.3266	No
SSA_2282	Phage infection protein, putative	107220.5308	Yes
SSA_2301	S-layer protein/ peptidoglycan endo-beta-N-acetylglucosaminidase, putative	21506.3392	No
SSA_2307	Hypothetical protein	56585.6564	No
SSA_2313	Hypothetical protein	16101.3788	Yes
SSA_2320	Hypothetical protein	125568.1653	No
SSA_2321	Cation (Co/Zn/Cd) efflux protein, putative	32405.2702	Yes
SSA_2338	Conserved uncharacterized protein	35945.4865	Yes
SSA_2340	Conserved hypothetical protein	43813.4783	No
SSA_2364	Immunodominant staphylococcal antigen A precursor, putative	19901.7084	No
SSA_2381	DegP protein, putative	40827.0835	No

Table 5. Potential targets for SPase I that affect biofilm formation

Table 5.

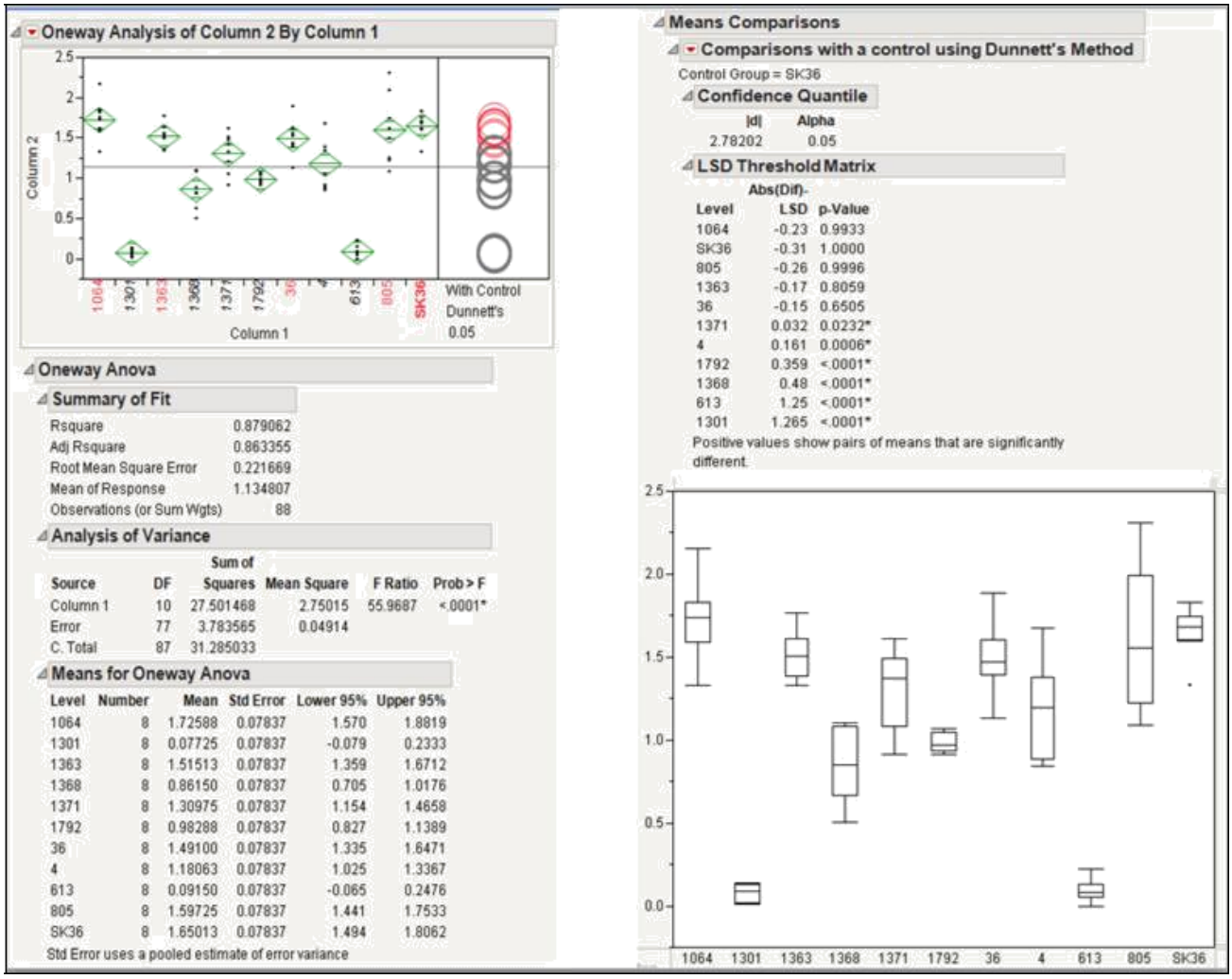
Gene ID	Gene Annotation
XG2_0036	Secreted protein, possible function in cell-wall metabolism (amidase), putative (N-acetylmuramoyl-L-alanine amidase)
XG2_0613	Glucosyltransferase, putative
XG2_0805	Collagen-binding surface protein, putative
XG2_1064	Hypothetical protein (dextransucrase/glucansucrase, N-acetylmuramoyl-L-alanine amidase, family 2)
XG2_1219	Sortase, putative
XG2_1301	Conserved uncharacterized protein, possible surface protein
XG2_1363	FmtA-like protein, putative
XG2_1368	Hypothetical protein (beta-lactamase)
XG2_1371	FmtA-like protein, putative (beta-lactamase family protein)
XG2_1744	Iron compound ABC transporter, permease protein, putative
XG2_1984	Cell surface SD repeat antigen precursor, putative
XG2_2320	Uncharacterized protein
XG2_2364	Immunodominant staphylococcal antigen A precursor, putative

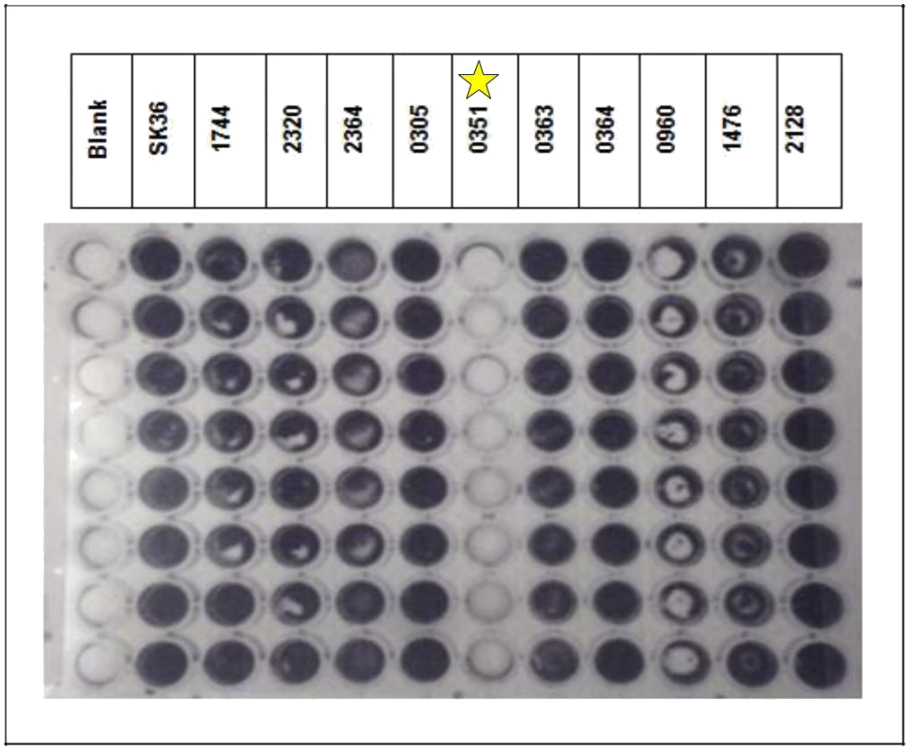
Figure 2. Crystal violet assay of 51 biofilm mutants. Yellow star signifies selected mutant, XG2_0351. Panel 1 of each plate contains the blank. Panel 2 of each plate contains SK36. Panels 3-12 of each plate contain the mutants.



	Blank	SK36	0004	1792	0036	0613	0805	1064	1301	1363	1368	1371
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.046	0.896	0.78	0.8	0.594	0.971	0.858	0.946	0.749	0.57	0.796	0.888
B	0.045	0.868	0.869	0.735	0.838	0.809	0.738	0.783	0.791	0.621	0.745	1.027
C	0.048	1.018	0.852	0.79	0.708	0.859	0.926	0.95	0.808	0.73	0.748	1.123
D	0.048	0.998	0.857	0.876	0.836	0.93	1.026	1.171	0.866	0.831	0.965	1.495
E	0.045	0.989	1.003	0.99	0.859	0.944	1.022	0.996	0.873	0.883	0.835	1.615
F	0.045	1.054	0.973	1.012	1.048	1.006	1.034	1.079	0.94	0.954	1.039	1.543
G	0.047	1.02	1	1.01	1.035	1.088	1.05	1.225	0.884	1.053	1.075	1.231
H	0.097	1.216	0.985	0.899	0.96	0.89	0.719	0.842	0.832	0.853	0.844	0.811

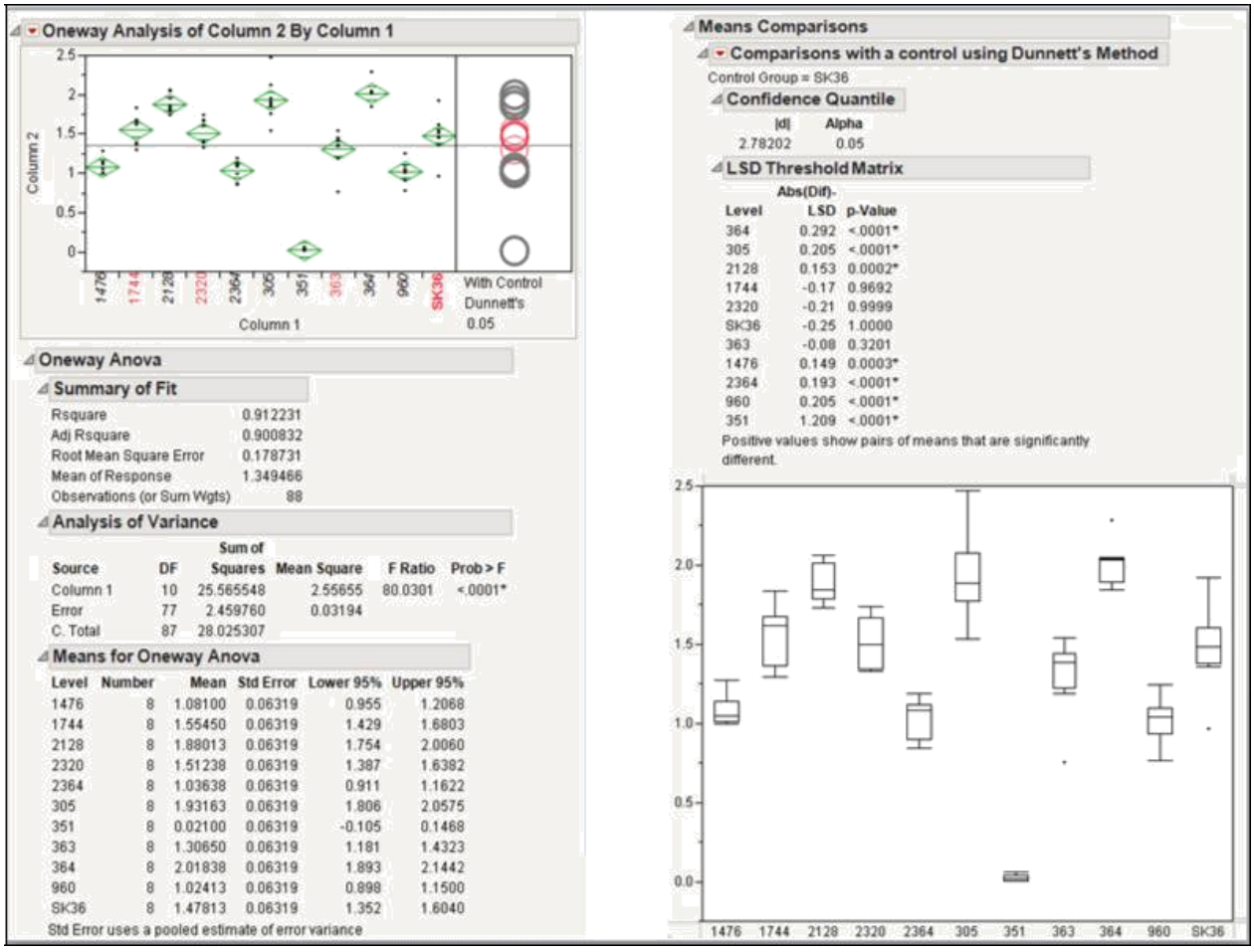
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.083	1.807	1.472	1.005	1.978	0.09	1.181	1.845	0.171	1.611	0.974	1.709
B	0.103	1.694	0.974	1.049	1.225	0.187	1.705	1.679	0.103	1.731	0.894	1.293
C	0.085	1.849	1.003	1.025	1.649	0.163	1.83	1.822	0.107	1.578	0.6	1.138
D	0.089	1.78	0.939	1.146	1.504	0.18	1.59	1.93	0.21	1.585	0.718	1.596
E	0.098	1.707	1.769	1.032	1.611	0.235	1.338	1.422	0.128	1.426	0.919	1.509
F	0.104	1.772	1.428	1.124	1.522	0.141	1.311	1.707	0.226	1.444	1.196	1.565
G	0.105	1.926	1.466	1.076	1.48	0.178	2.176	1.907	0.198	1.643	1.173	1.417
H	0.089	1.422	1.15	1.162	1.715	0.314	2.403	2.251	0.231	1.859	1.174	1.007

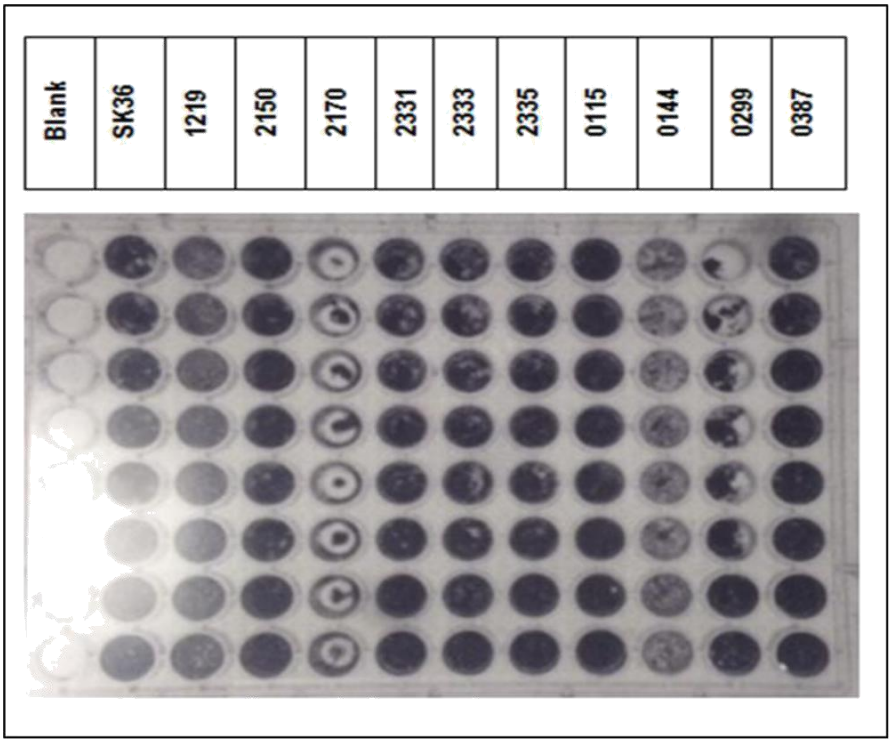

Oneway Analysis of Column 2 By Column 1
Column 2
Column 1
With Control
Dunnett's
0.05
Oneway Anova
Summary of Fit
Rsquare 0.879062
Adj Rsquare 0.863355
Root Mean Square Error 0.221669
Mean of Response 1.134807
Observations (or Sum Wgts) 88
Analysis of Variance
Source DF Sum of Squares Mean Square F Ratio Prob > F
Column 1 10 27.501468 2.75015 55.9687 <.0001*
Error 77 3.783565 0.04914
C. Total 87 31.285033
Means for Oneway Anova
Level Number Mean Std Error Lower 95% Upper 95%
1064 8 1.72588 0.07837 1.570 1.8819
1301 8 0.07725 0.07837 -0.079 0.2333
1363 8 1.51513 0.07837 1.359 1.6712
1368 8 0.86150 0.07837 0.705 1.0176
1371 8 1.30975 0.07837 1.154 1.4658
1792 8 0.98288 0.07837 0.827 1.1389
36 8 1.49100 0.07837 1.335 1.6471
4 8 1.18063 0.07837 1.025 1.3367
613 8 0.09150 0.07837 -0.065 0.2476
805 8 1.59725 0.07837 1.441 1.7533
SK36 8 1.65013 0.07837 1.494 1.8062
Std Error uses a pooled estimate of error variance
Means Comparisons
Comparisons with a control using Dunnett's Method
Control Group = SK36
Confidence Quantile
|d| Alpha
2.78202 0.05
LSD Threshold Matrix
Abs(Dif)-
Level LSD p-Value
1064 -0.23 0.9933
SK36 -0.31 1.0000
805 -0.26 0.9996
1363 -0.17 0.8059
36 -0.15 0.6505
1371 0.032 0.0232*
4 0.161 0.0006*
1792 0.359 <.0001*
1368 0.48 <.0001*
613 1.25 <.0001*
1301 1.265 <.0001*
Positive values show pairs of means that are significantly different.
1064 1301 1363 1368 1371 1792 36 4 613 805 SK36



	Blank	SK36	1744	2320	2364	0305	0351	0363	0364	0960	1476	2128
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.047	0.902	0.914	1.041	0.396	1.038	0.056	0.904	0.932	0.852	0.869	0.831
B	0.048	0.885	0.625	0.955	0.422	0.91	0.129	0.92	1.074	0.879	1.006	1.1
C	0.049	0.905	0.894	1.065	0.356	0.95	0.348	0.859	1.06	1.002	0.995	1.215
D	0.049	1.016	1.326	1.129	0.389	0.915	0.621	0.871	1.067	1.162	1.019	1.178
E	0.047	1.312	1.395	1.308	0.5	1.165	0.362	0.799	1.148	1.104	0.99	1.25
F	0.049	1.168	1.09	1.149	0.444	1.189	0.27	0.761	1.19	1.141	0.911	1.072
G	0.051	1.034	0.924	1.026	0.399	0.873	0.358	0.782	1.089	0.945	0.993	0.969
H	0.052	1.164	1.48	0.92	0.398	0.98	0.458	0.708	0.993	0.874	0.82	0.822

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.078	1.603	1.479	1.488	0.941	2.57	0.111	1.514	2.136	0.864	1.242	1.961
B	0.105	1.717	1.93	1.423	1.199	1.627	0.1	1.282	1.943	1.342	1.14	2.132
C	0.074	1.056	1.455	1.476	1.174	2.211	0.102	1.447	2.135	1.122	1.21	2.04
D	0.08	1.456	1.389	1.835	1.22	1.849	0.096	1.543	1.945	1.144	1.146	1.825
E	0.125	1.64	1.751	1.694	0.96	2.006	0.11	1.516	2.141	1.192	1.105	1.917
F	0.121	1.561	1.713	1.768	1.286	2.048	0.157	1.636	2.376	1.192	1.093	2.155
G	0.088	1.538	1.774	1.424	1.09	1.964	0.135	1.429	2.11	1.098	1.106	1.875
H	0.094	2.019	1.71	1.756	1.186	1.943	0.122	0.85	2.126	1.004	1.371	1.901

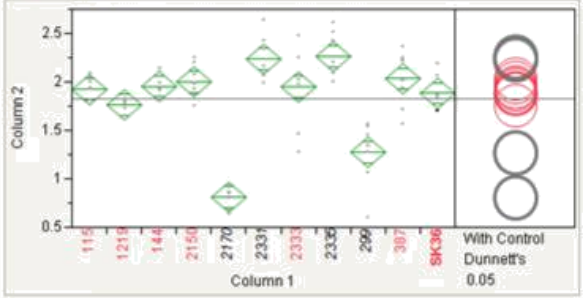

Oneway Analysis of Column 2 By Column 1
Column 2
Column 1
With Control
Dunnett's
0.05
Oneway Anova
Summary of Fit
Rsquare 0.912231
Adj Rsquare 0.900832
Root Mean Square Error 0.178731
Mean of Response 1.349466
Observations (or Sum Wgts) 88
Analysis of Variance
Source DF Sum of Squares Mean Square F Ratio Prob > F
Column 1 10 25.565548 2.55655 80.0301 <.0001*
Error 77 2.459760 0.03194
C. Total 87 28.025307
Means for Oneway Anova
Level Number Mean Std Error Lower 95% Upper 95%
1476 8 1.08100 0.06319 0.955 1.2068
1744 8 1.55450 0.06319 1.429 1.6803
2128 8 1.88013 0.06319 1.754 2.0060
2320 8 1.51238 0.06319 1.387 1.6382
2364 8 1.03638 0.06319 0.911 1.1622
305 8 1.93163 0.06319 1.806 2.0575
351 8 0.02100 0.06319 -0.105 0.1468
363 8 1.30650 0.06319 1.181 1.4323
364 8 2.01838 0.06319 1.893 2.1442
960 8 1.02413 0.06319 0.898 1.1500
SK36 8 1.47813 0.06319 1.352 1.6040
Std Error uses a pooled estimate of error variance
Means Comparisons
Comparisons with a control using Dunnett's Method
Control Group = SK36
Confidence Quantile
|d| Alpha
2.78202 0.05
LSD Threshold Matrix
Abs(Dif)-
Level LSD p-Value
364 0.292 <.0001*
305 0.205 <.0001*
2128 0.153 0.0002*
1744 -0.17 0.9692
2320 -0.21 0.9999
SK36 -0.25 1.0000
363 -0.08 0.3201
1476 0.149 0.0003*
2364 0.193 <.0001*
960 0.205 <.0001*
351 1.209 <.0001*
Positive values show pairs of means that are significantly different.
1476 1744 2128 2320 2364 305 351 363 364 960 SK36



	Blank	SK36	1219	2150	2170	2331	2333	2335	0115	0144	0299	0387
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.046	0.929	0.361	0.753	1.074	1.368	0.545	1.638	0.657	0.41	0.599	0.686
B	0.045	1.708	0.424	0.983	0.768	0.981	0.64	1.15	0.933	0.351	0.489	1.188
C	0.047	1.906	0.427	1.183	1.227	1.564	1.093	0.89	0.775	0.347	1.785	1.276
D	0.046	1.849	0.481	1.185	1.628	1.66	0.837	0.805	0.651	0.381	1.222	0.887
E	0.045	1.192	0.438	0.959	1.482	0.674	1.179	1.396	0.672	0.626	1.434	1.05
F	0.05	1.726	0.504	0.873	1.591	1.05	1.574	0.934	0.938	0.601	1.747	1.121
G	0.049	1.301	0.473	1.25	1.467	0.947	1.068	0.985	0.916	0.529	0.869	0.925
H	0.048	0.993	0.454	0.853	1.238	0.978	0.961	0.898	1.01	0.376	0.629	0.846

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.077	1.78	1.802	2.003	0.714	2.168	1.363	2.23	2.018	2.07	0.682	1.64
B	0.072	1.854	1.861	2.06	0.946	2.074	2.149	2.076	2.172	2.008	1.363	1.794
C	0.082	1.916	1.881	2.28	0.936	2.249	1.542	2.592	2.011	2.009	1.462	2.297
D	0.088	1.95	1.797	2.074	1.017	2.149	2.147	2.236	1.891	1.996	1.417	2.307
E	0.087	1.849	1.737	1.831	0.76	2.245	2.053	2.164	1.839	1.932	1.149	2.021
F	0.088	1.946	1.897	1.939	0.892	2.425	2.101	2.276	2.128	1.883	1.645	2.092
G	0.081	2.148	1.912	2.329	0.992	2.509	2.331	2.501	1.886	2.184	1.62	2.33
H	0.08	2.275	1.889	2.158	0.884	2.72	2.563	2.698	2.087	2.222	1.527	2.448

Oneway Analysis of Column 2 By Column 1



Oneway Anova

Summary of Fit

Rsquare	0.796115
Adj Rsquare	0.769637
Root Mean Square Error	0.219812
Mean of Response	1.827216
Observations (or Sum Wgts)	88

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Column 1	10	14.527304	1.45273	30.0664	<.0001*
Error	77	3.720439	0.04832		
C. Total	87	18.247743			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
115	8	1.92213	0.07772	1.7674	2.0769
1219	8	1.76513	0.07772	1.6104	1.9199
144	8	1.95613	0.07772	1.8014	2.1109
2150	8	2.00238	0.07772	1.8476	2.1571
2170	8	0.81075	0.07772	0.6560	0.9655
2331	8	2.23550	0.07772	2.0807	2.3903
2333	8	1.94925	0.07772	1.7945	2.1040
2335	8	2.26475	0.07772	2.1100	2.4195
299	8	1.27625	0.07772	1.1215	1.4310
387	8	2.03425	0.07772	1.8795	2.1890
SK36	8	1.88288	0.07772	1.7281	2.0376

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons with a control using Dunnett's Method

Control Group = SK36

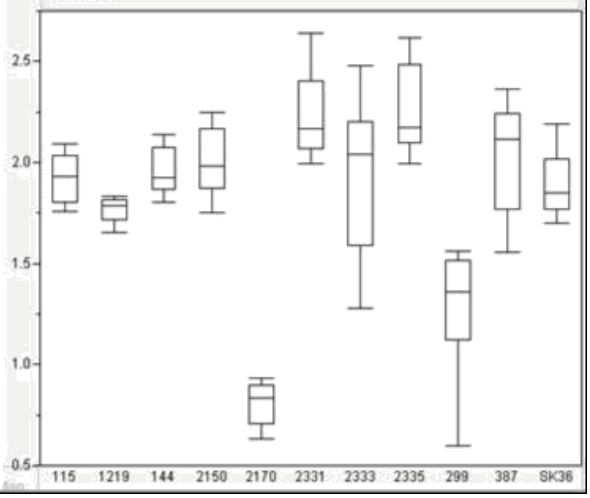
Confidence Quantile

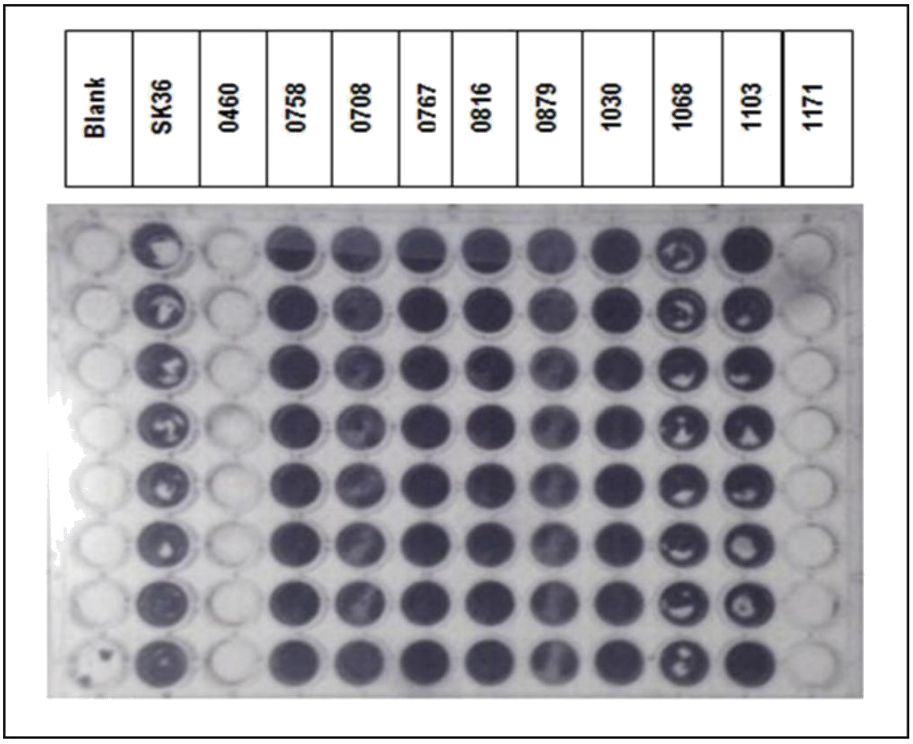
|d|	Alpha
2.78202	0.05

LSD Threshold Matrix

Level	Abs(Dif)-LSD	p-Value
2335	0.076	0.0071*
2331	0.047	0.0157*
387	-0.15	0.6940
2150	-0.19	0.8829
144	-0.23	0.9944
2333	-0.24	0.9974
115	-0.27	1.0000
SK36	-0.31	1.0000
1219	-0.19	0.8910
299	0.301	<.0001*
2170	0.766	<.0001*

Positive values show pairs of means that are significantly different.

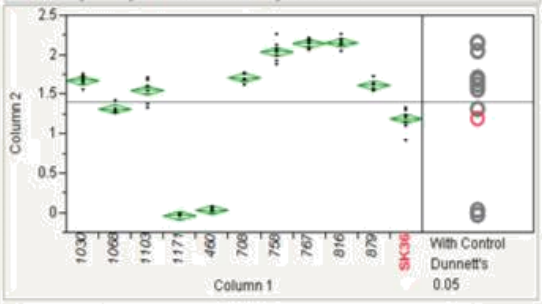




	Blank	SK36	0460	0758	0708	0767	0816	0879	1030	1068	1103	1171
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.045	0.943	0.051	1.062	0.507	1.042	1.083	0.543	0.924	0.952	0.957	0.046
B	0.045	0.989	0.055	0.989	0.501	1.012	1.132	0.678	1.03	1.027	1.015	0.046
C	0.048	0.946	0.053	0.879	0.468	0.928	0.851	0.592	0.89	0.982	1.009	0.047
D	0.049	1.128	0.057	0.912	0.552	0.986	0.886	0.491	0.904	0.905	0.837	0.047
E	0.062	0.897	0.056	0.93	0.437	0.945	0.78	0.494	1.084	0.982	1.044	0.046
F	0.046	0.868	0.054	0.857	0.381	0.957	0.874	0.481	1.006	1.067	0.855	0.046
G	0.837	1.012	0.053	0.784	0.388	0.861	0.804	0.458	0.918	0.954	0.864	0.047
H	0.576	1.009	0.059	1.025	0.554	1.023	0.936	0.38	0.897	0.891	0.889	0.052

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.083	1.033	0.133	1.816	2.385	2.176	2.257	1.837	1.785	1.414	1.83	0.104
B	0.073	1.223	0.145	1.891	2.17	2.285	2.248	1.755	1.839	1.425	1.702	0.083
C	0.075	1.352	0.148	1.861	2.043	2.291	2.253	1.659	1.837	1.387	1.665	0.085
D	0.071	1.331	0.19	1.795	2.149	2.296	2.285	1.653	1.681	1.388	1.727	0.082
E	0.082	1.257	0.161	1.74	2.109	2.211	2.315	1.684	1.877	1.526	1.72	0.072
F	0.086	1.418	0.169	1.813	2.192	2.261	2.159	1.732	1.765	1.43	1.488	0.075
G	0.176	1.444	0.161	1.868	1.995	2.282	2.308	1.837	1.764	1.532	1.446	0.079
H	0.342	1.442	0.133	1.879	2.239	2.333	2.38	1.739	1.813	1.384	1.793	0.08

Oneway Analysis of Column 2 By Column 1



Oneway Anova

Summary of Fit

Rsquare	0.988755
Adj Rsquare	0.987295
Root Mean Square Error	0.082512
Mean of Response	1.396943
Observations (or Sum Wgts)	88

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Column 1	10	46.096662	4.60967	677.0782	<.0001*
Error	77	0.524229	0.00681		
C. Total	87	46.620892			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1030	8	1.6716	0.02917	1.614	1.7297
1068	8	1.3123	0.02917	1.254	1.3703
1103	8	1.5479	0.02917	1.490	1.6060
1171	8	-0.0410	0.02917	-0.099	0.0171
460	8	0.0315	0.02917	-0.027	0.0896
708	8	1.7094	0.02917	1.651	1.7675
758	8	2.0368	0.02917	1.979	2.0948
767	8	2.1434	0.02917	2.085	2.2015
816	8	2.1521	0.02917	2.094	2.2102
879	8	1.6135	0.02917	1.555	1.6716
SK36	8	1.1890	0.02917	1.131	1.2471

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons with a control using Dunnett's Method

Control Group = SK36

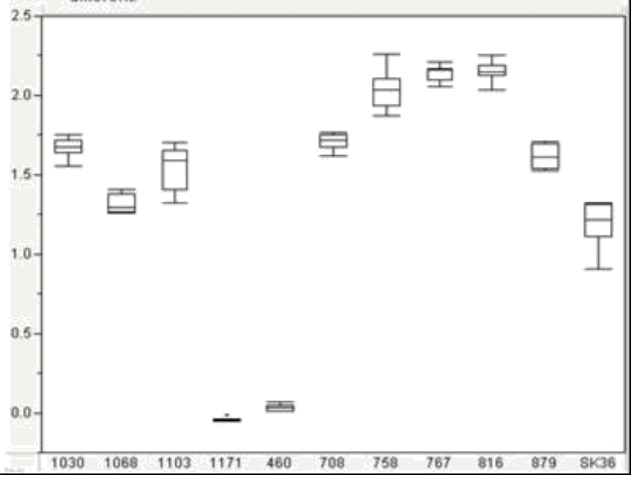
Confidence Quantile

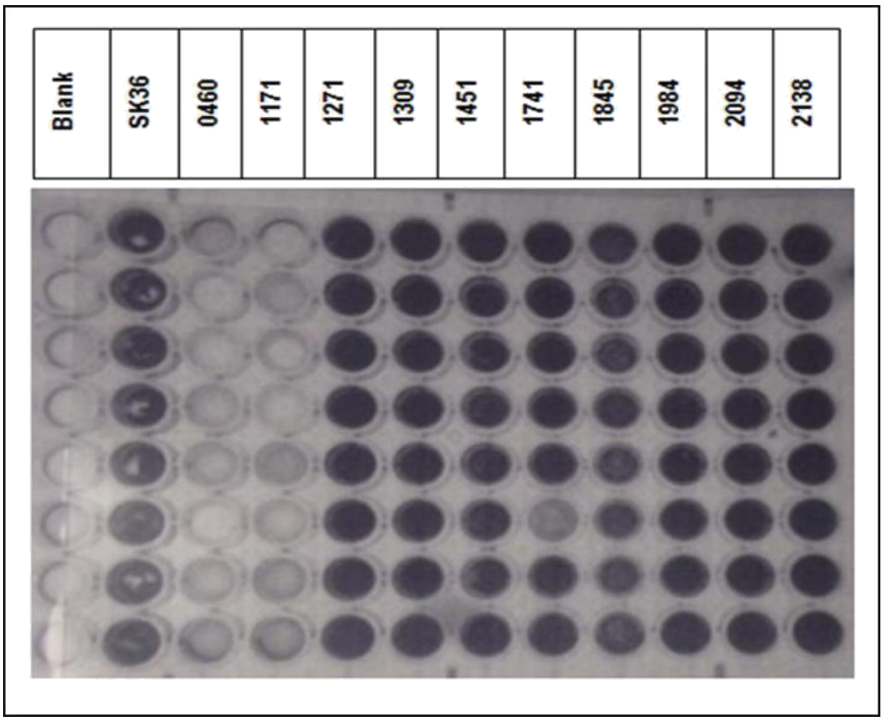
|d|	Alpha
2.78202	0.05

LSD Threshold Matrix

Level	Abs(Dif)-LSD	p-Value
816	0.848	<.0001*
767	0.84	<.0001*
758	0.733	<.0001*
708	0.406	<.0001*
1030	0.368	<.0001*
879	0.31	<.0001*
1103	0.244	<.0001*
1068	0.008	0.0292*
SK36	-0.11	1.0000
460	1.043	<.0001*
1171	1.115	<.0001*

Positive values show pairs of means that are significantly different.

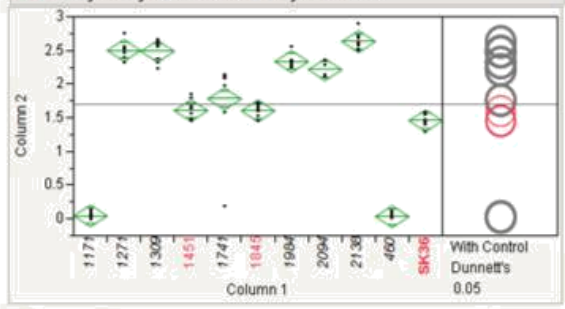




	Blank	SK36	0460	1171	1271	1309	1451	1741	1845	1984	2094	2138
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.047	1.089	0.047	0.048	1.205	1.181	1.122	1.123	0.49	1.185	1.026	1.097
B	0.046	1.109	0.046	0.049	1.242	1.153	1.421	1.109	0.553	1.085	1.089	1.084
C	0.049	1.159	0.046	0.046	1.072	1.122	1.121	1.367	0.544	1.119	0.993	1.152
D	0.049	0.975	0.046	0.047	1.079	1.11	1.07	1.009	0.609	1.054	1.017	1.111
E	0.046	1.013	0.051	0.047	1.13	1.13	0.962	1.051	0.481	1.143	1.032	1.033
F	0.047	1.072	0.048	0.051	1.228	1.149	0.935	1.066	0.604	1.049	0.961	1.022
G	0.048	0.988	0.05	0.05	1.189	1.208	0.976	1.092	0.519	1.121	1.055	0.998
H	0.056	1.01	0.054	0.051	1.183	1.129	1.019	1.016	0.536	1.073	1.018	1.027

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.074	1.566	0.134	0.085	2.405	2.736	1.884	2.064	1.55	2.417	2.437	2.719
B	0.083	1.488	0.088	0.139	2.559	2.697	1.581	2.177	1.735	2.438	2.439	2.588
C	0.106	1.534	0.094	0.127	2.652	2.463	1.544	2.195	1.736	2.348	2.222	2.615
D	0.09	1.637	0.13	0.091	2.629	2.314	1.724	2.193	1.82	2.375	2.197	2.684
E	0.087	1.631	0.139	0.149	2.489	2.415	1.567	2.228	1.711	2.364	2.2	2.657
F	0.132	1.535	0.086	0.124	2.603	2.646	1.755	0.285	1.682	2.454	2.194	2.803
G	0.101	1.674	0.17	0.197	2.615	2.663	1.651	2.217	1.609	2.419	2.378	2.779
H	0.107	1.374	0.216	0.232	2.84	2.75	1.937	1.671	1.773	2.651	2.442	2.995

Oneway Analysis of Column 2 By Column 1



Oneway Anova

Summary of Fit

Rsquare	0.943406
Adj Rsquare	0.936056
Root Mean Square Error	0.228775
Mean of Response	1.700455
Observations (or Sum Wgts)	88

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Column 1	10	67.179065	6.71791	128.3561	<.0001*
Error	77	4.030031	0.05234		
C. Total	87	71.209096			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1171	8	0.04550	0.08088	-0.116	0.2066
1271	8	2.50150	0.08088	2.340	2.6626
1309	8	2.48800	0.08088	2.327	2.6491
1451	8	1.60788	0.08088	1.447	1.7689
1741	8	1.78125	0.08088	1.620	1.9423
1845	8	1.60450	0.08088	1.443	1.7656
1984	8	2.33575	0.08088	2.175	2.4968
2094	8	2.21613	0.08088	2.055	2.3772
2138	8	2.63250	0.08088	2.471	2.7936
460	8	0.03463	0.08088	-0.126	0.1957
SK36	8	1.45738	0.08088	1.296	1.6184

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons with a control using Dunnett's Method

Control Group = SK36

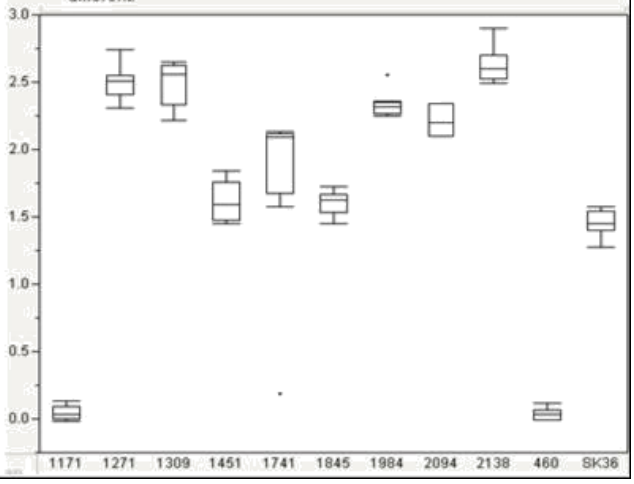
Confidence Quantile

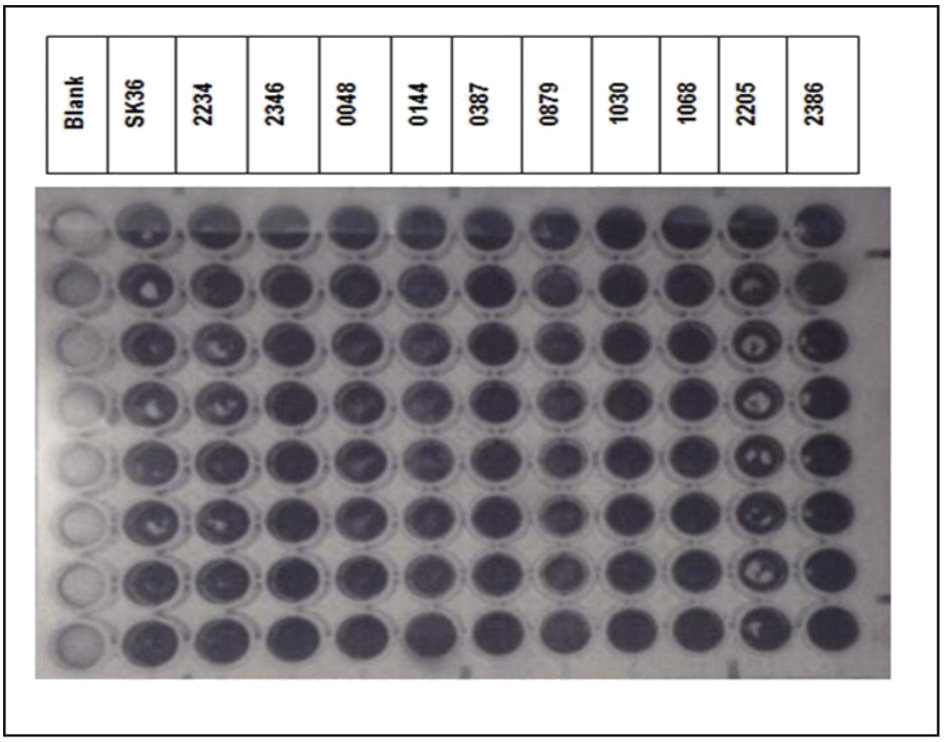
|d|	Alpha
2.78202	0.05

LSD Threshold Matrix

Level	Abs(Dif)-LSD	p-Value
2138	0.857	<.0001*
1271	0.726	<.0001*
1309	0.712	<.0001*
1984	0.56	<.0001*
2094	0.441	<.0001*
1741	0.006	0.0441*
1451	-0.17	0.7389
1845	-0.17	0.7598
SK36	-0.32	1.0000
1171	1.094	<.0001*
460	1.105	<.0001*

Positive values show pairs of means that are significantly different.

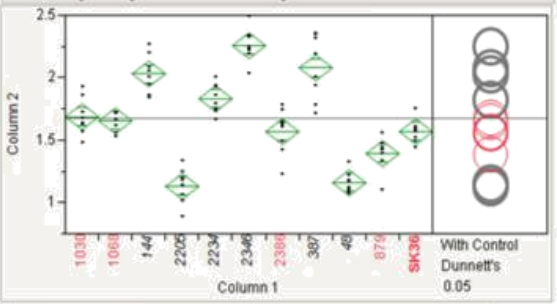




	Blank	SK36	2234	2346	0048	0144	0387	0879	1030	1068	2205	2386
	1	2	3	4	5	6	7	8	9	10	11	12
A	1.081	1.085	1.082	1.203	0.795	0.704	1.167	0.631	1.073	1.044	0.991	1.066
B	0.047	1.059	1.08	1.118	0.757	0.576	1.037	0.691	1.24	1.122	1.048	1.016
C	0.049	1.022	1.07	0.982	0.75	0.524	1.079	0.835	1.1	1.017	1.047	1.035
D	0.05	1.05	1.03	0.998	0.797	0.598	1.053	0.673	0.973	1.074	1.058	0.975
E	0.047	1.076	0.973	1.031	0.703	0.453	0.997	0.727	1.002	1.17	1.068	1.058
F	0.389	1.105	1.013	1.076	0.701	0.494	1.021	0.681	1.024	1.075	1.107	1.124
G	0.048	1.088	0.945	1.127	0.704	0.508	1.021	0.647	0.977	1.104	0.977	1.067
H	0.051	1.18	1.073	1.078	0.711	0.654	1.183	0.635	1.009	1.02	1.003	1.082

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.204	1.974	2.153	2.545	1.4	2.422	2.227	1.694	1.898	1.934	1.468	1.642
B	0.257	1.666	2.083	2.256	1.327	2.228	2.574	1.613	1.86	1.774	1.556	1.827
C	0.173	1.807	1.885	2.468	1.343	2.06	1.938	1.778	1.792	1.839	1.279	1.869
D	0.247	1.709	2.009	2.414	1.292	2.082	2.005	1.553	1.843	1.936	1.23	1.451
E	0.227	1.785	2.228	2.447	1.313	2.165	2.154	1.324	1.705	1.753	1.357	1.715
F	0.293	1.826	2.147	2.437	1.387	2.3	2.564	1.638	1.944	1.947	1.402	1.851
G	0.155	1.736	1.953	2.556	1.438	2.278	2.413	1.663	2.152	1.889	1.106	2.005
H	0.215	1.817	1.973	2.711	1.542	2.492	2.535	1.647	2.079	1.938	1.415	1.96

Oneway Analysis of Column 2 By Column 1



Oneway Anova

Summary of Fit

Rsquare	0.865951
Adj Rsquare	0.848542
Root Mean Square Error	0.145803
Mean of Response	1.66942
Observations (or Sum Wgts)	88

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Column 1	10	10.574345	1.05743	49.7418	<.0001*
Error	77	1.636903	0.02126		
C. Total	87	12.211248			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1030	8	1.68775	0.05155	1.5851	1.7904
1068	8	1.65488	0.05155	1.5522	1.7575
144	8	2.03200	0.05155	1.9294	2.1346
2205	8	1.13025	0.05155	1.0276	1.2329
2234	8	1.83250	0.05155	1.7299	1.9351
2346	8	2.25788	0.05155	2.1552	2.3605
2386	8	1.56863	0.05155	1.4660	1.6713
387	8	2.07988	0.05155	1.9772	2.1825
48	8	1.15888	0.05155	1.0562	1.2615
879	8	1.39238	0.05155	1.2897	1.4950
SK36	8	1.56863	0.05155	1.4660	1.6713

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons with a control using Dunnett's Method

Control Group = SK36

Confidence Quantile

|d|	Alpha
2.78202	0.05

LSD Threshold Matrix

Level	Abs(Dif)-LSD	p-Value
2346	0.486	<.0001*
387	0.308	<.0001*
144	0.261	<.0001*
2234	0.061	0.0045*
1030	-0.08	0.5038
1068	-0.12	0.8281
2386	-0.2	1.0000
SK36	-0.2	1.0000
879	-0.03	0.1192
48	0.207	<.0001*
2205	0.236	<.0001*

Positive values show pairs of means that are significantly

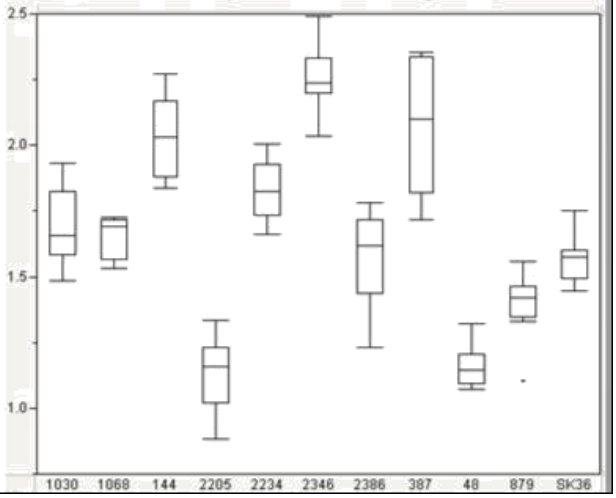
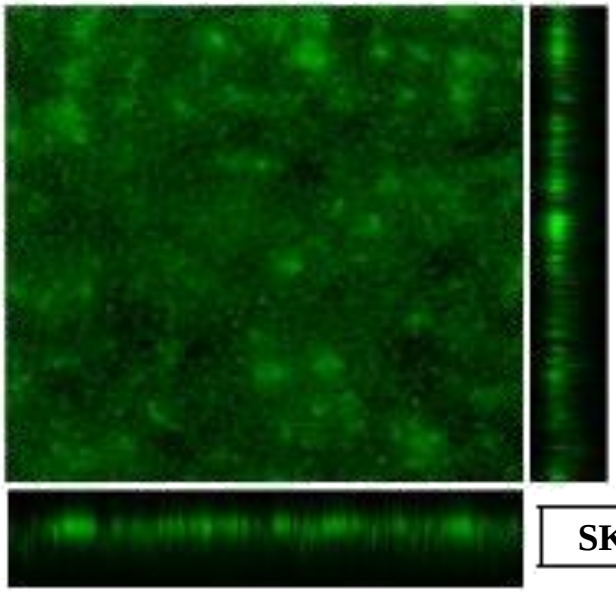
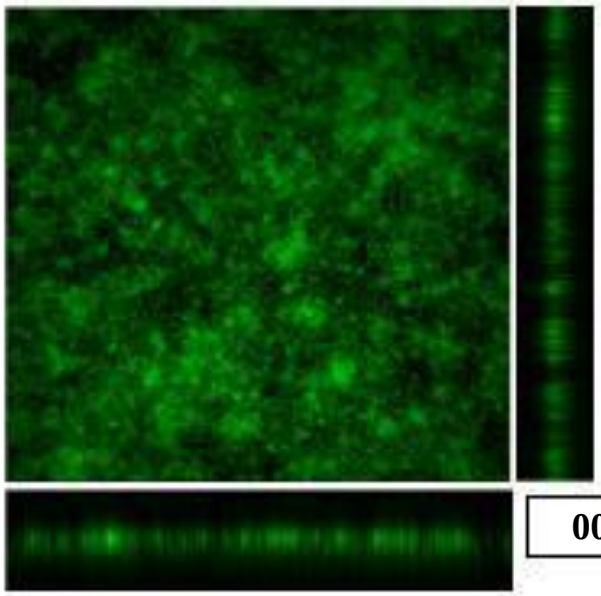


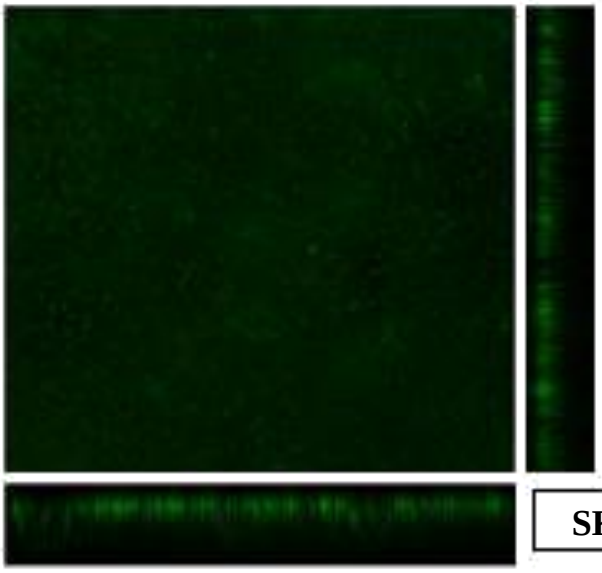
Figure 3. Biofilm imaging by confocal laser scanning microscopy (CLSM) of 24 mutants that have shown to affect biofilm formation through preliminary data. Each WT and mutant was repeated 3 times each (A-C).



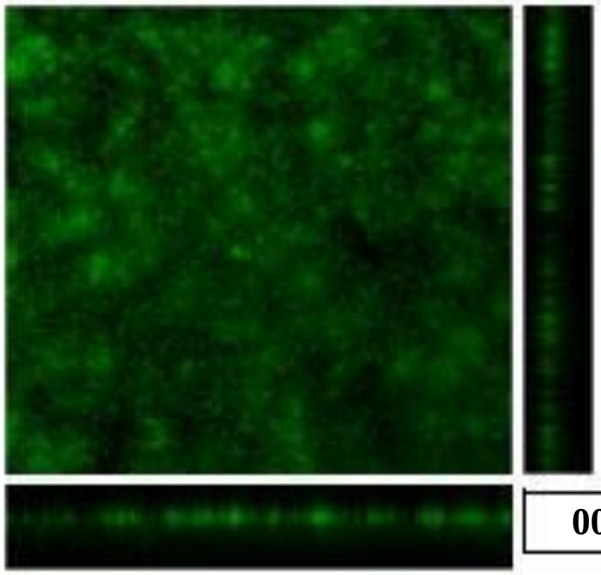
SK36-A



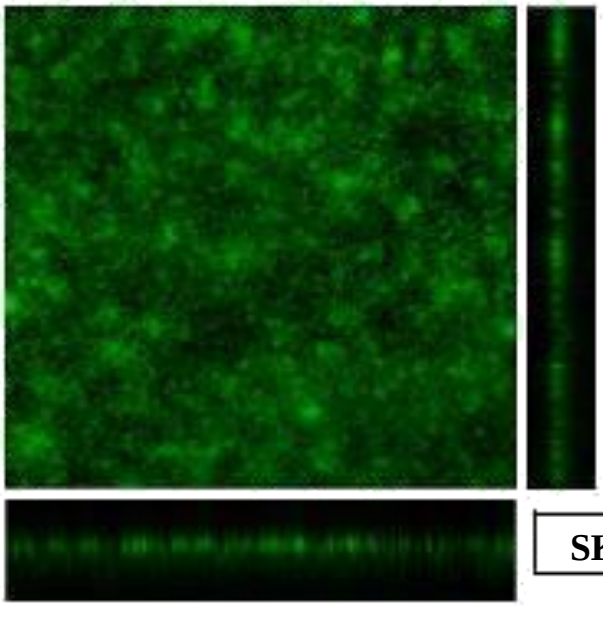
0036-A



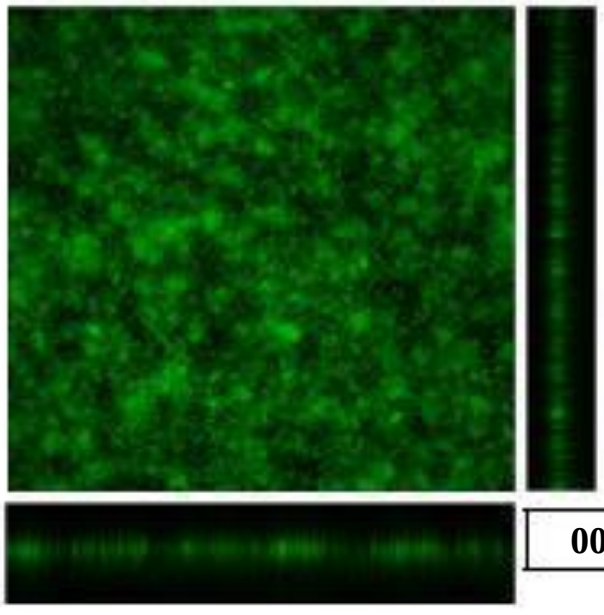
SK36-B



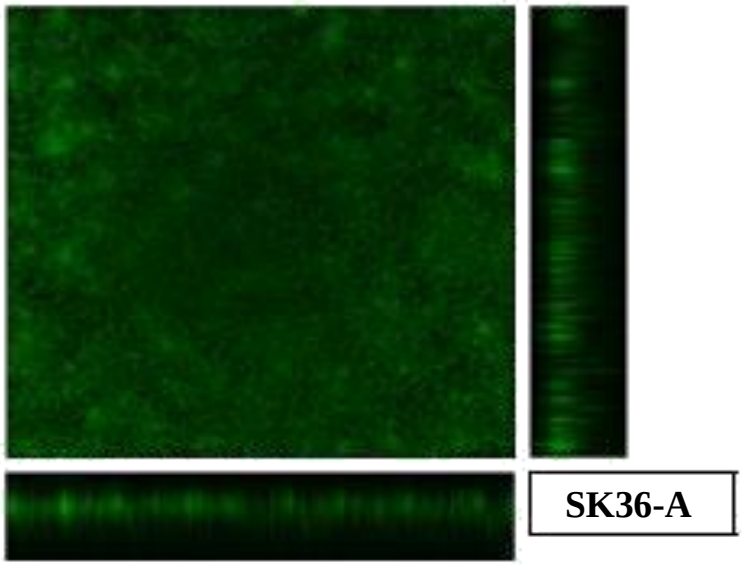
0036-B



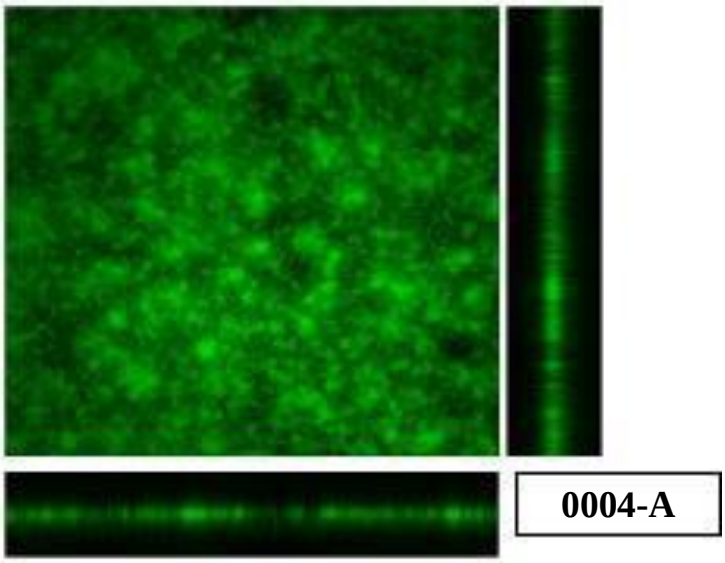
SK36-C



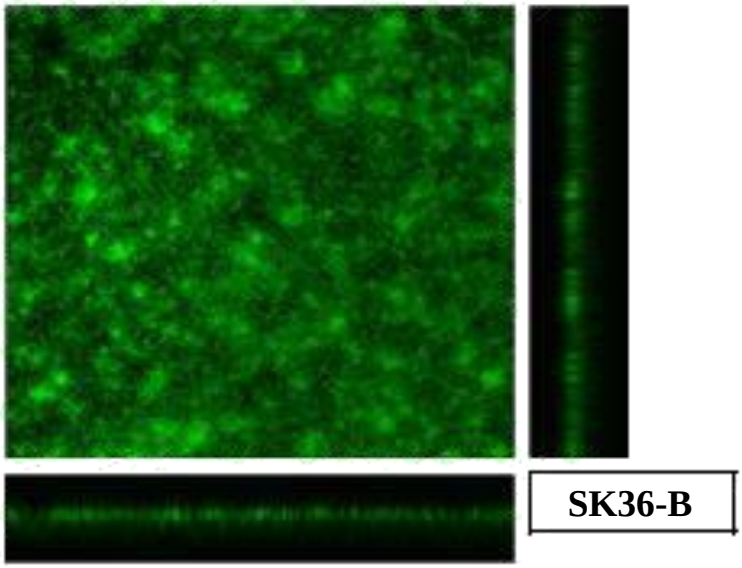
0036-C



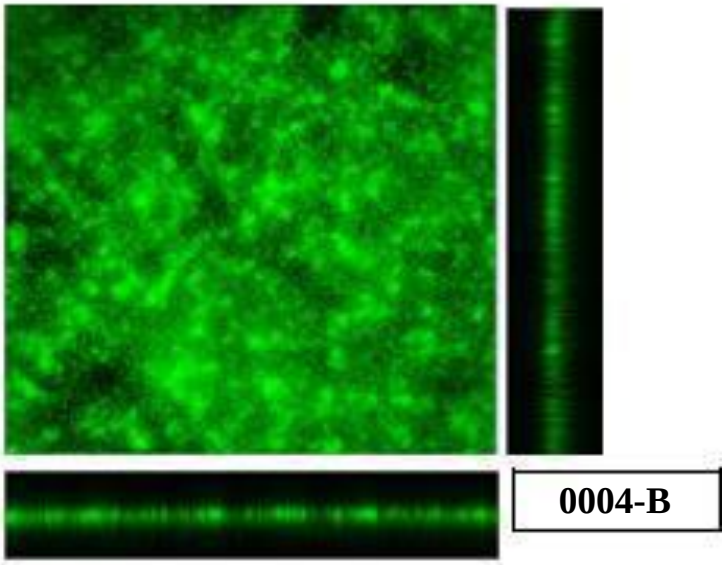
SK36-A



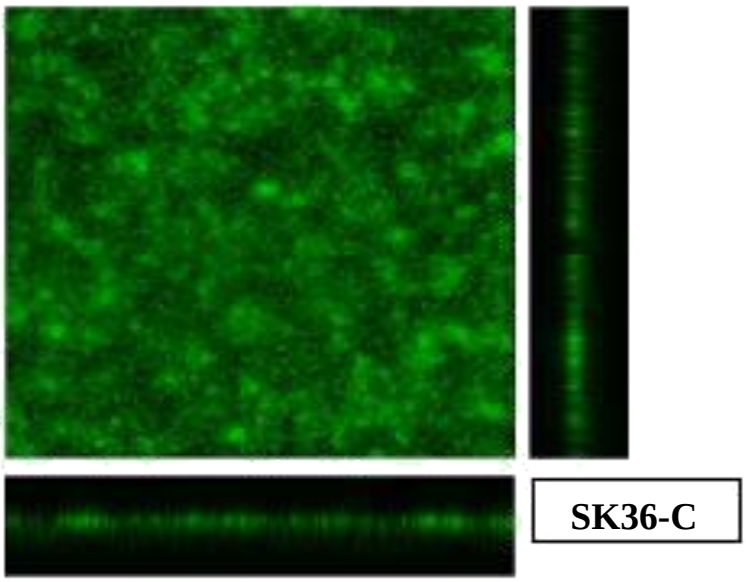
0004-A



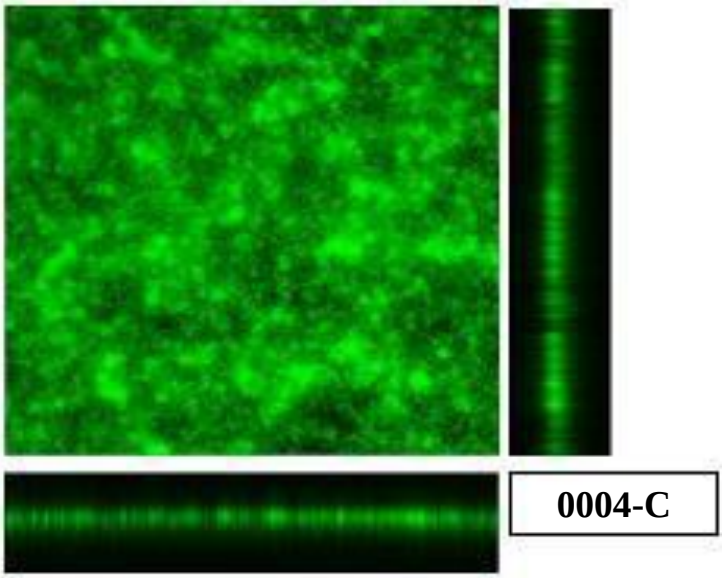
SK36-B



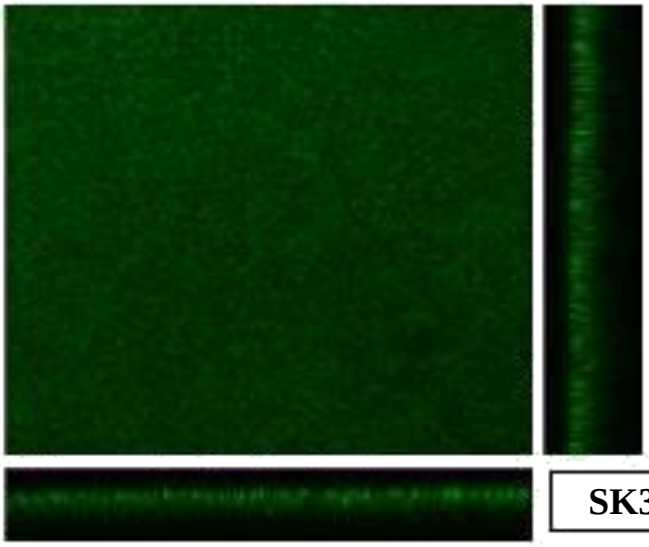
0004-B



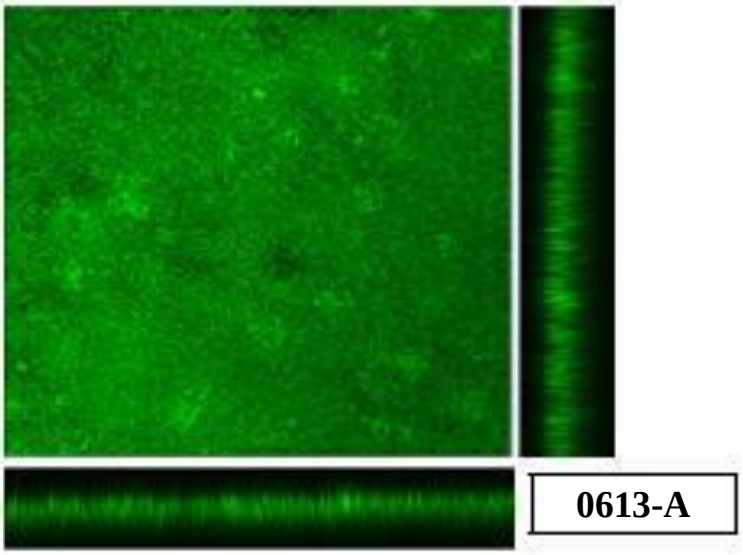
SK36-C



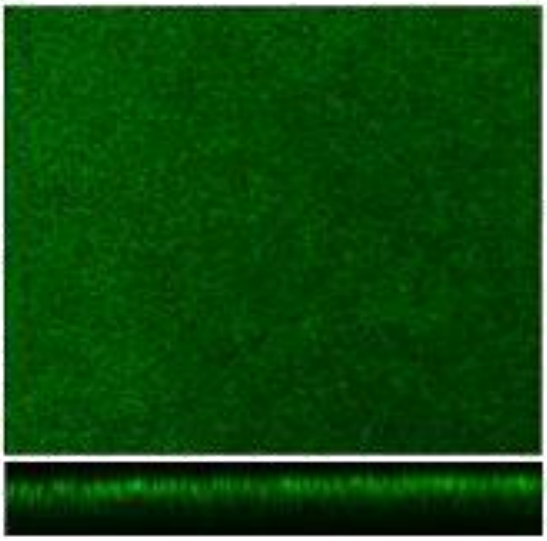
0004-C



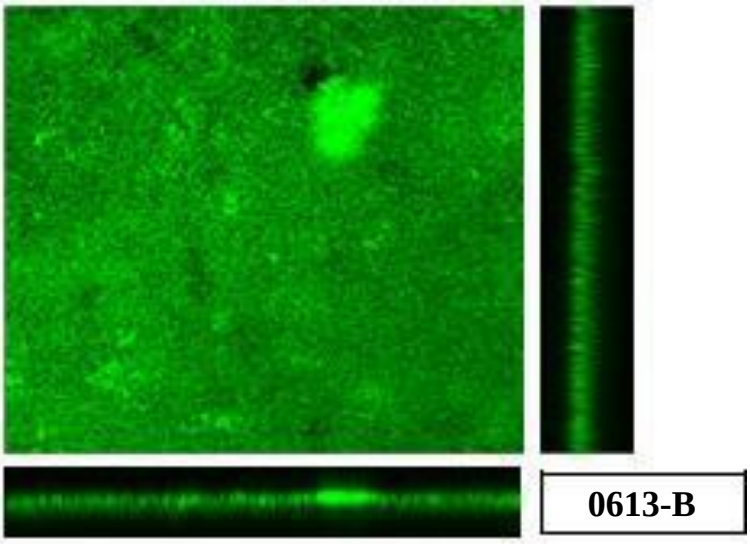
SK36-A



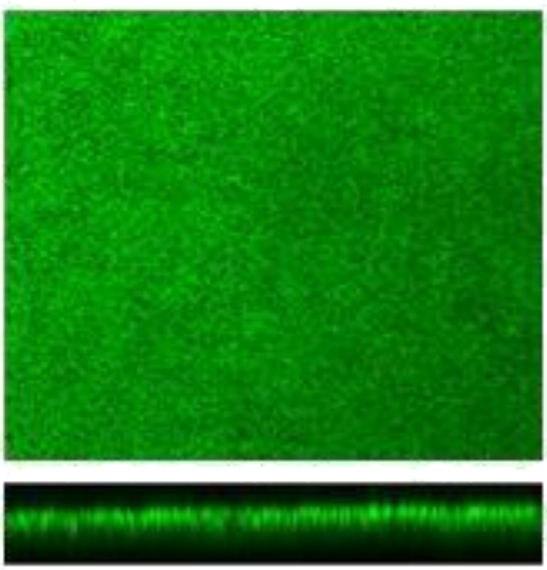
0613-A



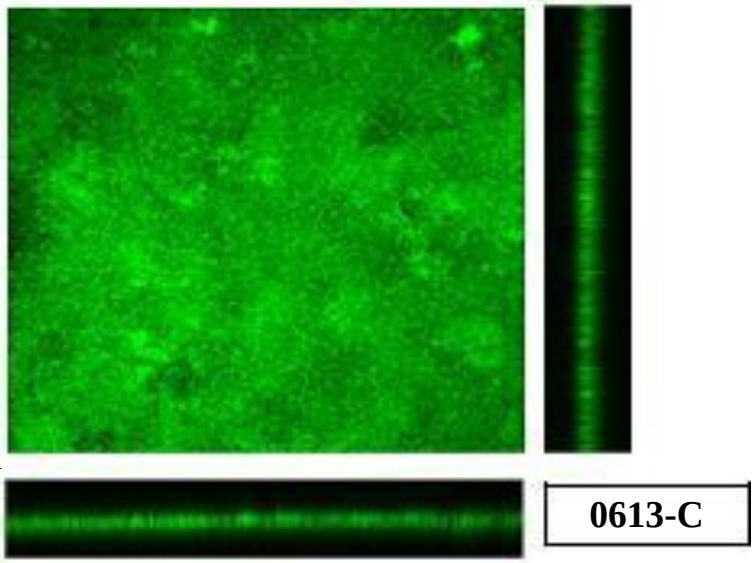
SK36-B



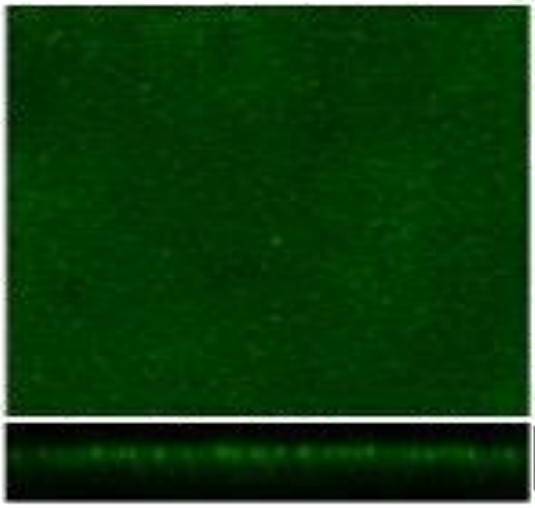
0613-B



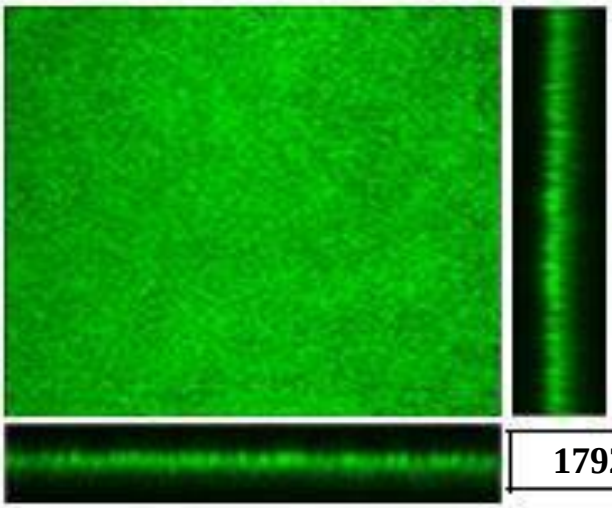
SK36-C



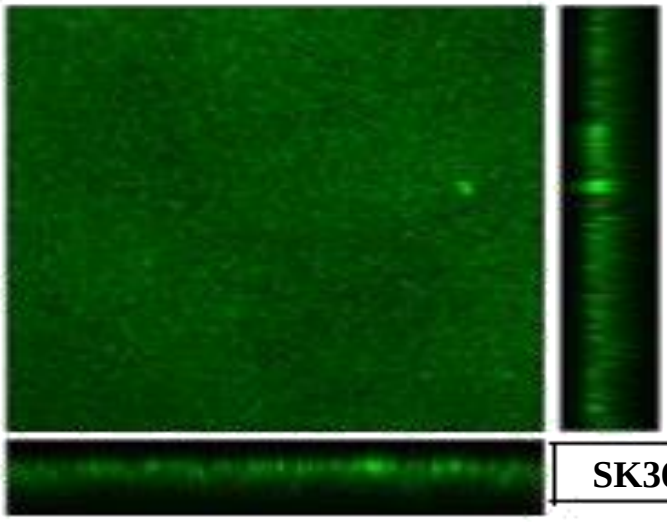
0613-C



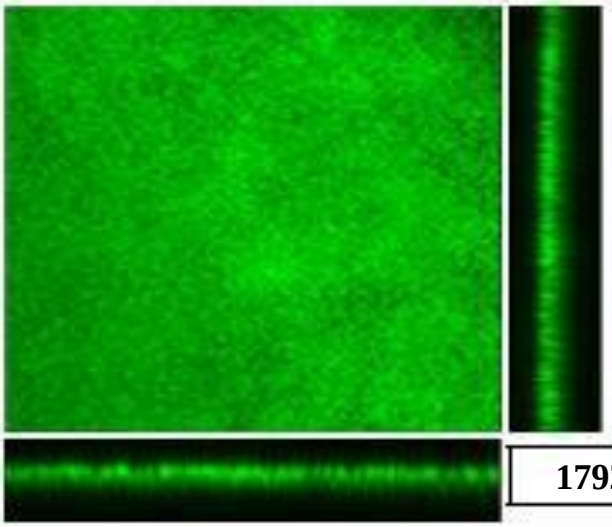
SK36-A



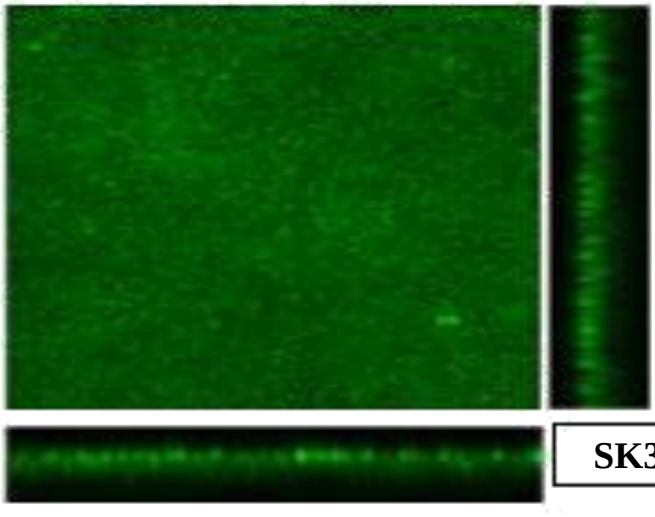
1792-A



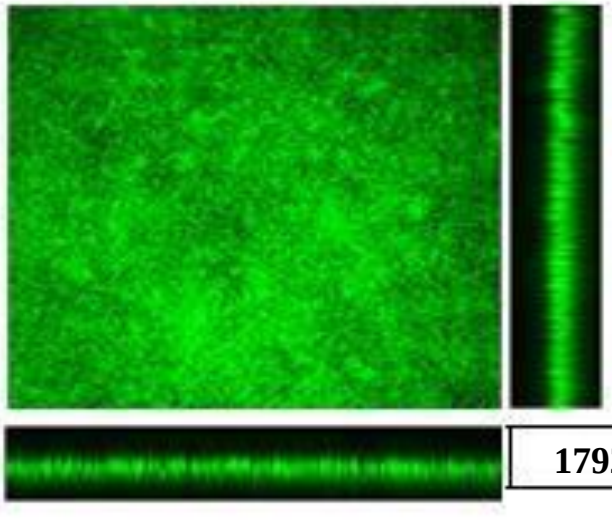
SK36-B



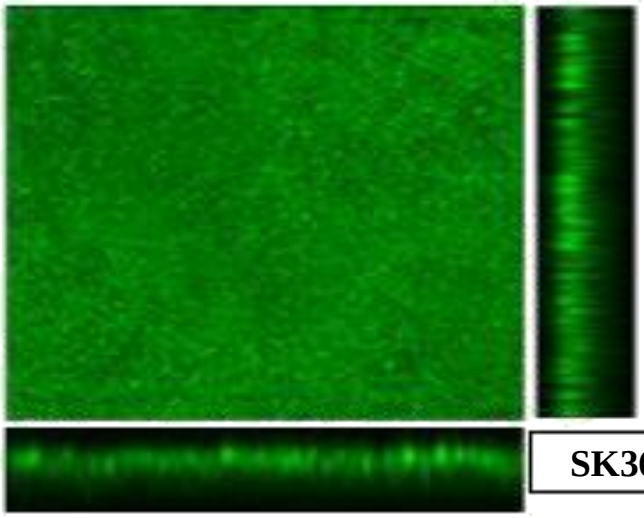
1792-B



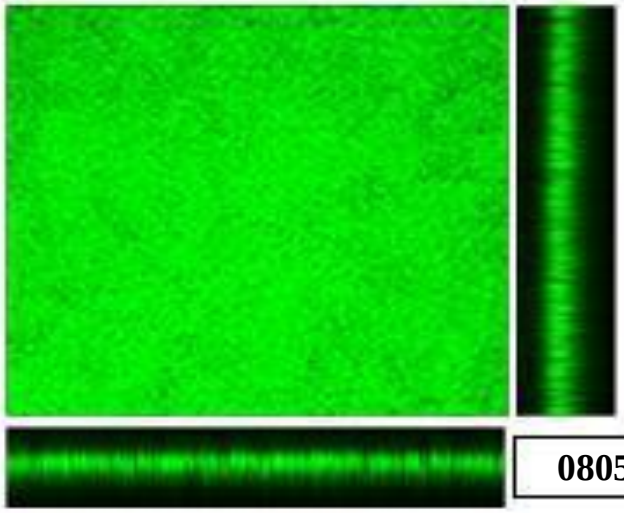
SK36-C



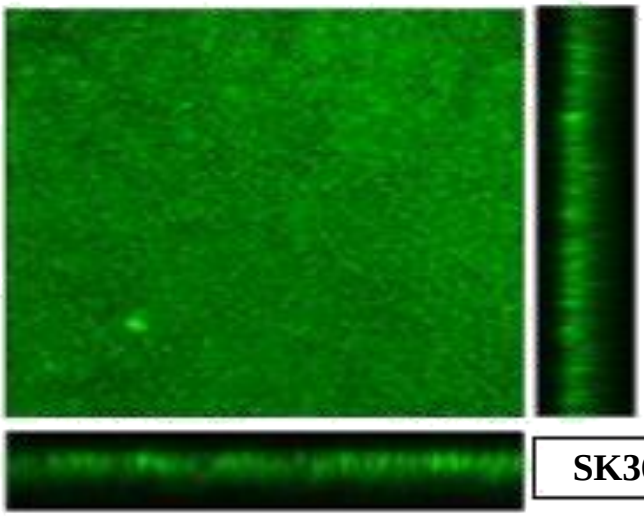
1792-C



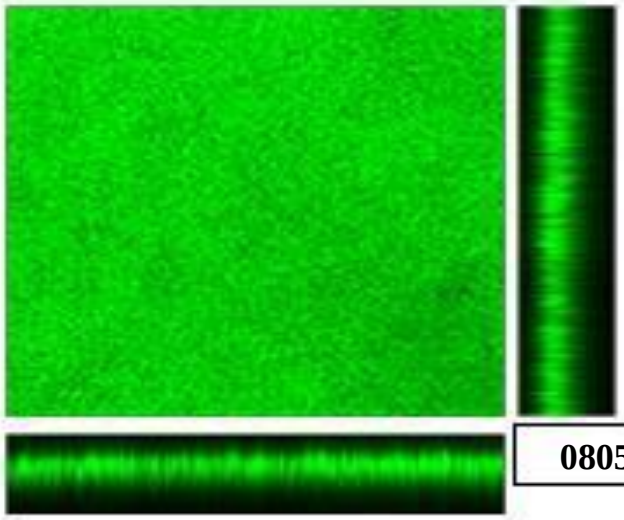
SK36-A



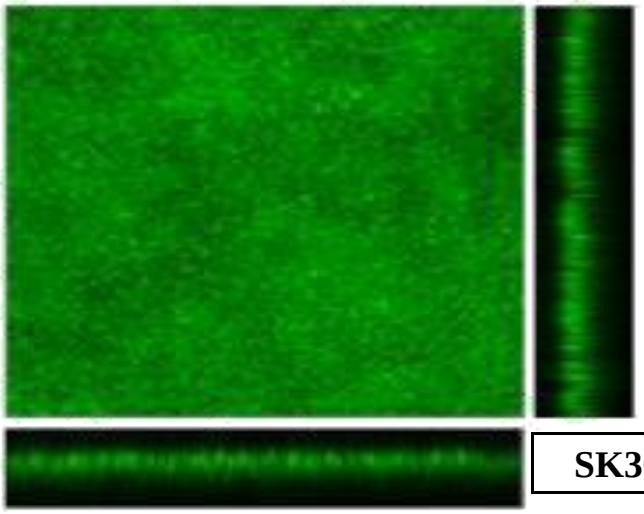
0805-A



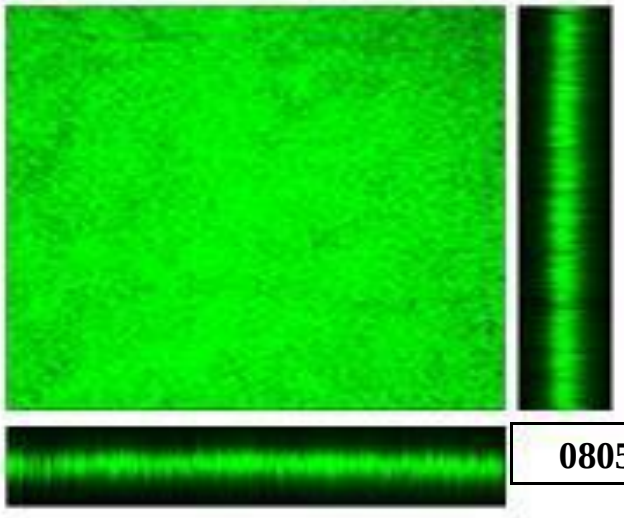
SK36-B



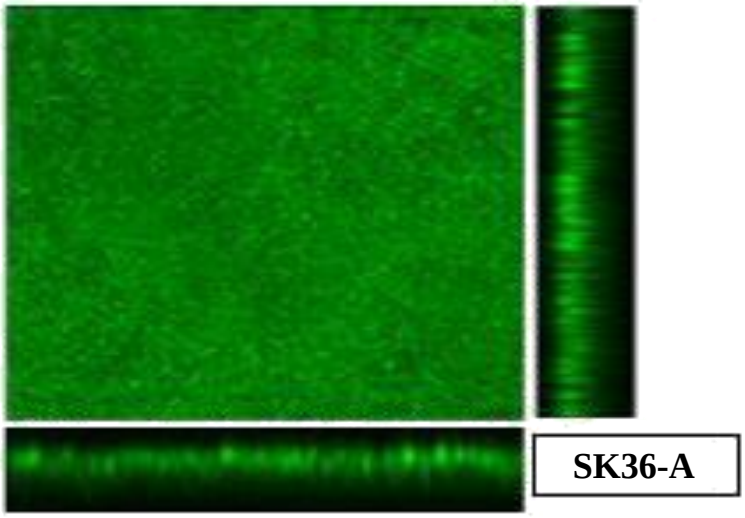
0805-B



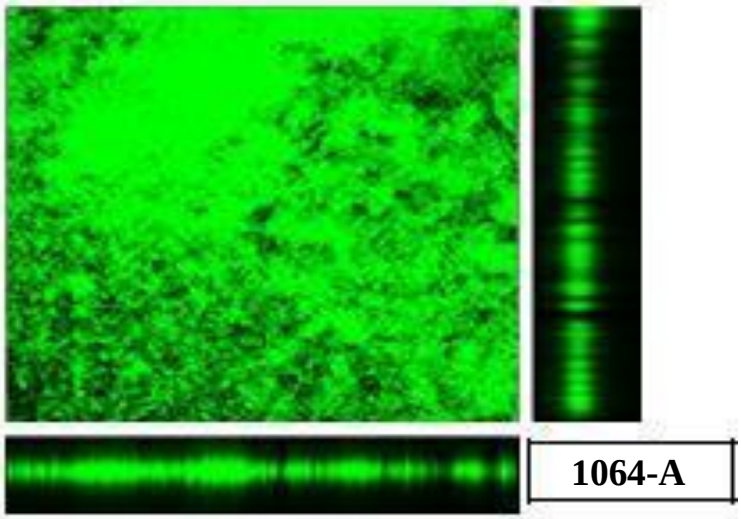
SK36-C



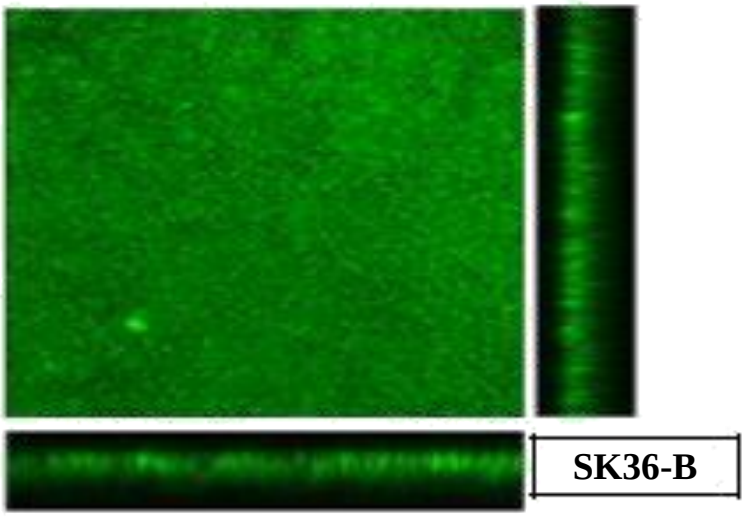
0805-C



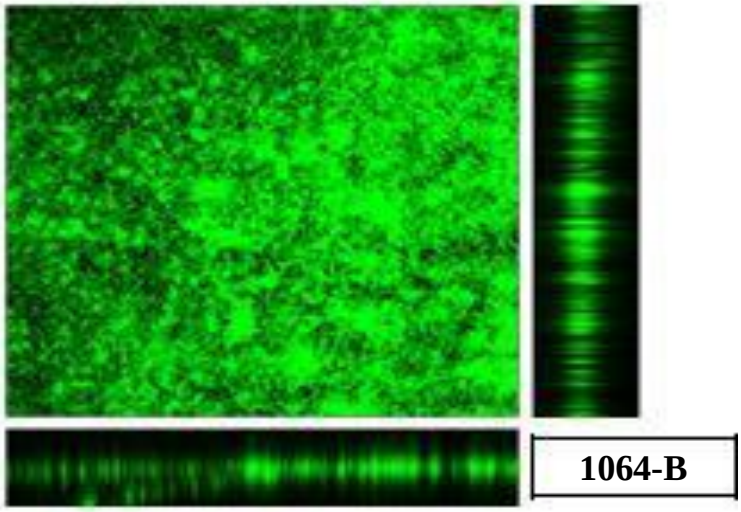
SK36-A



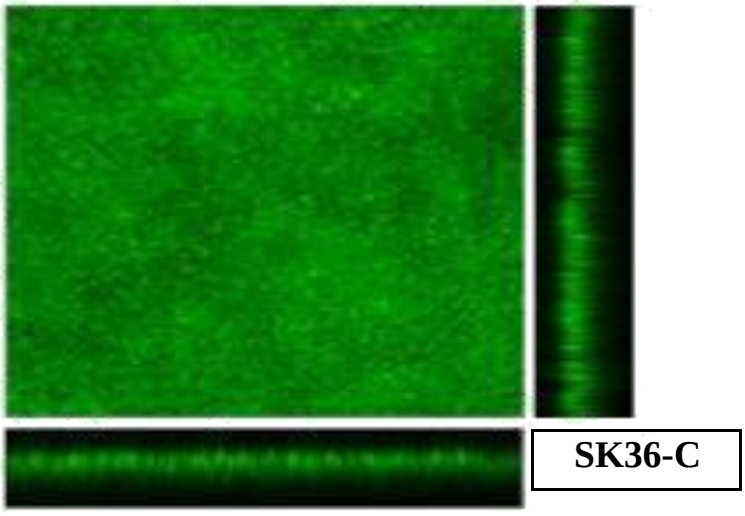
1064-A



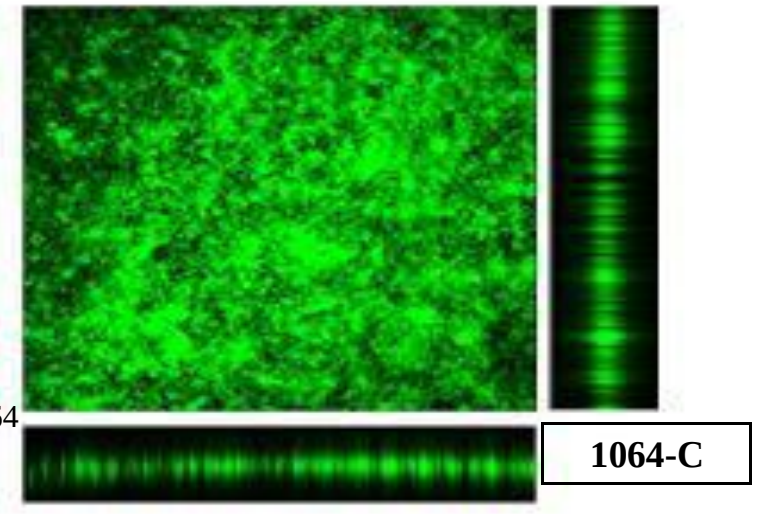
SK36-B



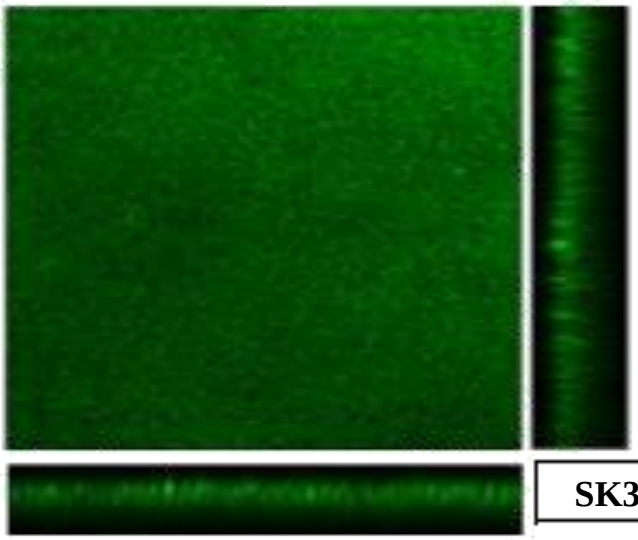
1064-B



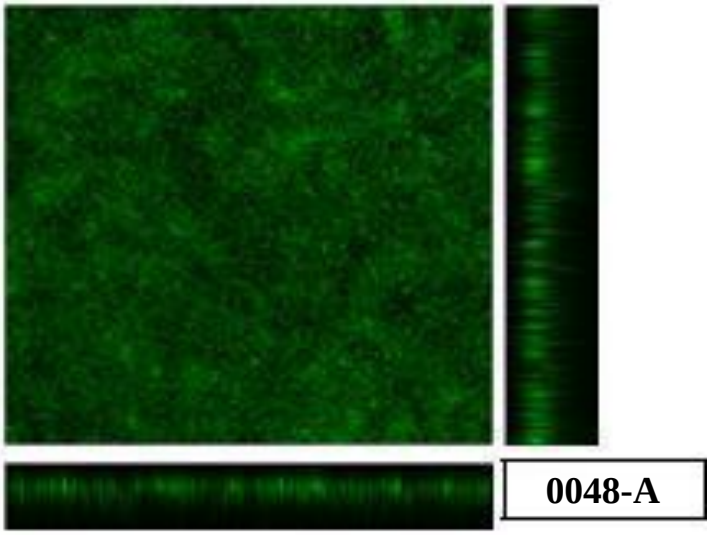
SK36-C



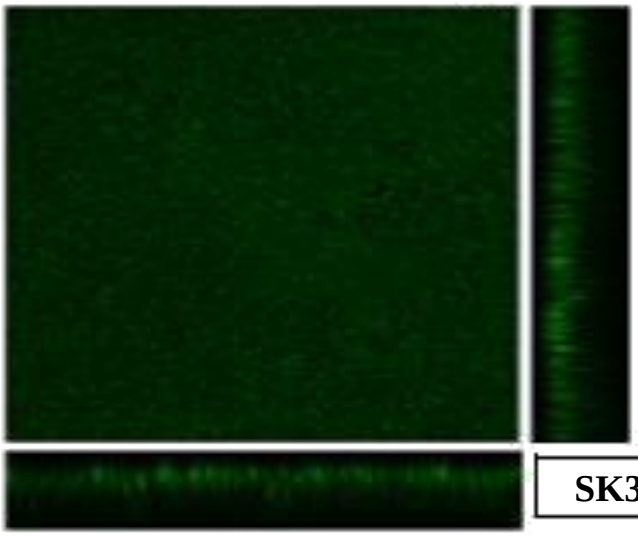
1064-C



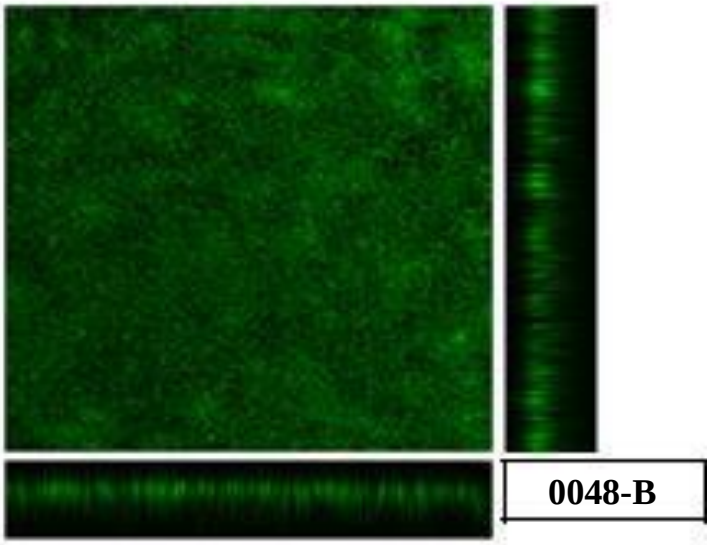
SK36-A



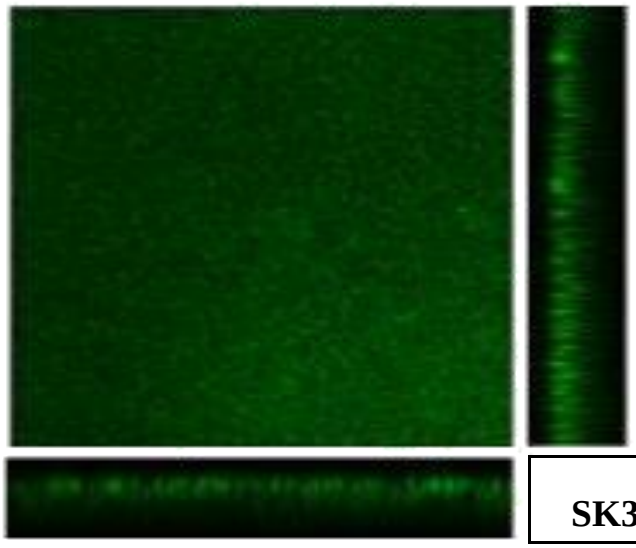
0048-A



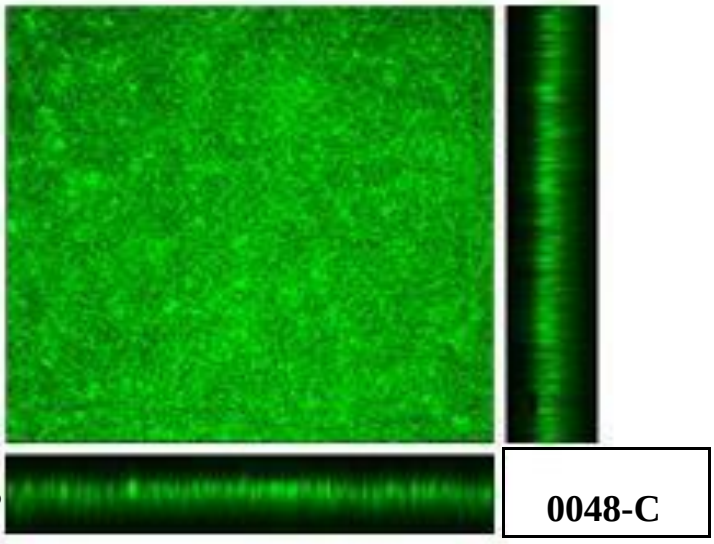
SK36-B



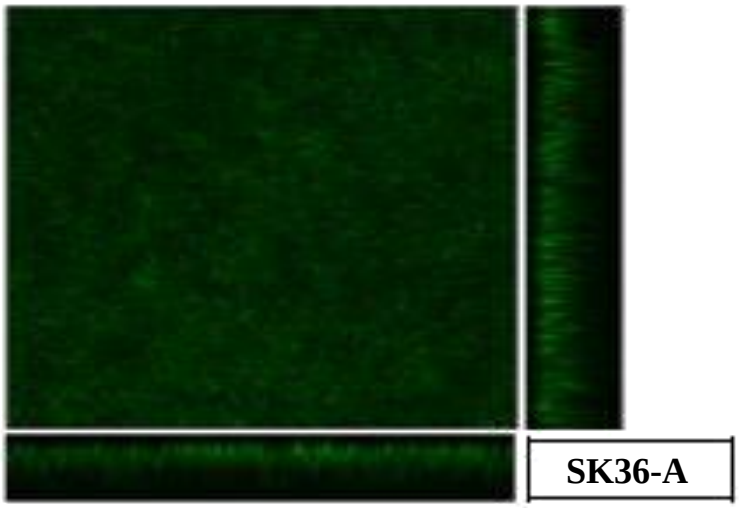
0048-B



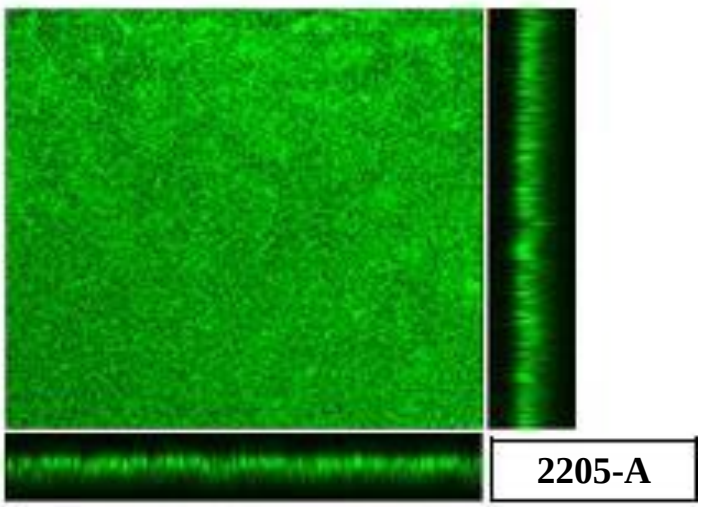
SK36-C



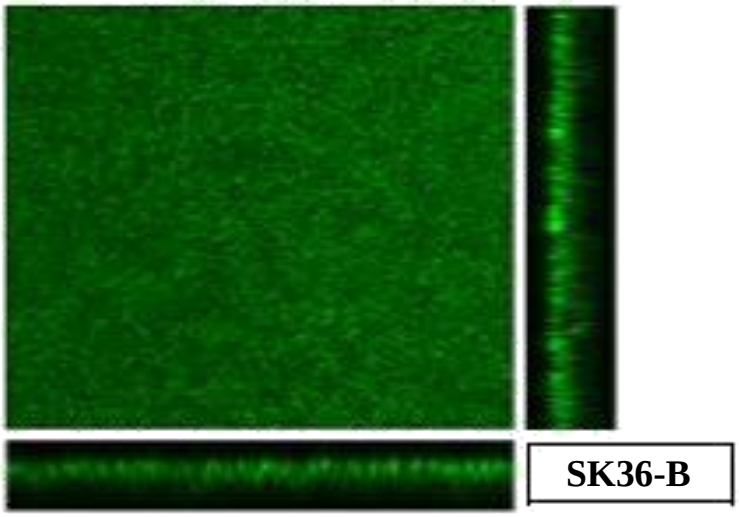
0048-C



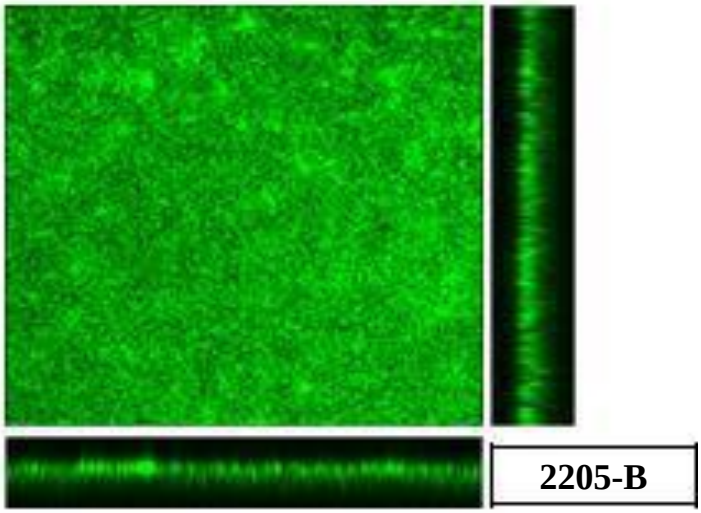
SK36-A



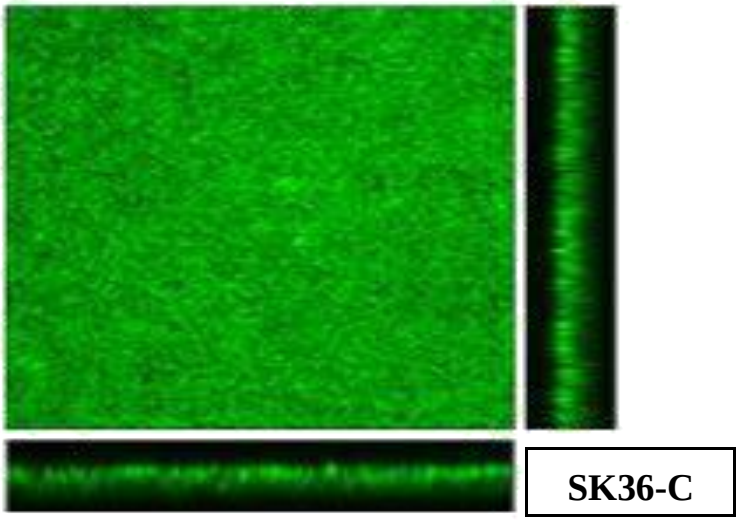
2205-A



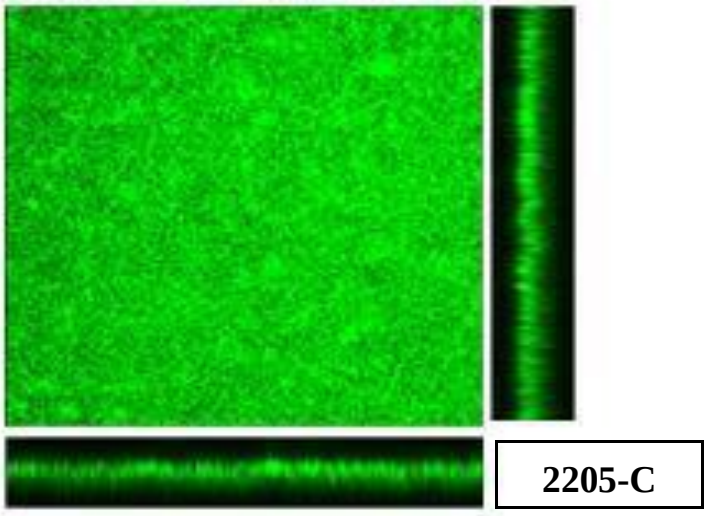
SK36-B



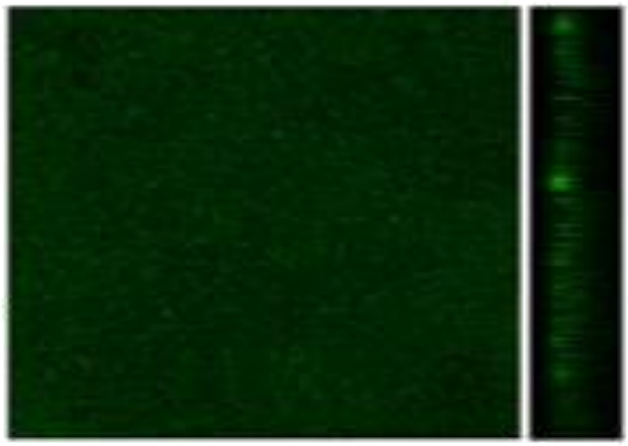
2205-B



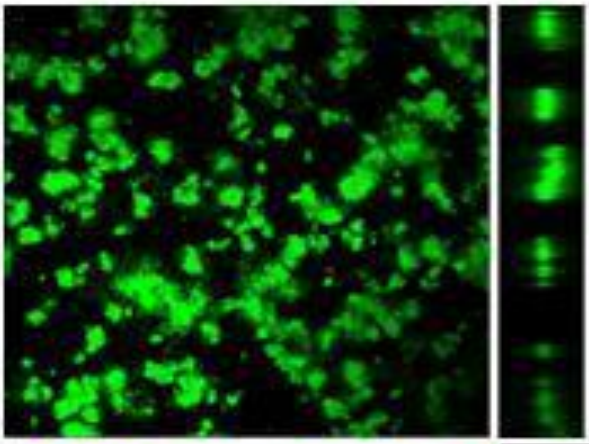
SK36-C



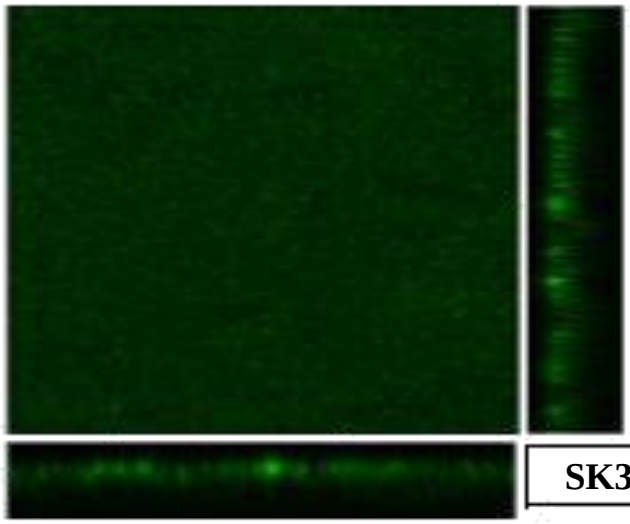
2205-C



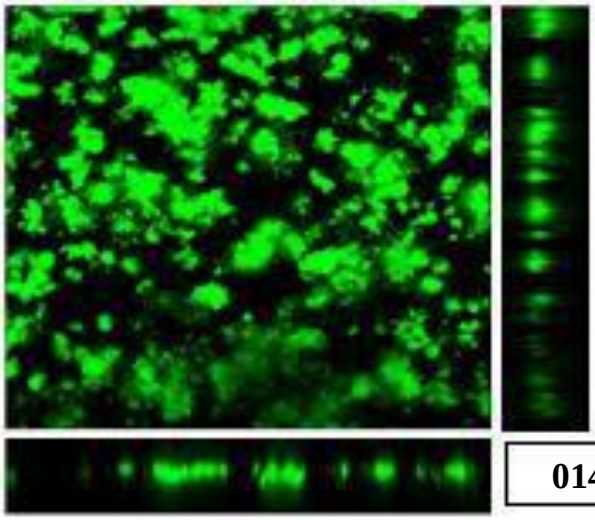
SK36-A



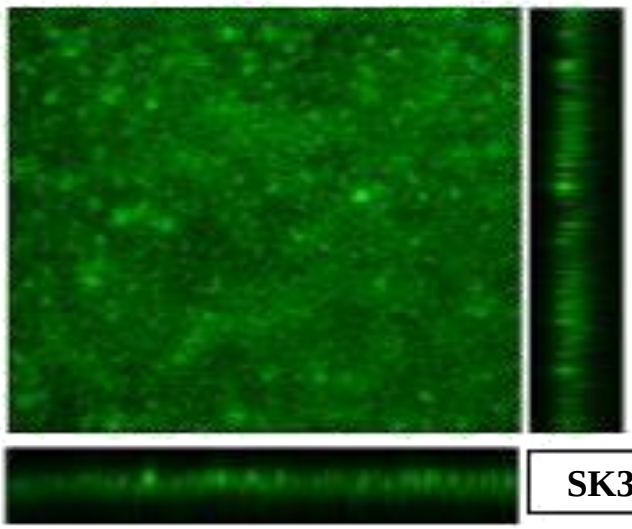
0144-A



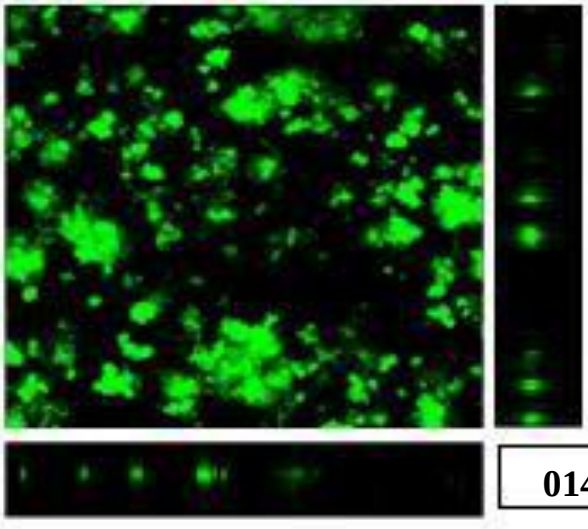
SK36-B



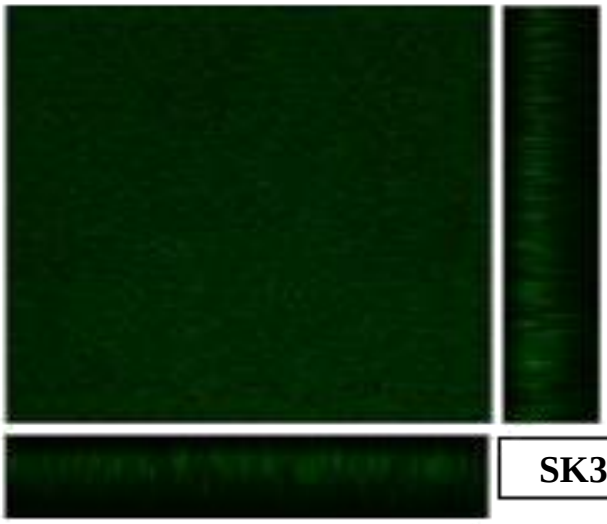
0144-B



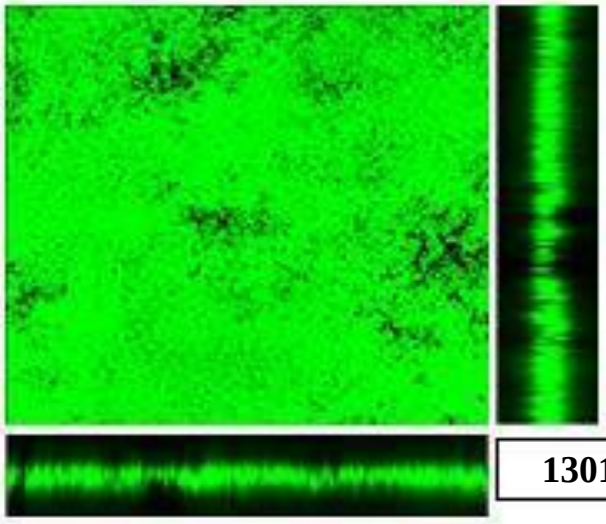
SK36-C



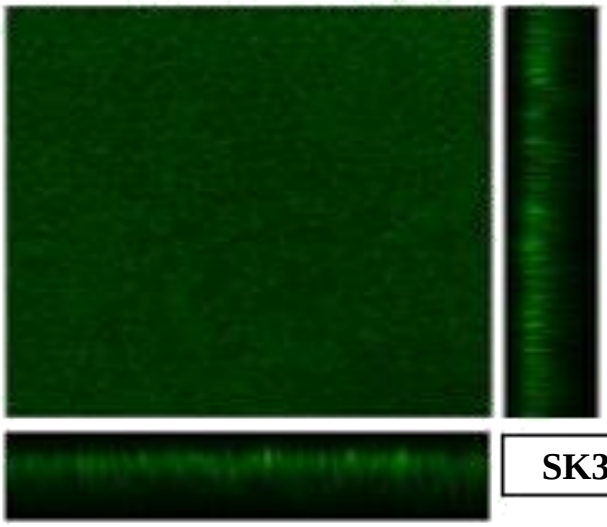
0144-C



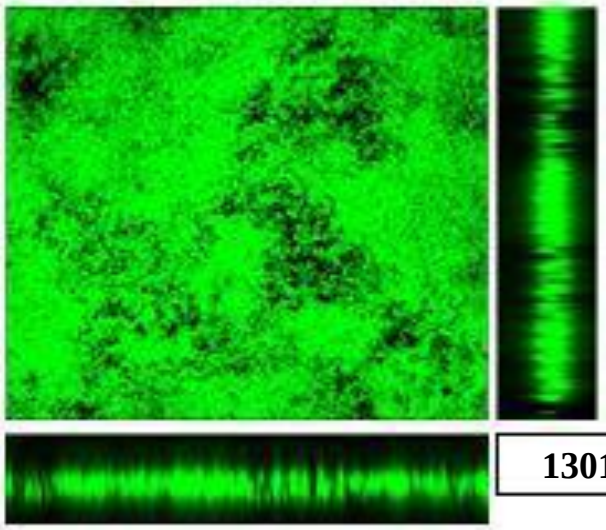
SK36-A



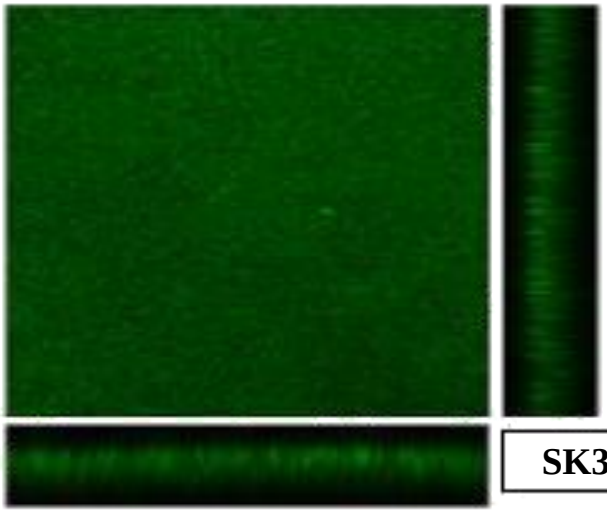
1301-A



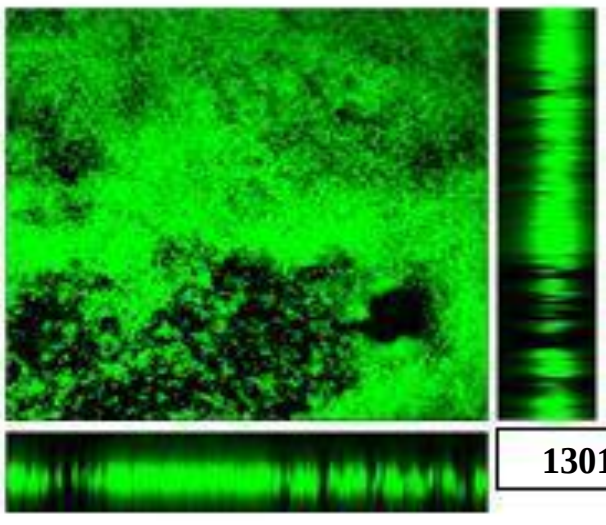
SK36-B



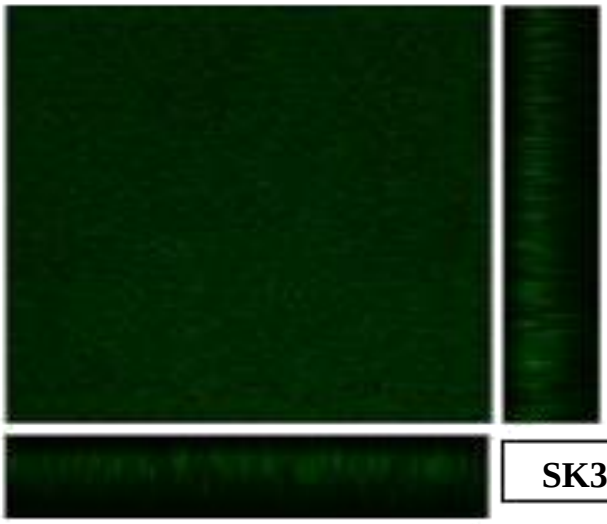
1301-B



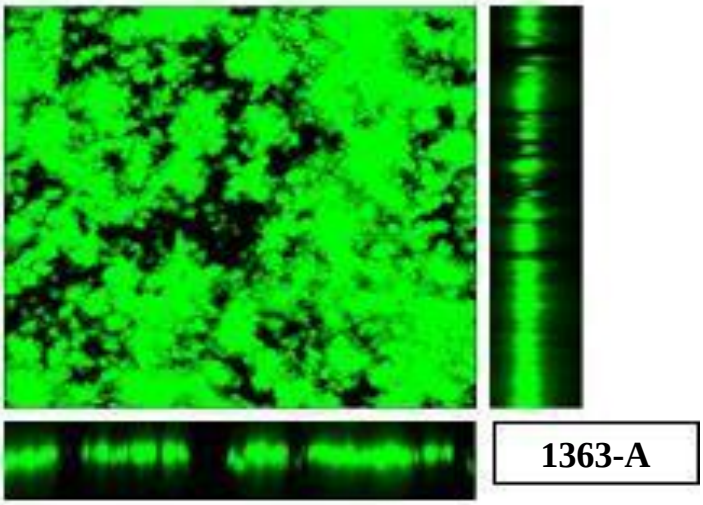
SK36-C



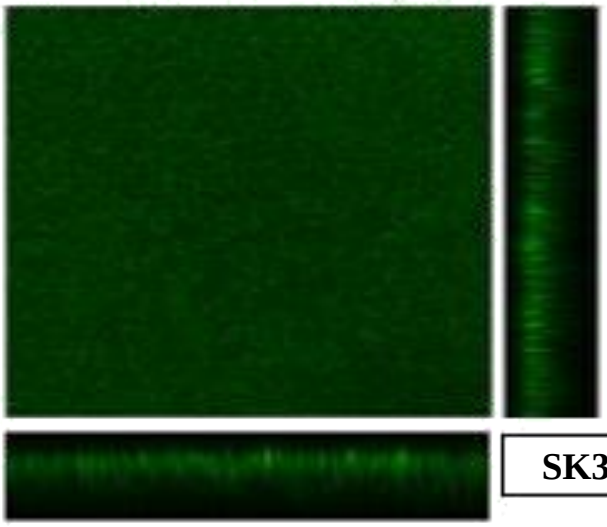
1301-C



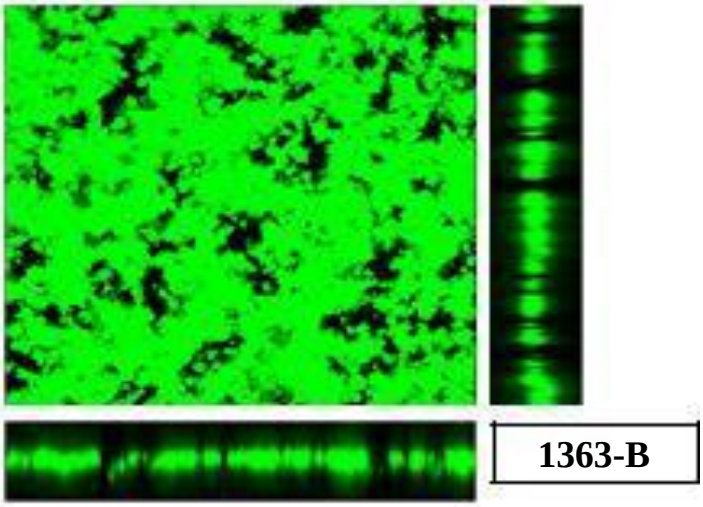
SK36-A



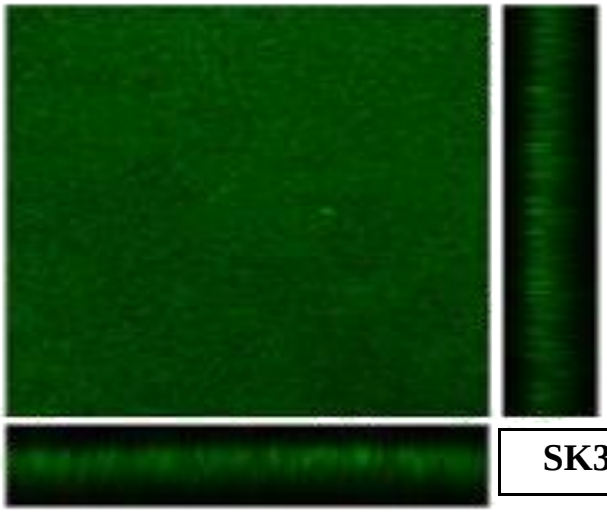
1363-A



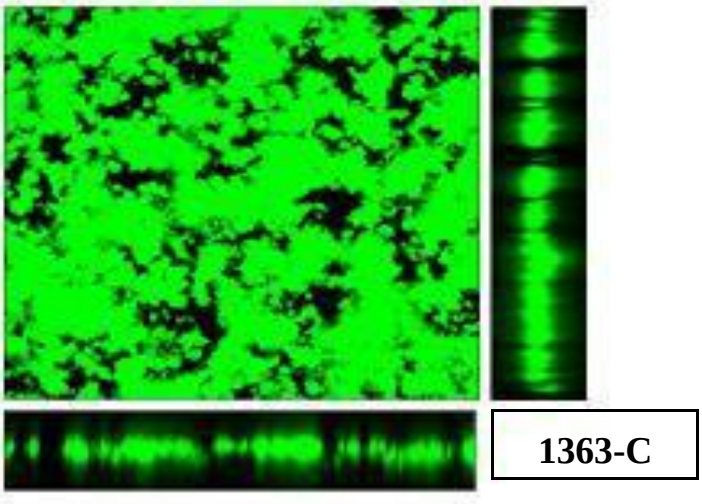
SK36-B



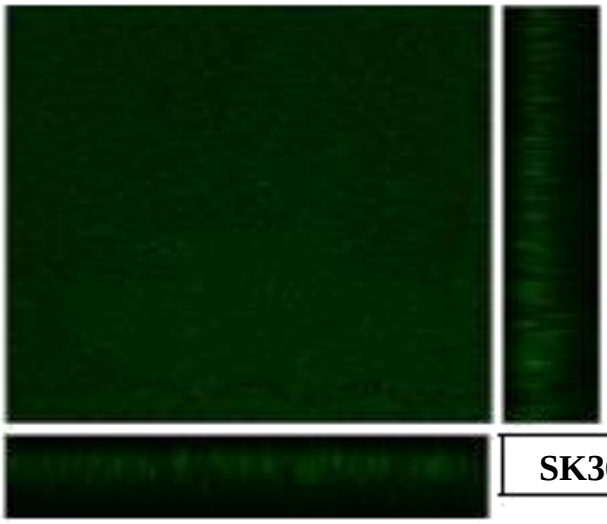
1363-B



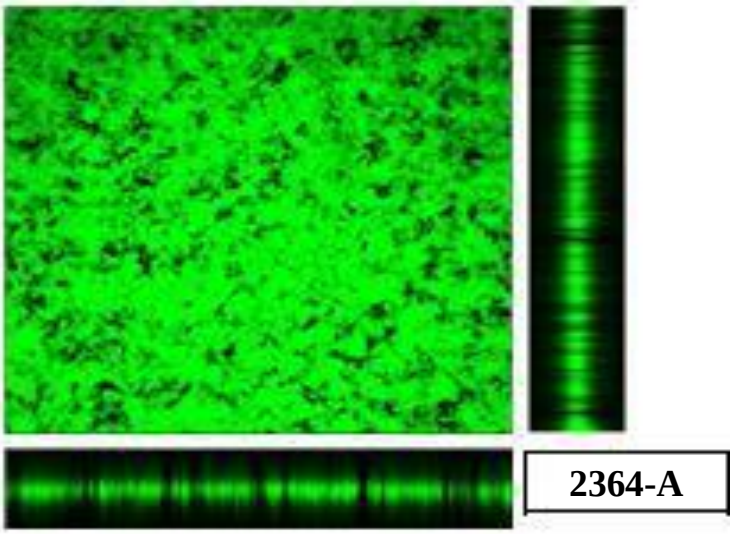
SK36-C



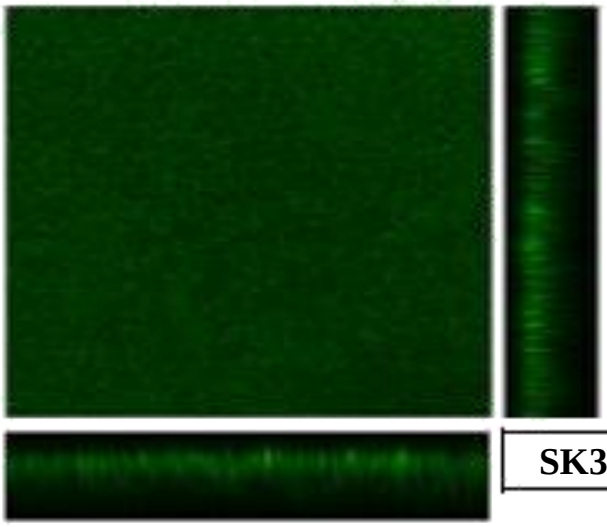
1363-C



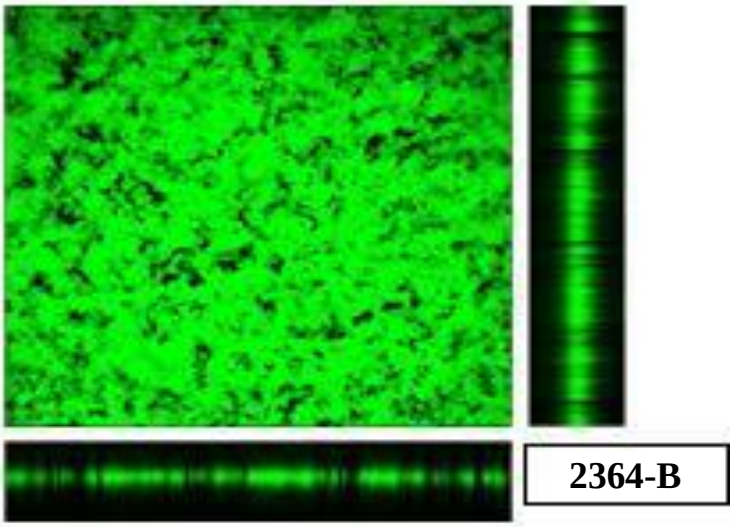
SK36-A



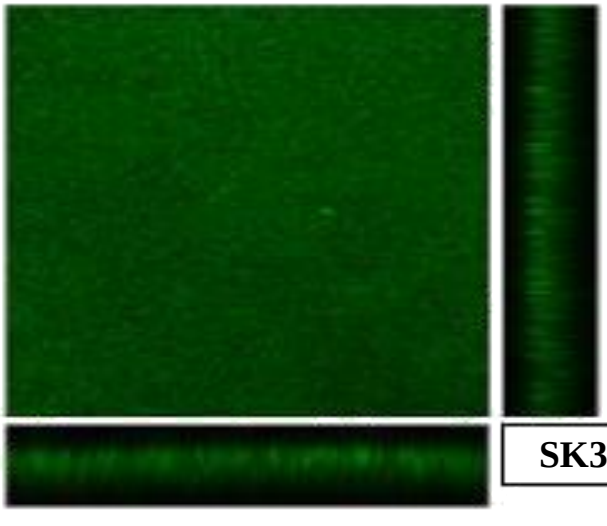
2364-A



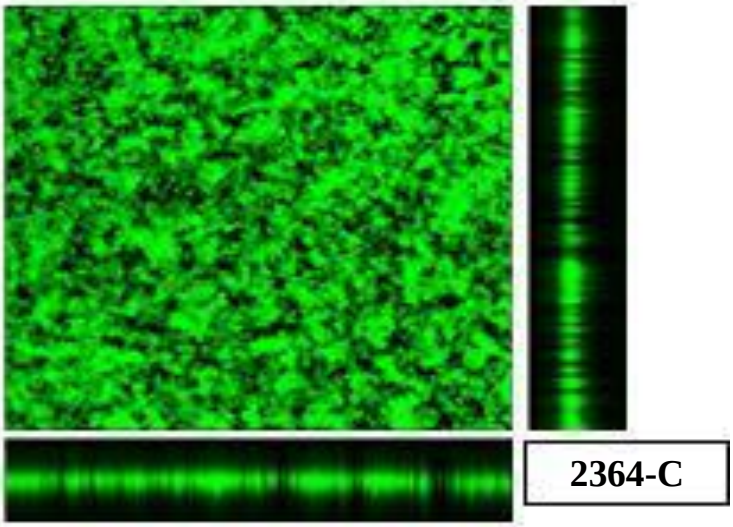
SK36-B



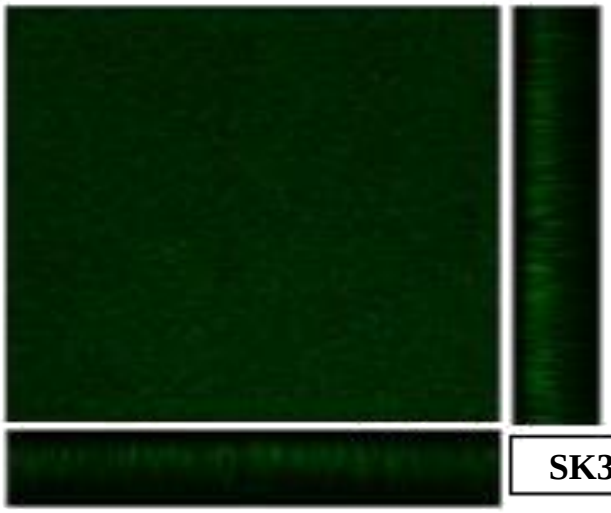
2364-B



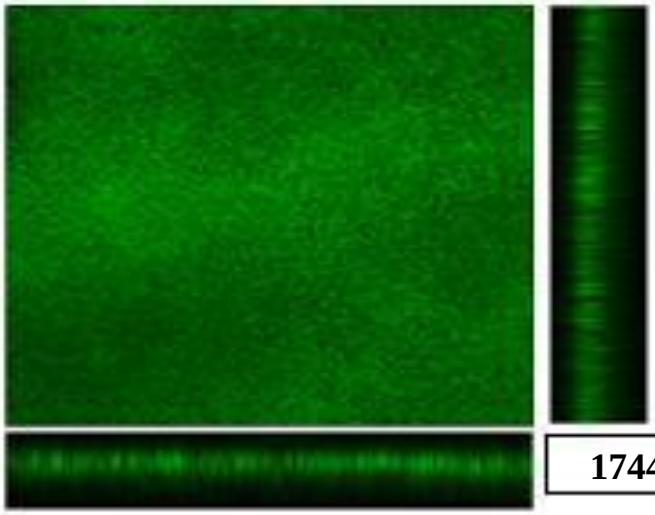
SK36-C



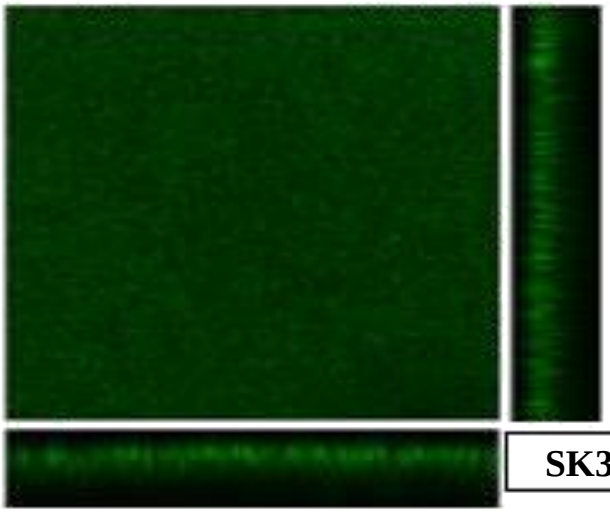
2364-C



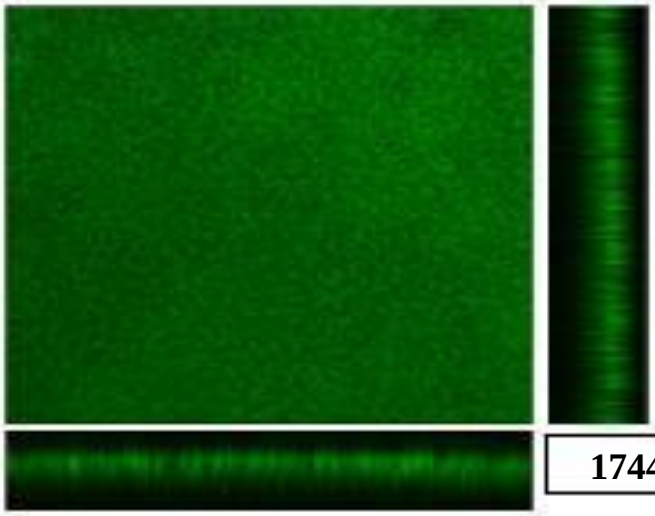
SK36-A



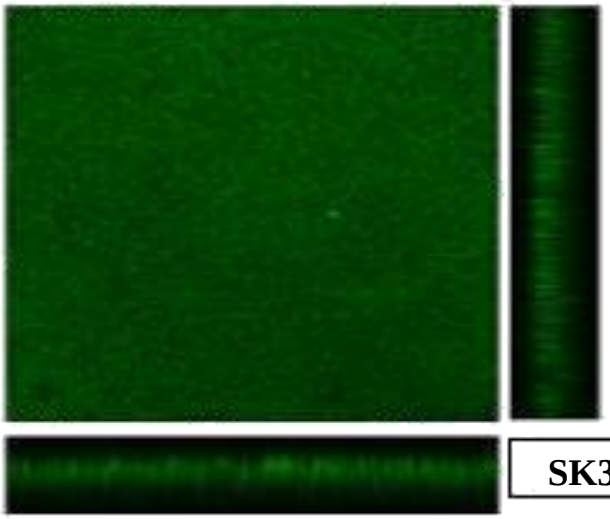
1744-A



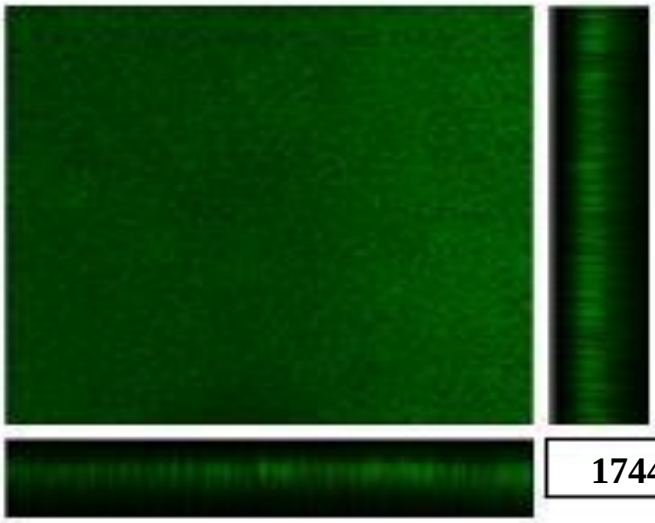
SK36-B



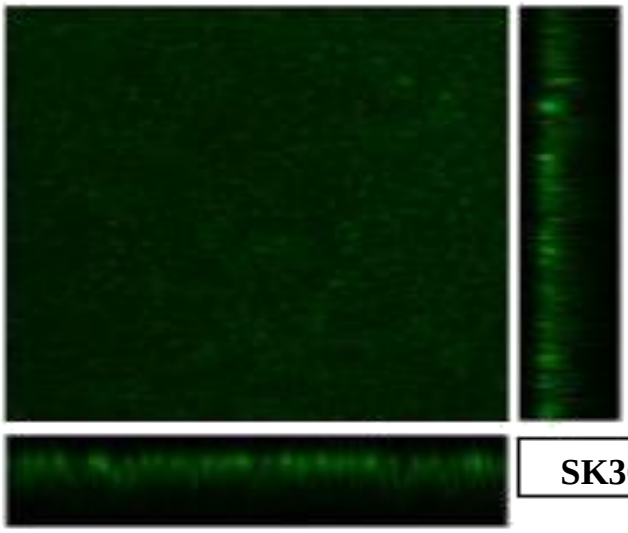
1744-B



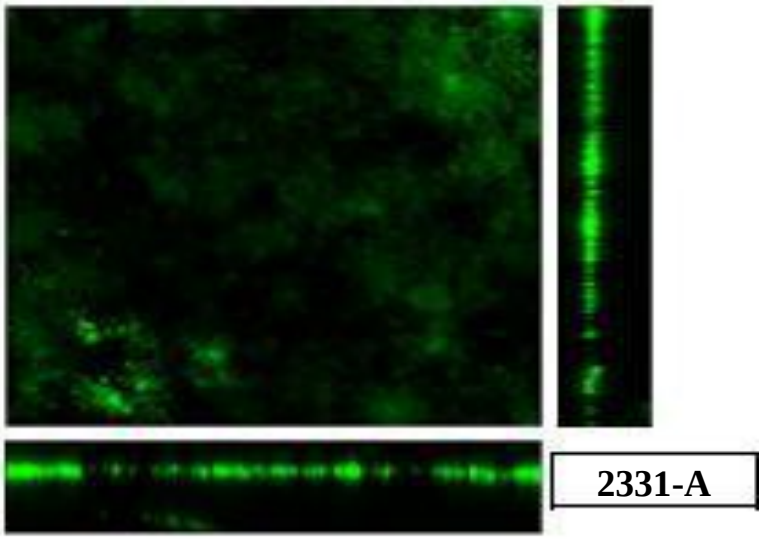
SK36-A



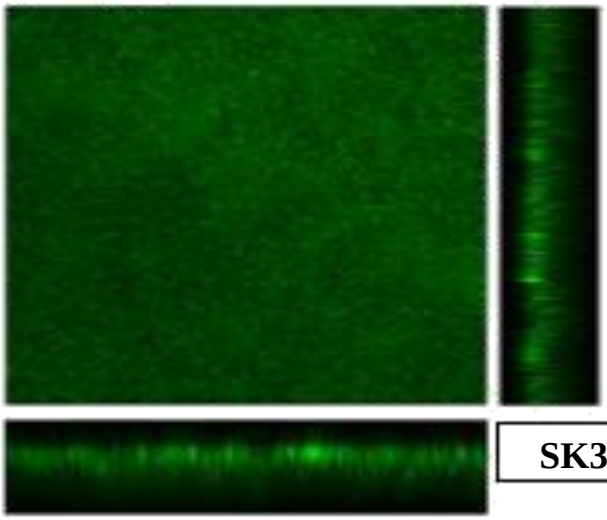
1744-C



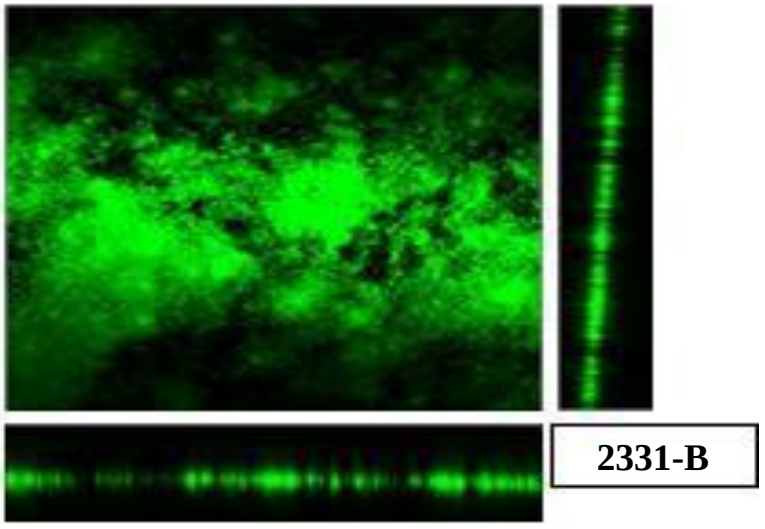
SK36-A



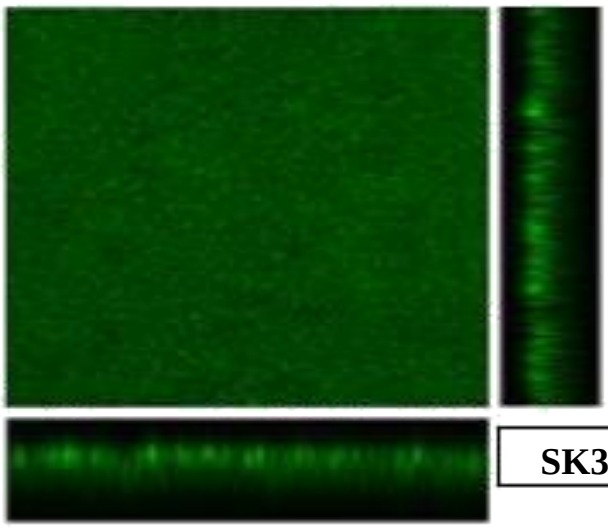
2331-A



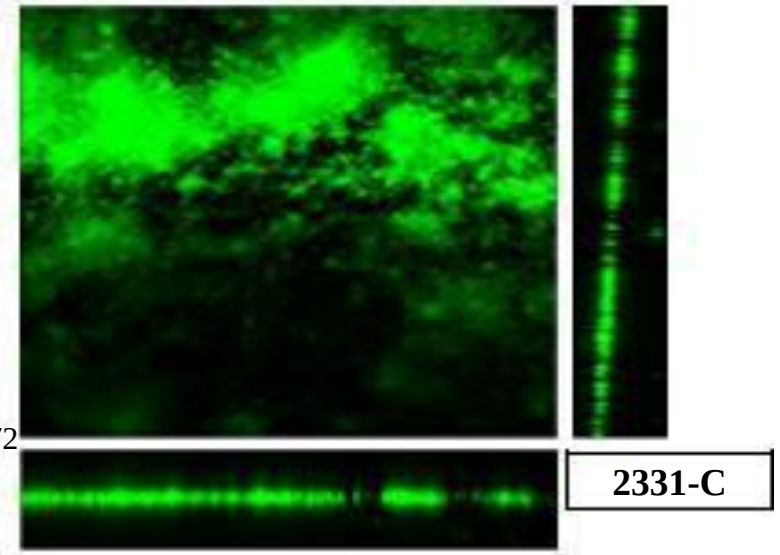
SK36-B



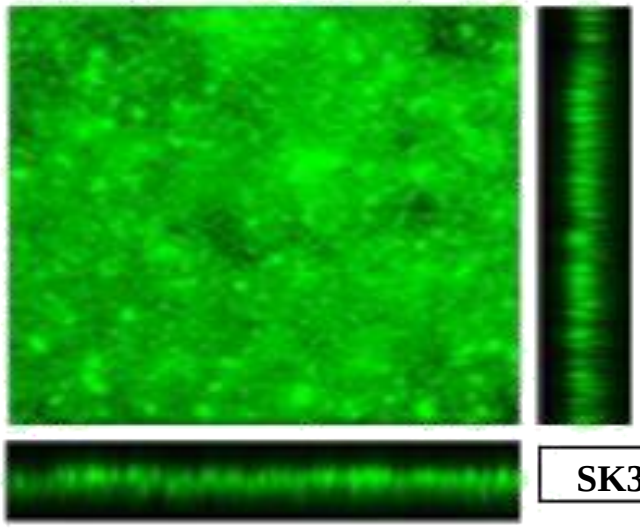
2331-B



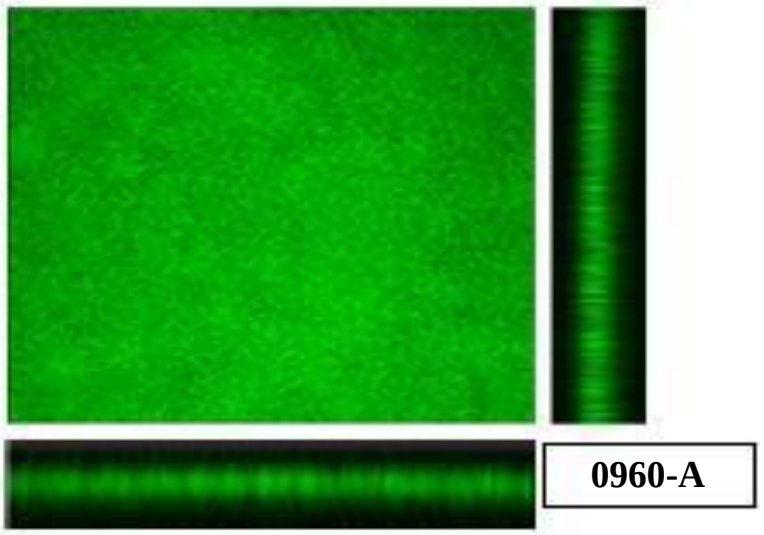
SK36-C



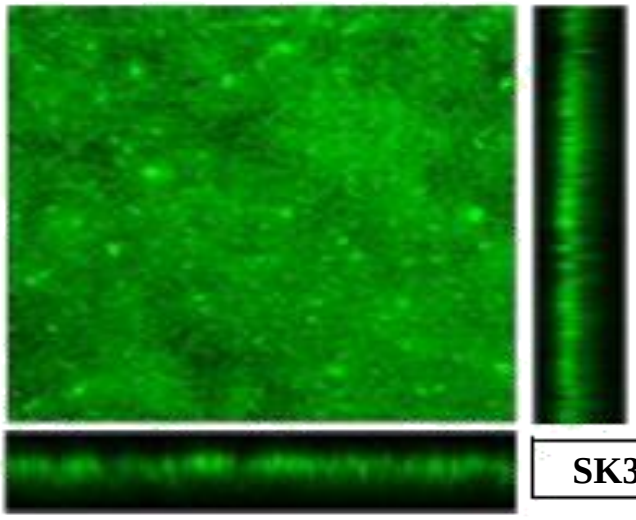
2331-C



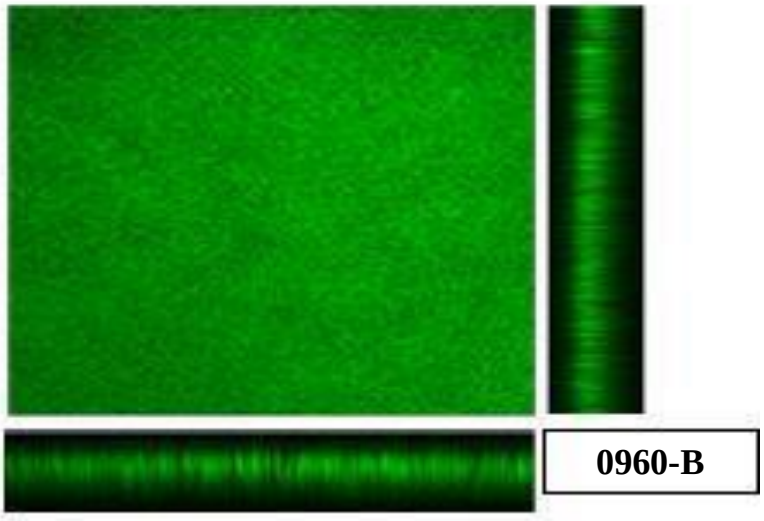
SK36-A



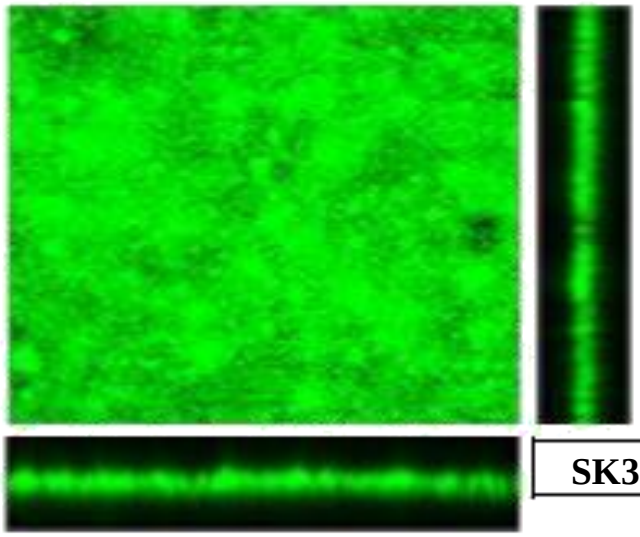
0960-A



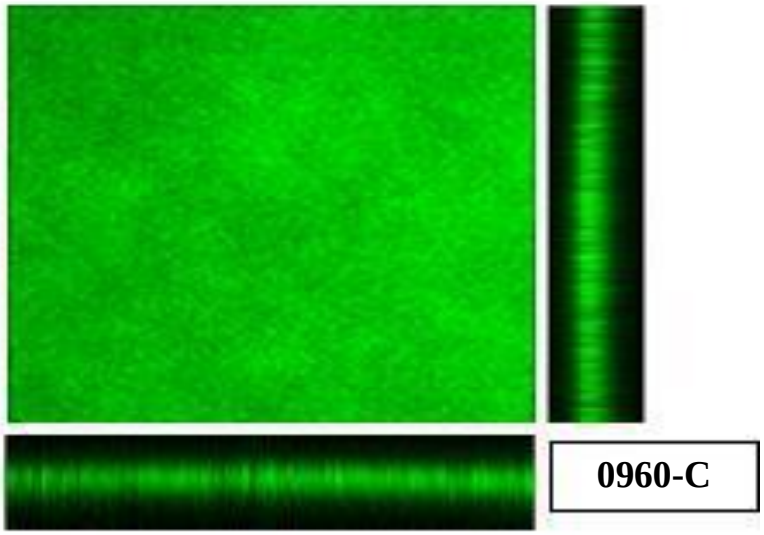
SK36-B



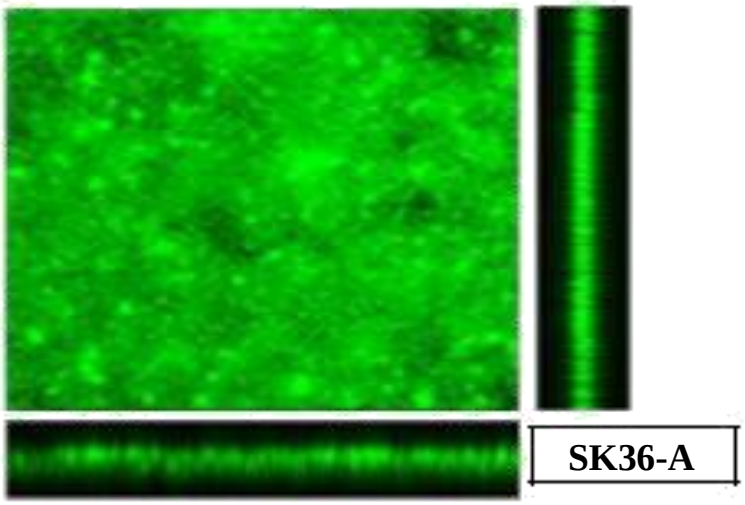
0960-B

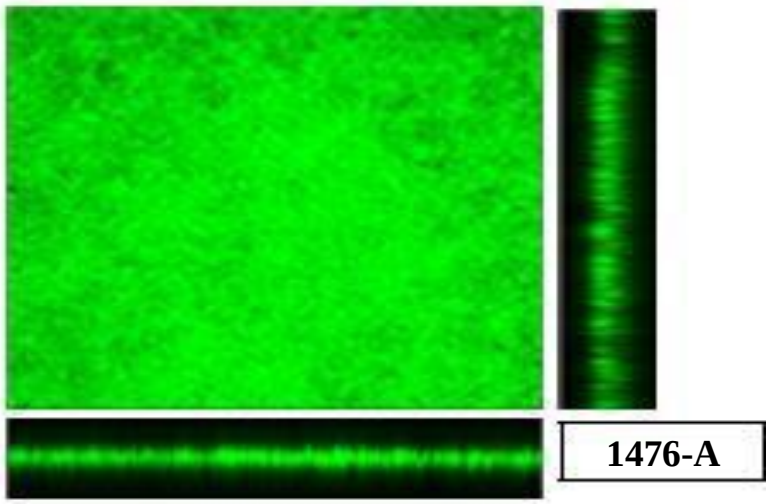


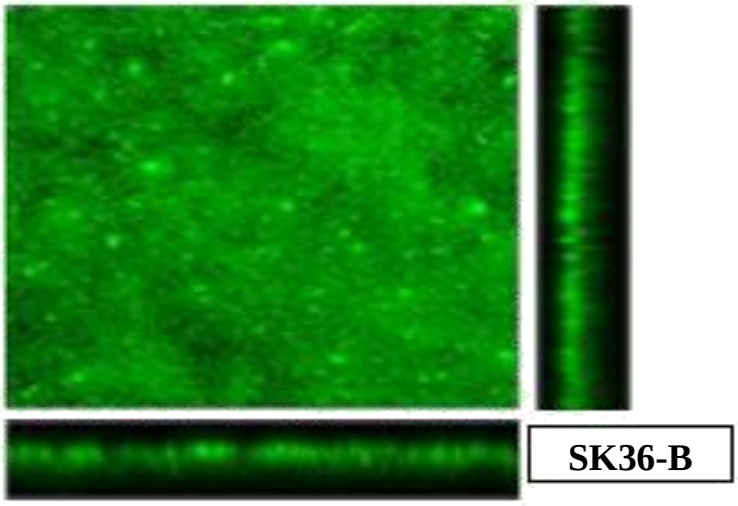
SK36-C

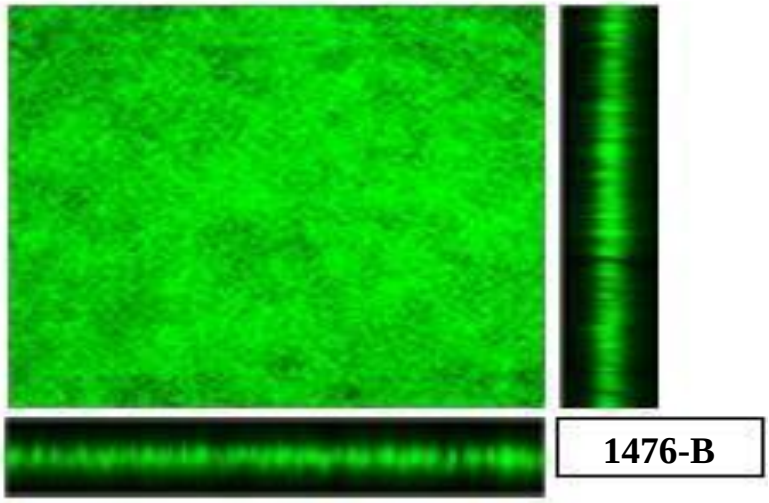


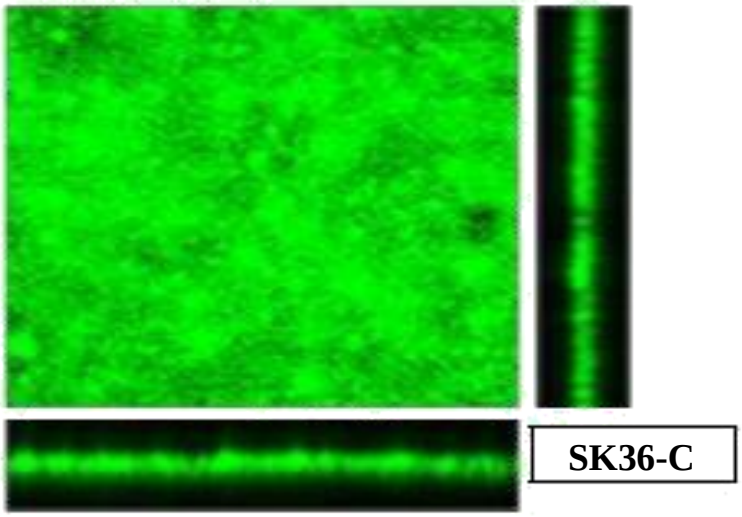
0960-C

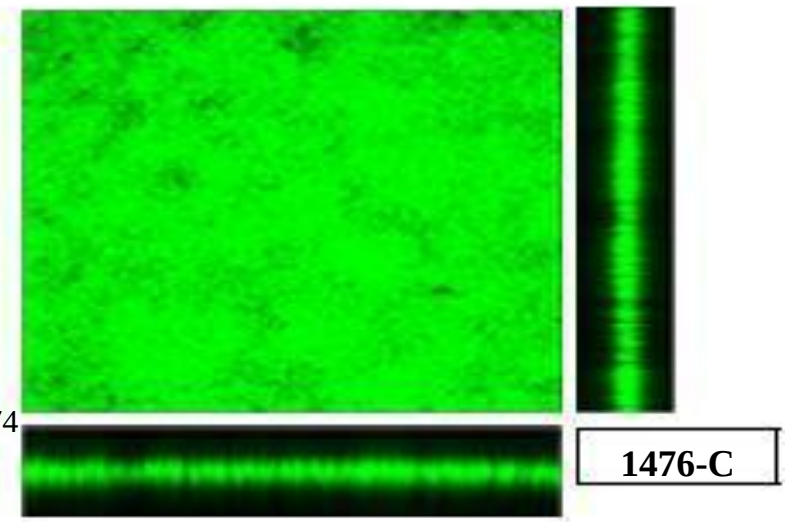

SK36-A

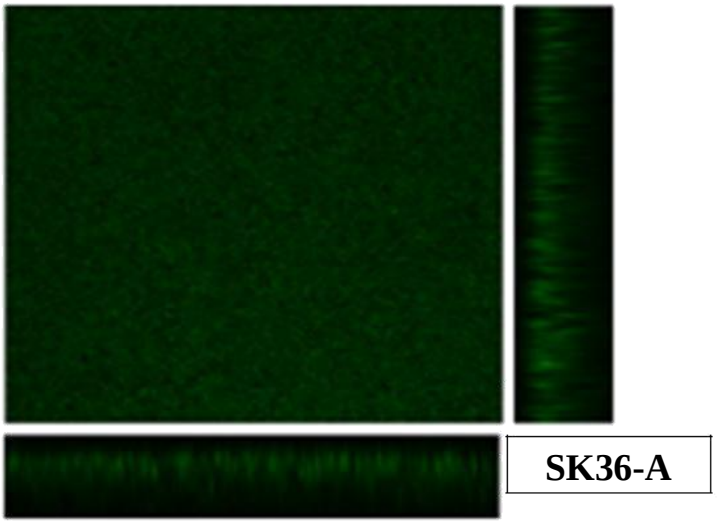

1476-A


SK36-B

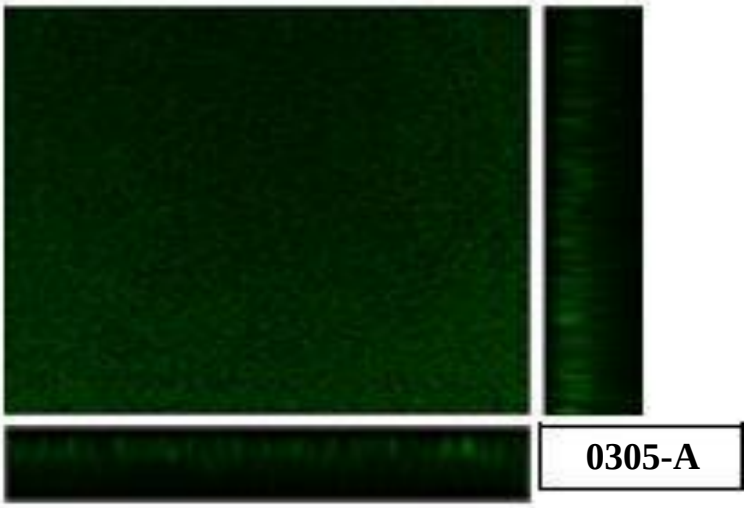

1476-B


SK36-C

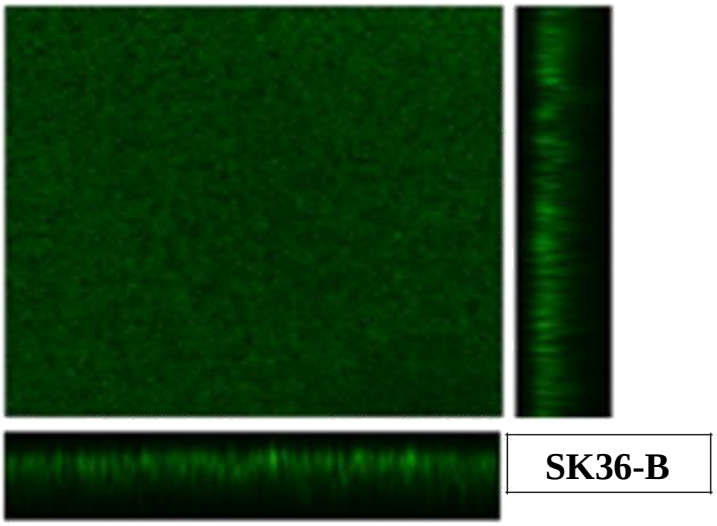

1476-C



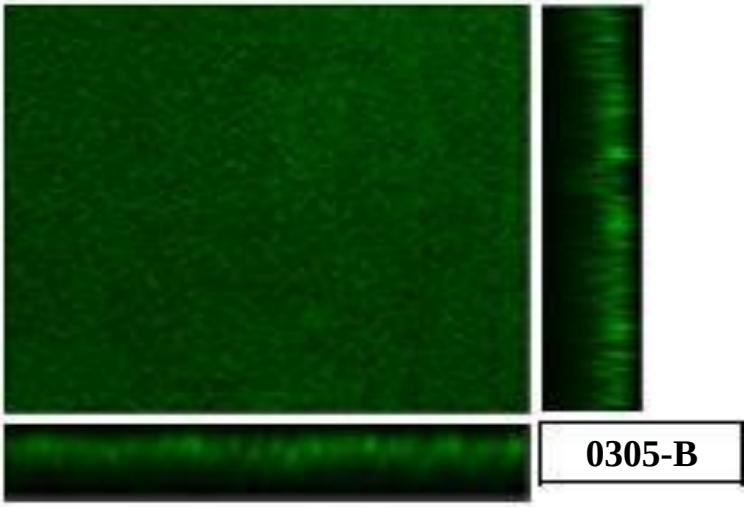
SK36-A



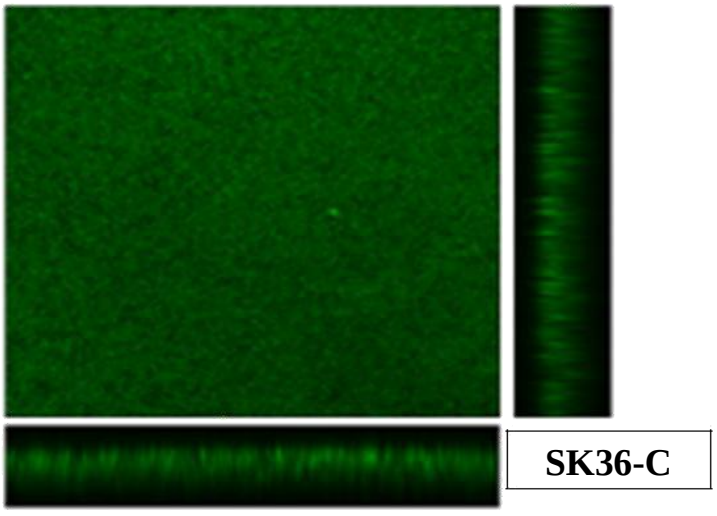
0305-A



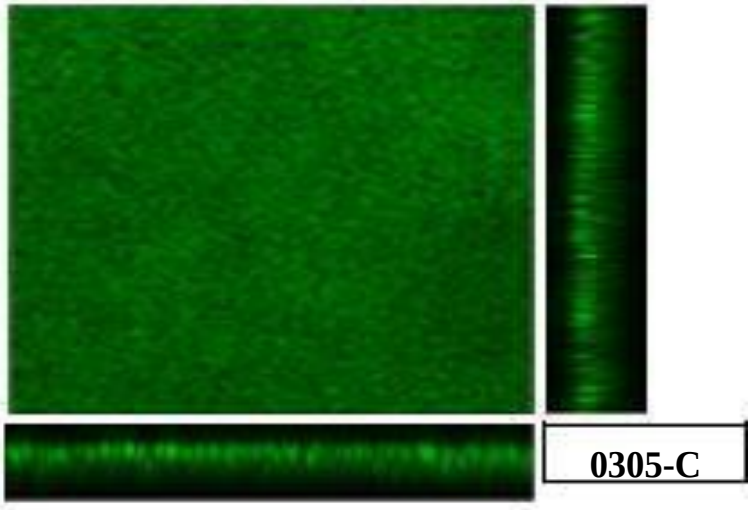
SK36-B



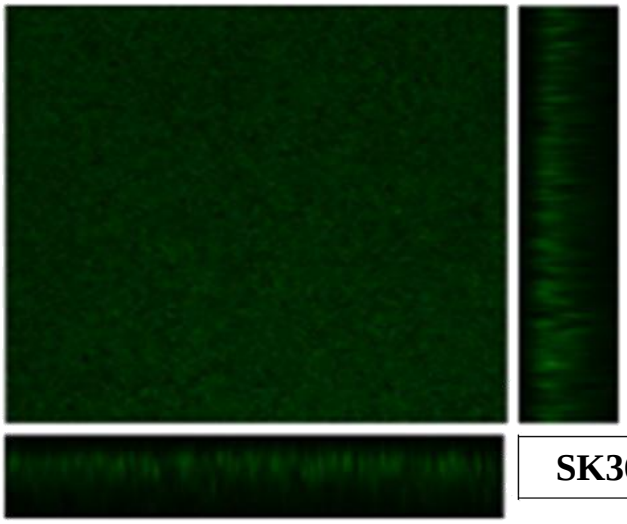
0305-B



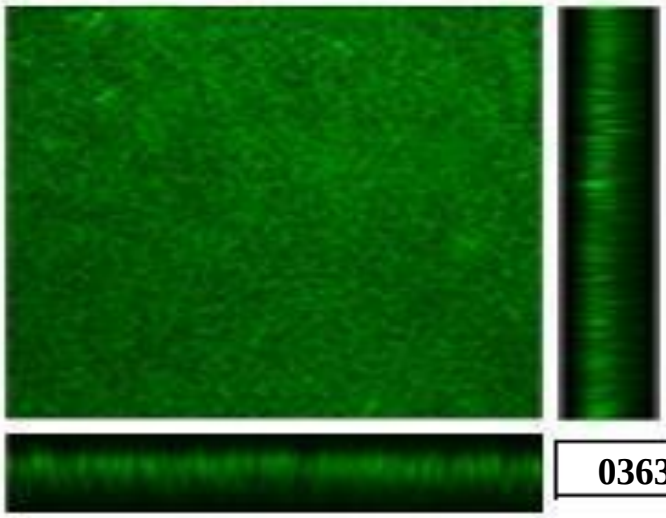
SK36-C



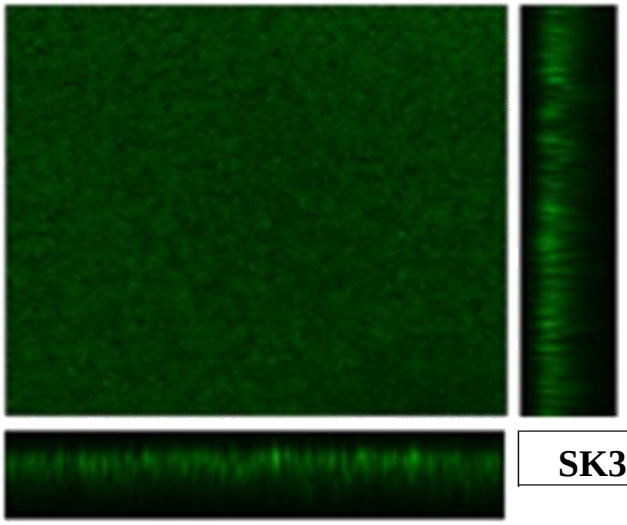
0305-C



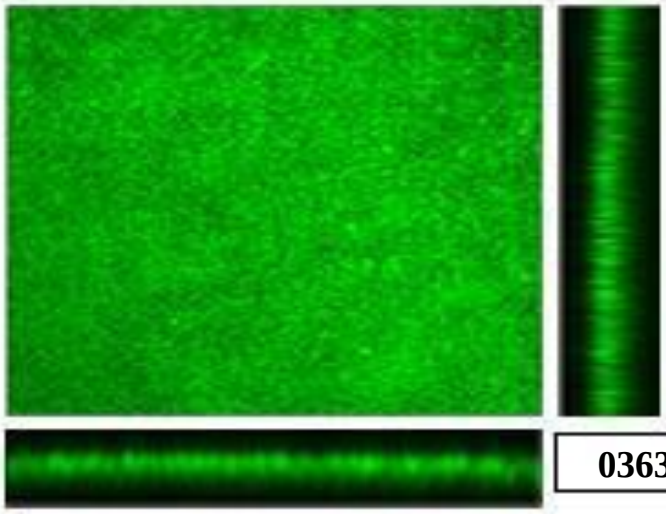
SK36-A



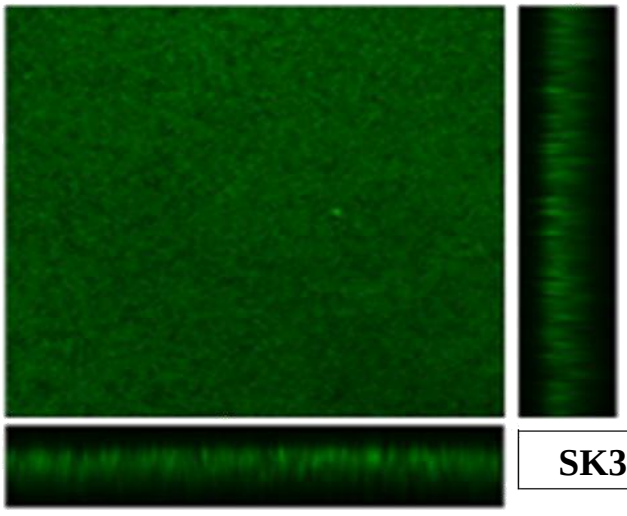
0363-A



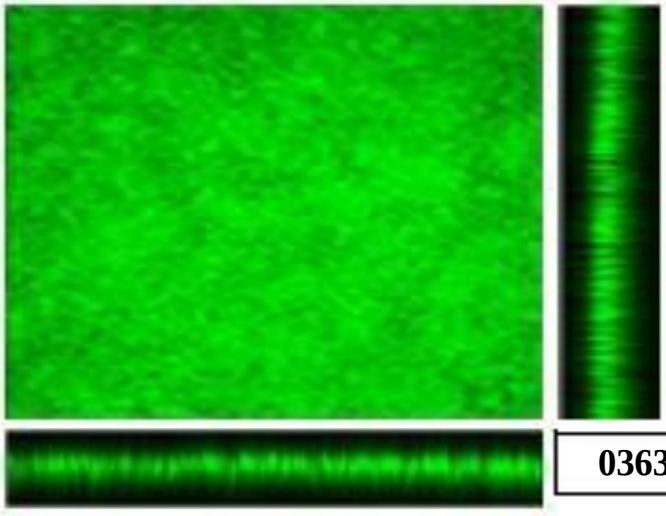
SK36-B



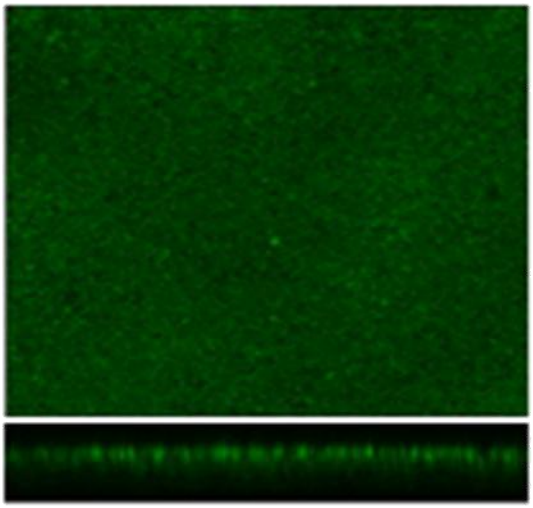
0363-B



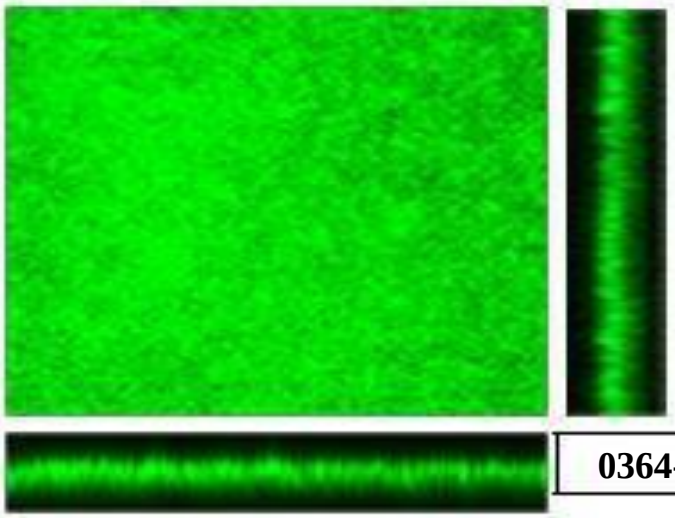
SK36-C



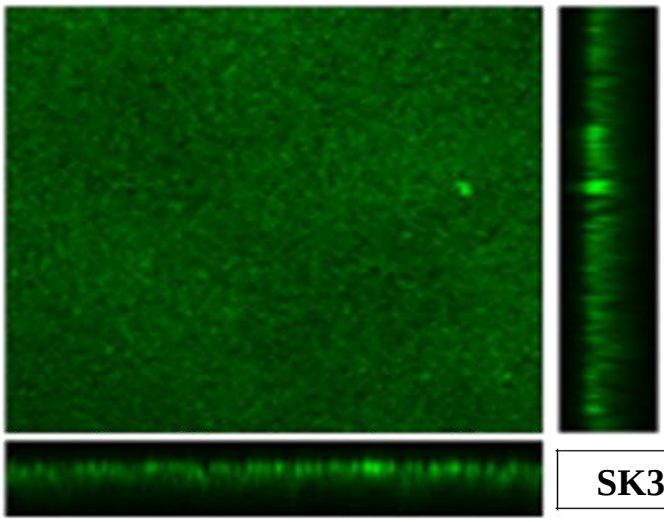
0363-C



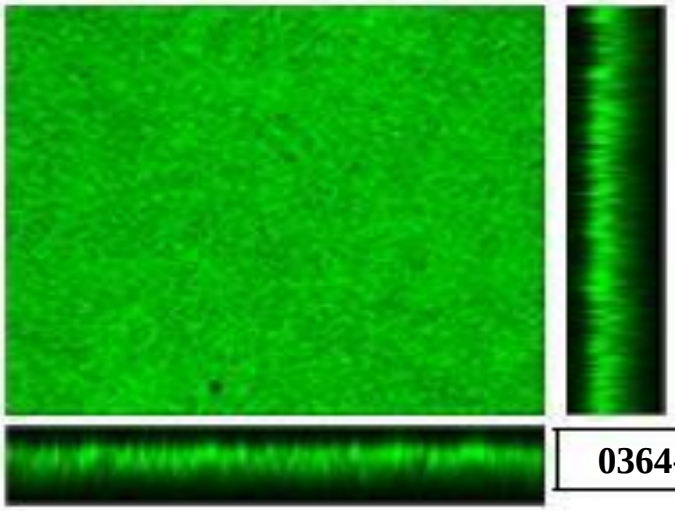
SK36-A



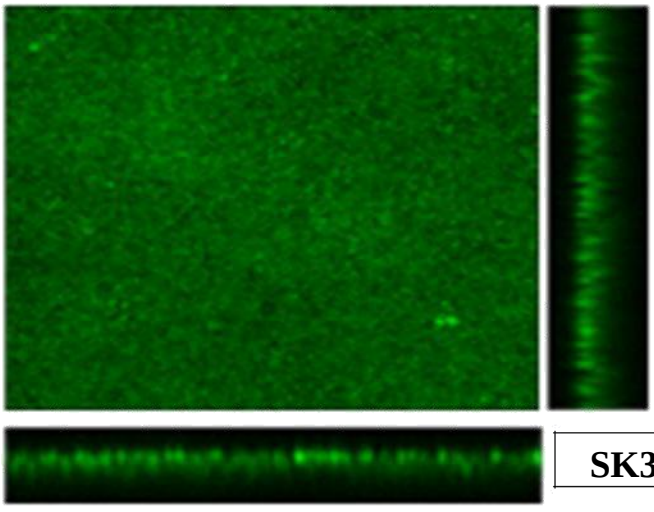
0364-A



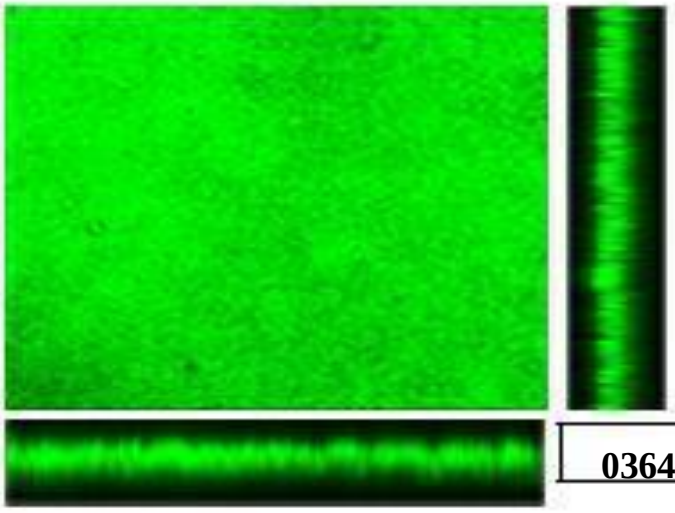
SK36-B



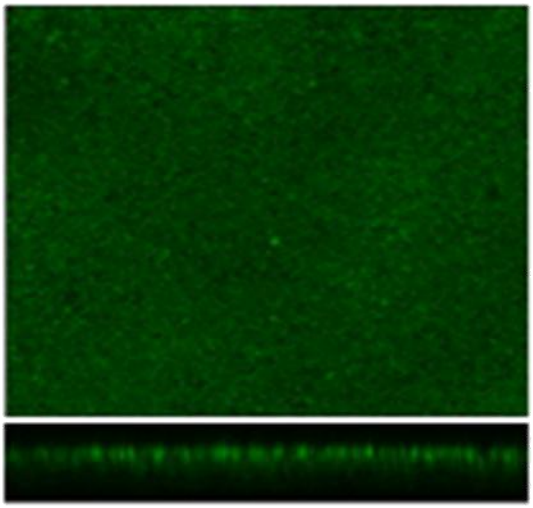
0364-B



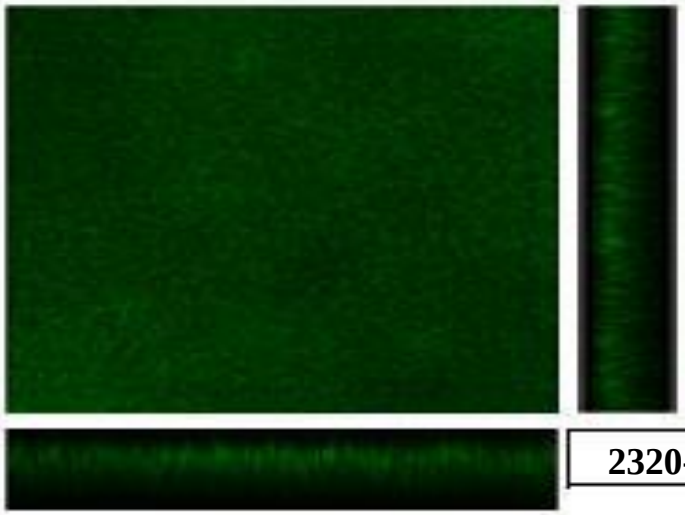
SK36-C



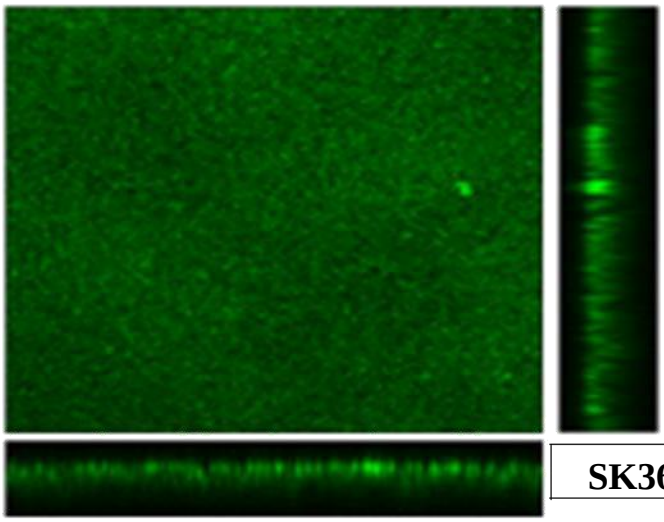
0364-C



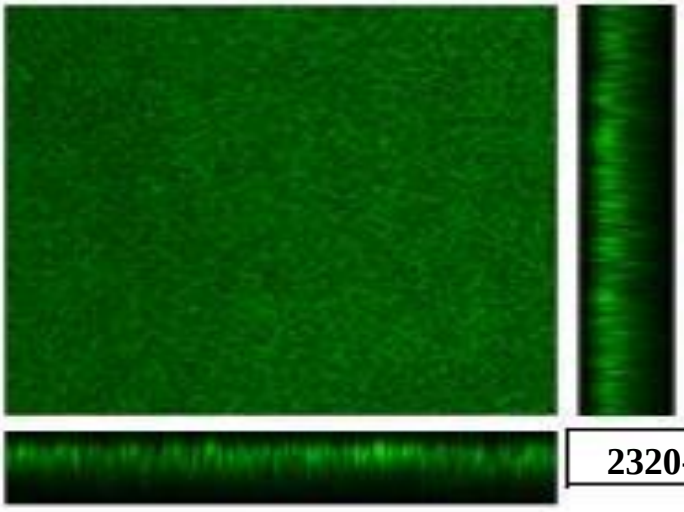
SK36-A



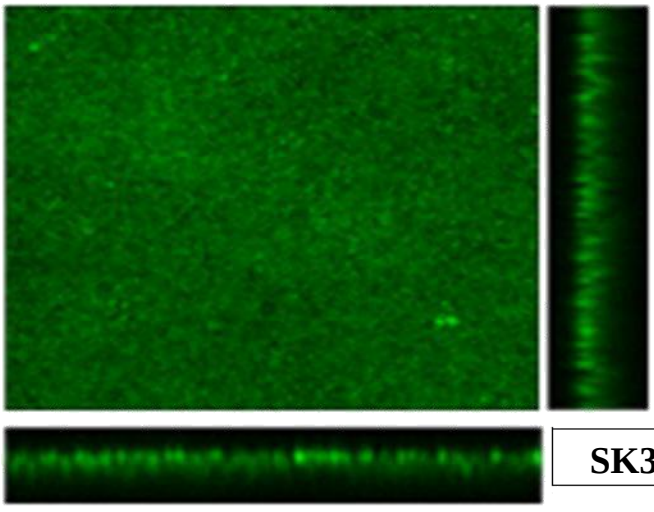
2320-A



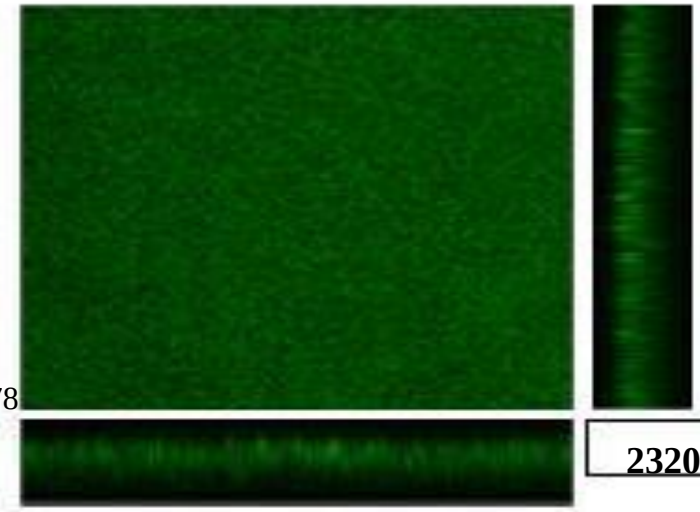
SK36-B



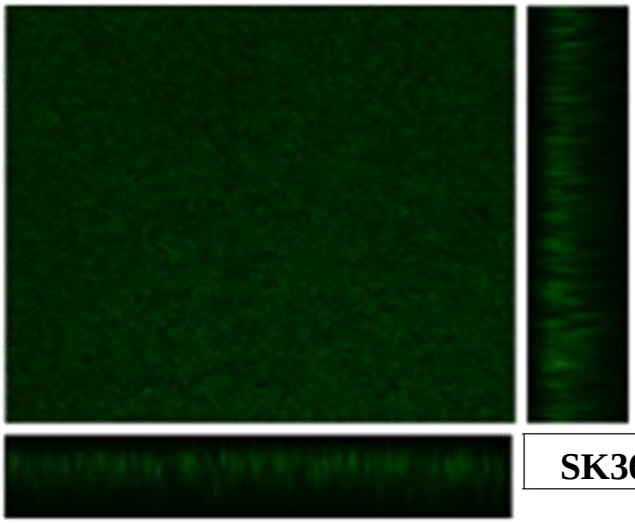
2320-B



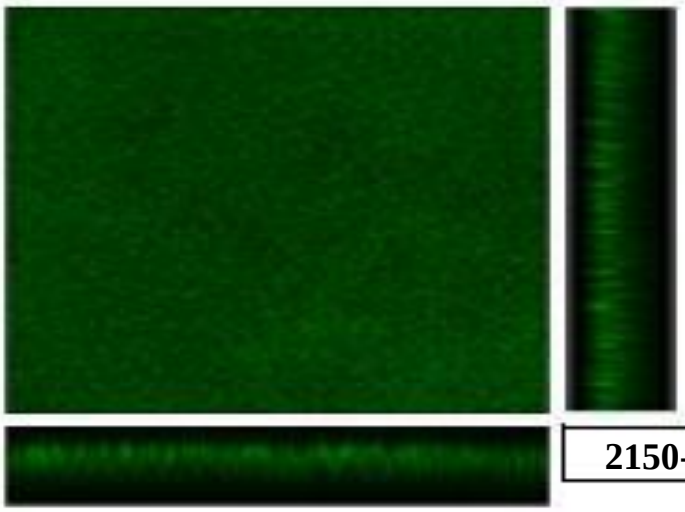
SK36-C



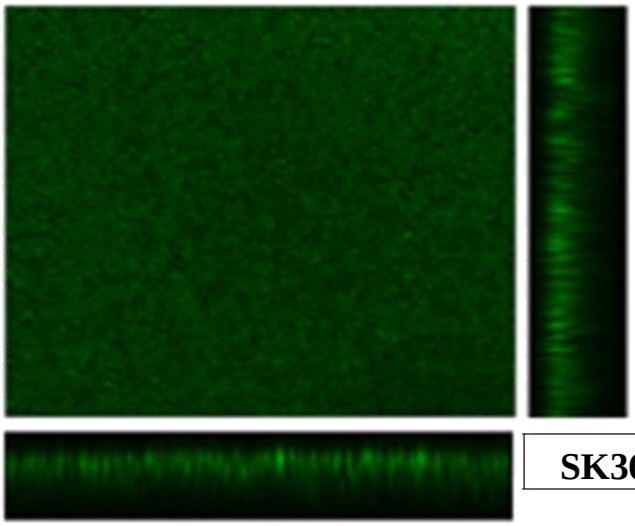
2320-C



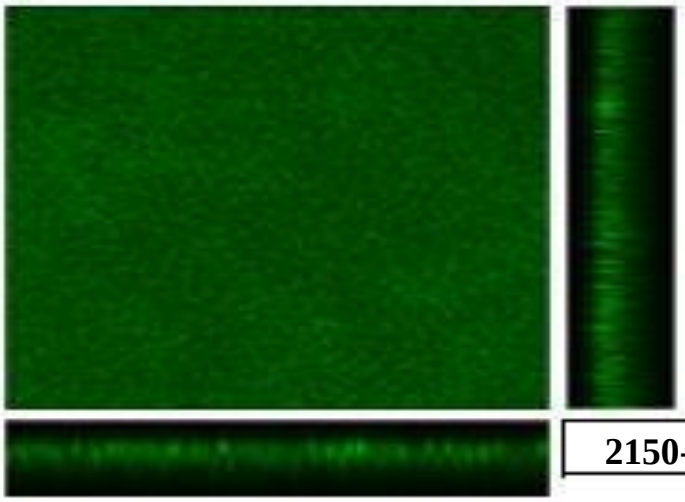
SK36-A



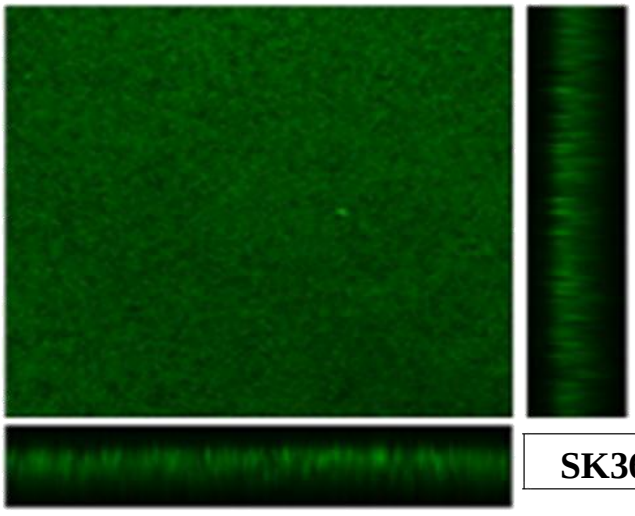
2150-A



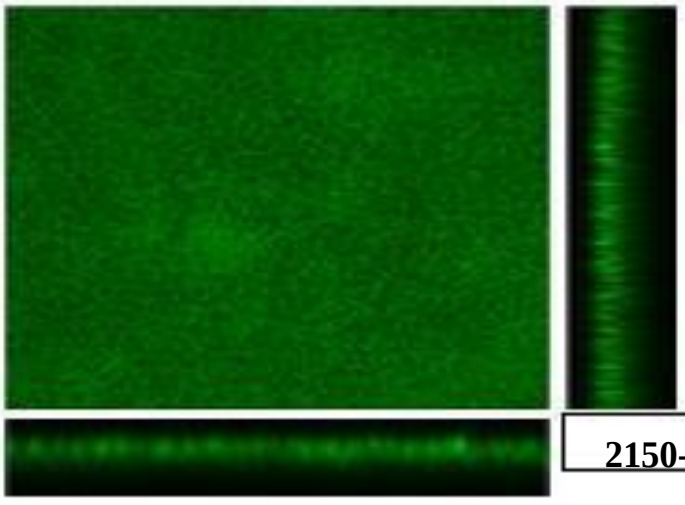
SK36-B



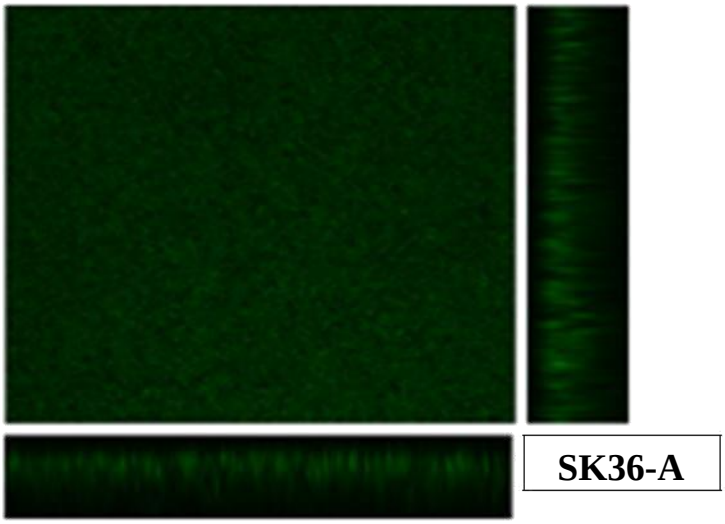
2150-B



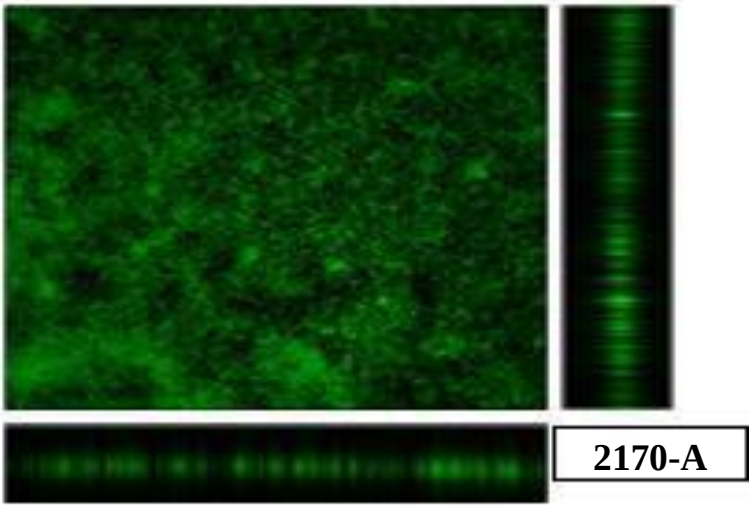
SK36-C



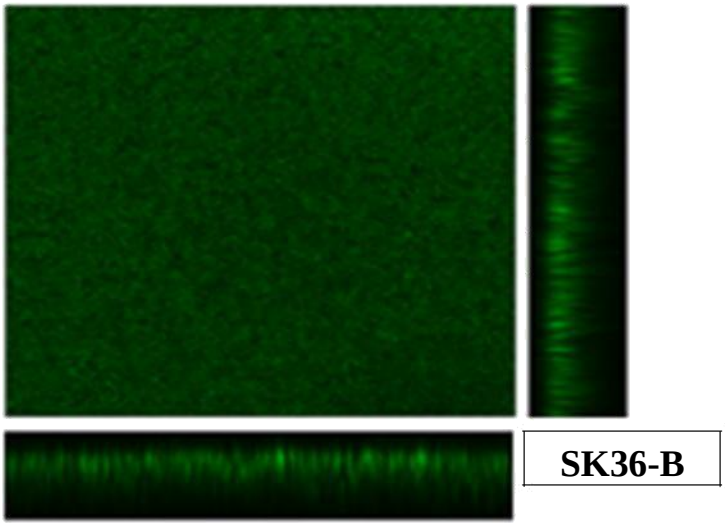
2150-C



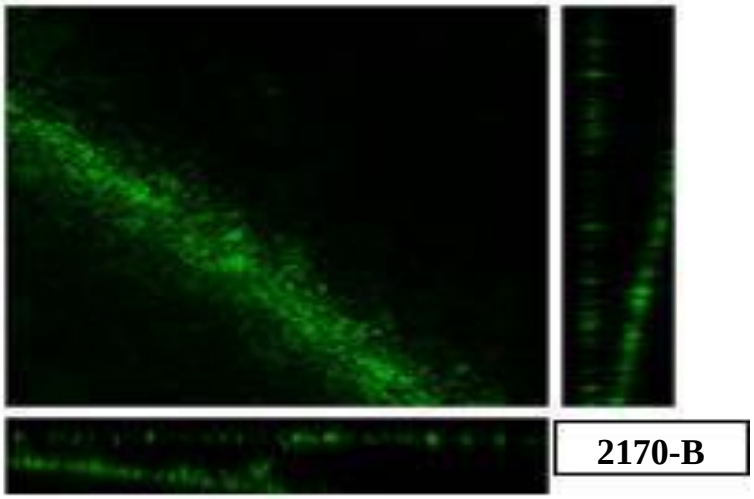
SK36-A



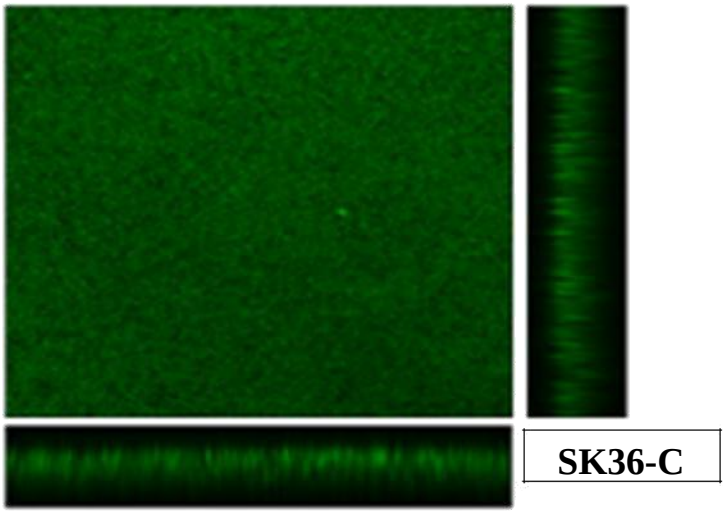
2170-A



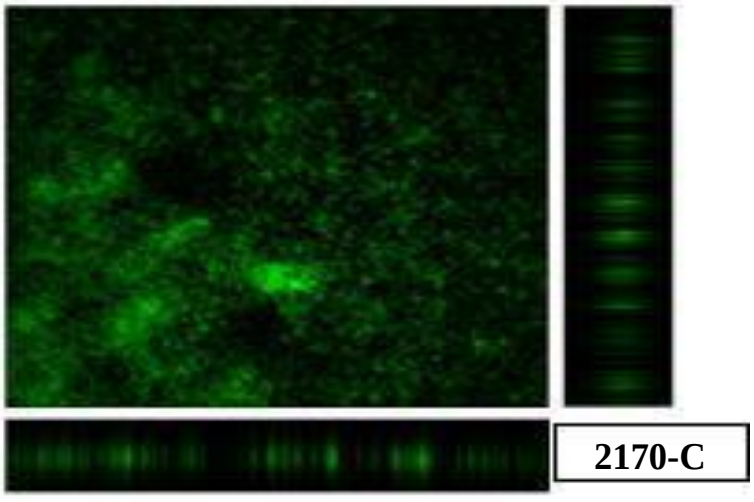
SK36-B



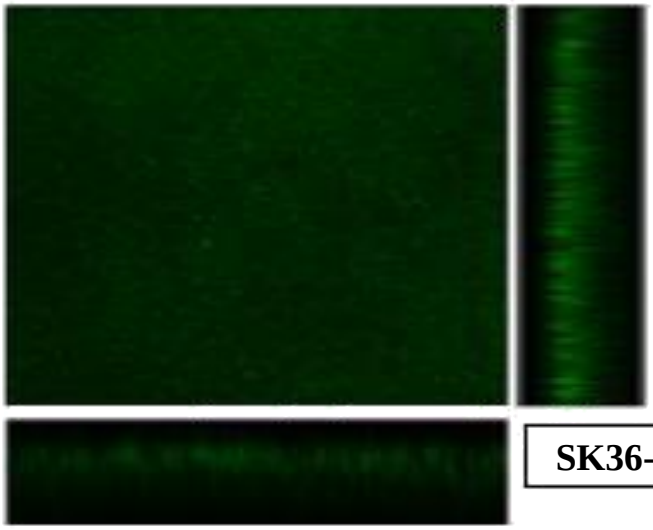
2170-B



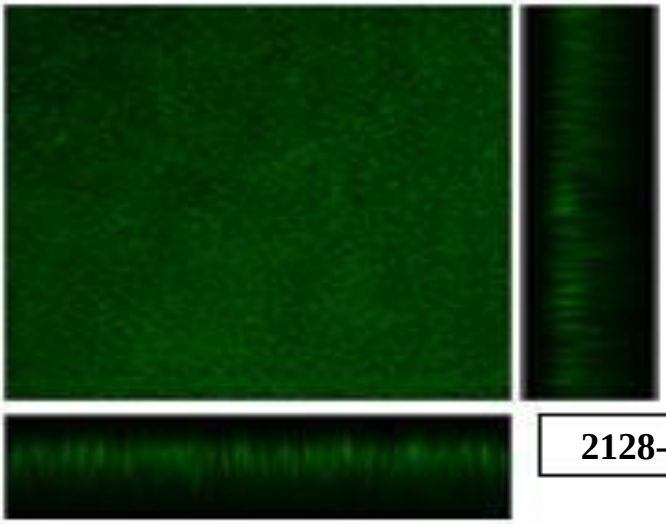
SK36-C



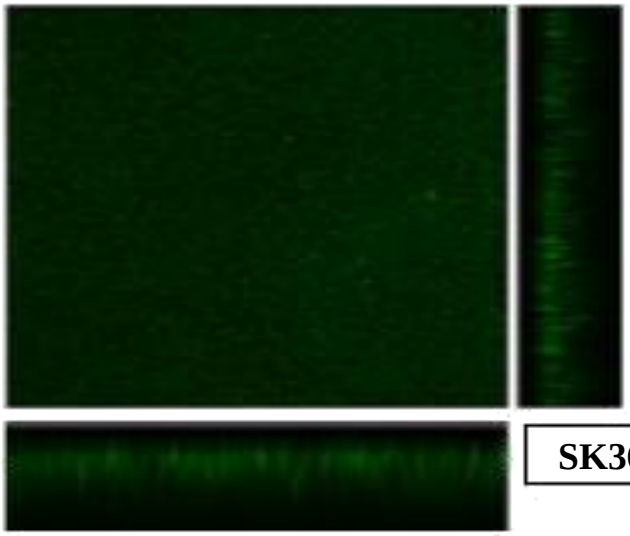
2170-C



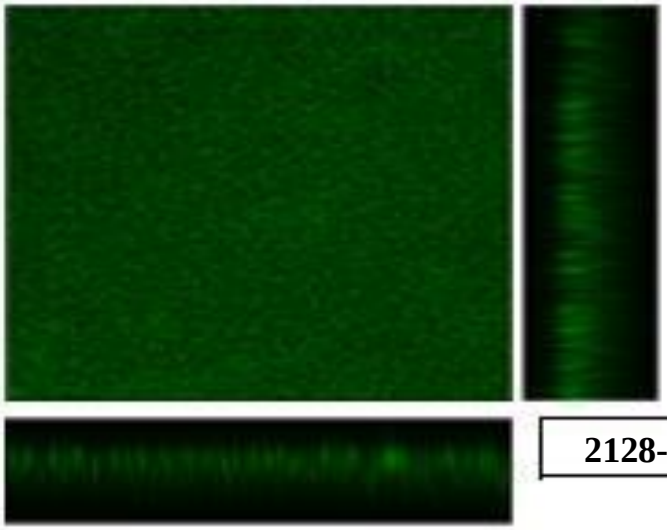
SK36-A



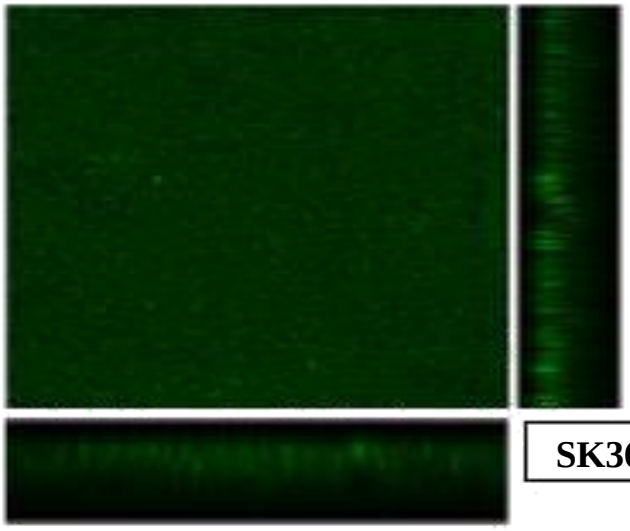
2128-A



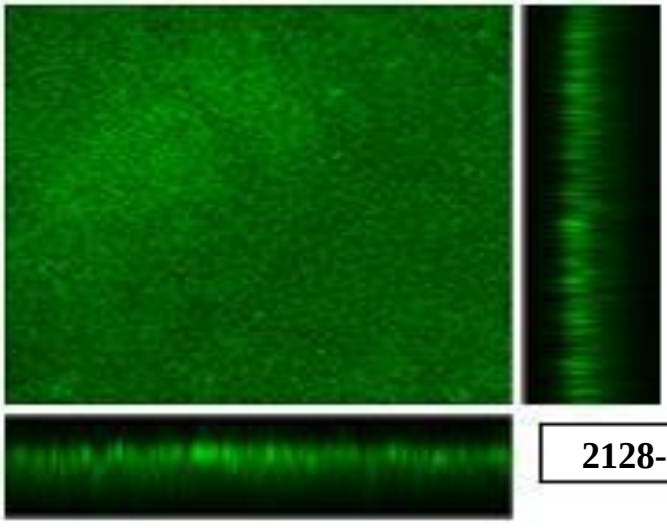
SK36-B



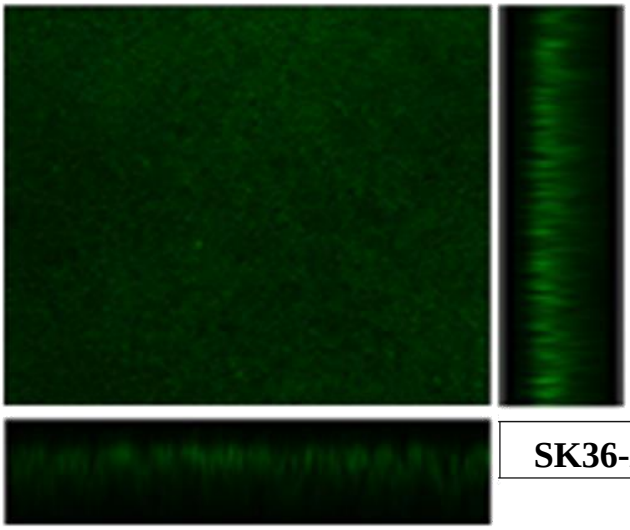
2128-B



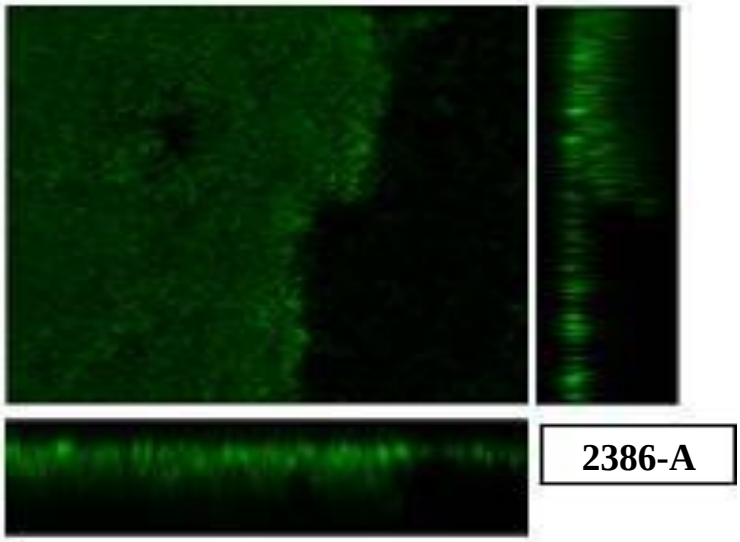
SK36-C



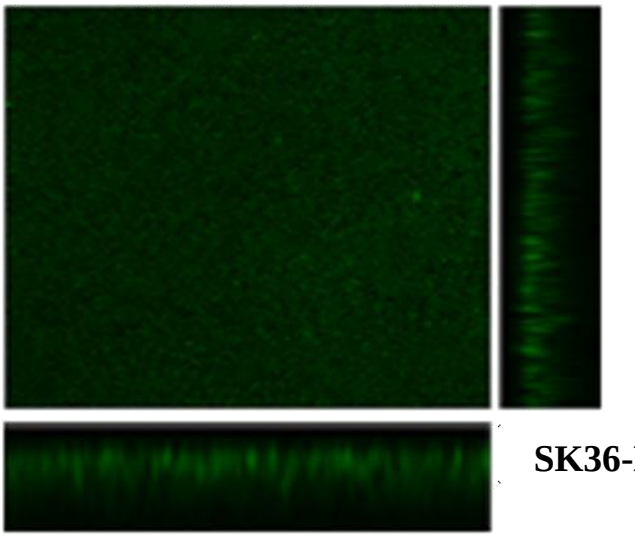
2128-C



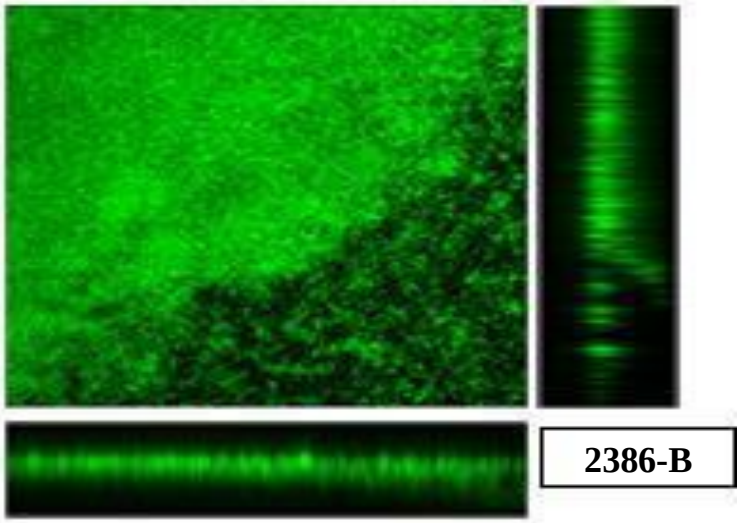
SK36-A



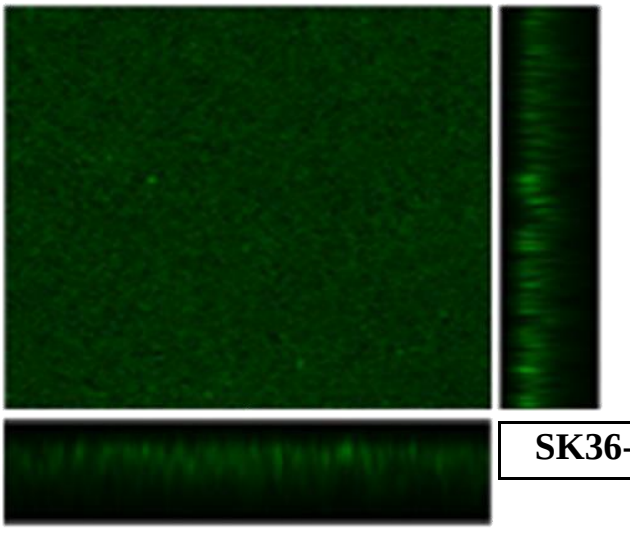
2386-A



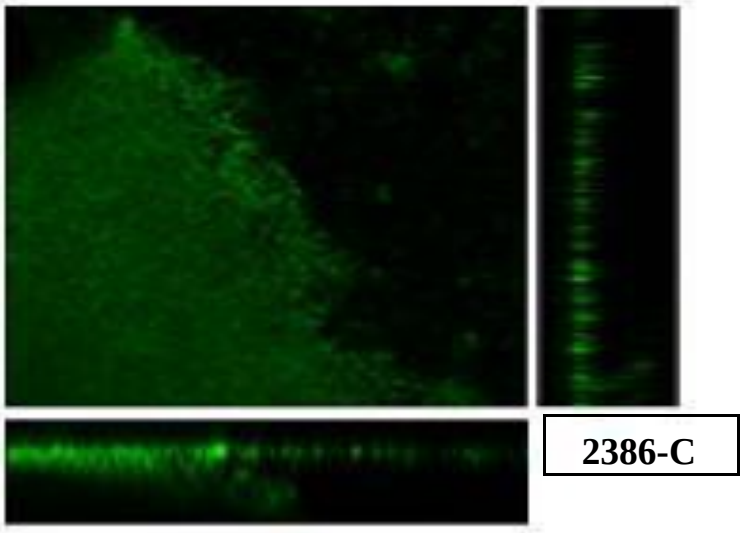
SK36-B



2386-B

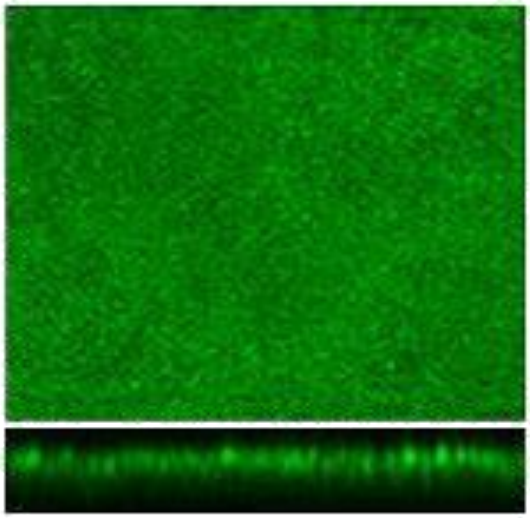


SK36-C

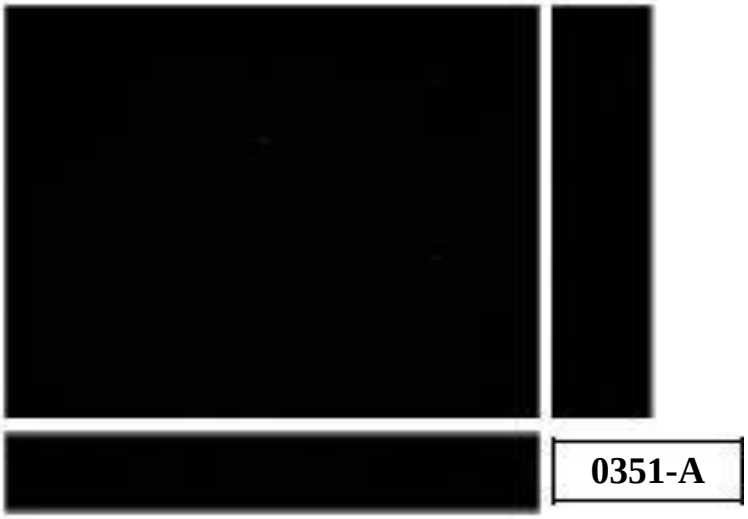


2386-C

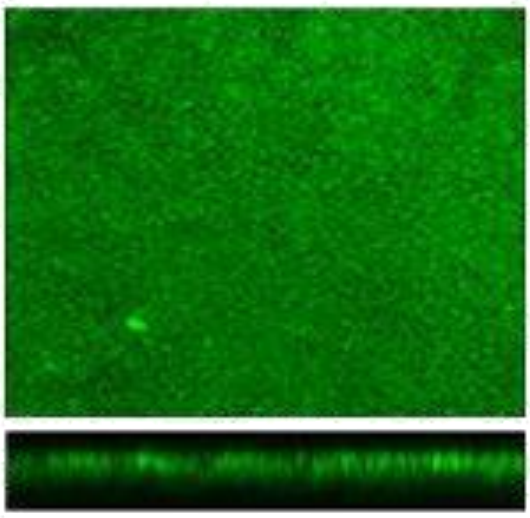
Figure 4. Biofilm imaging by confocal laser scanning microscopy (CLSM) of mutant selected for further study, XG2_0351. This mutant is one of two type I signal peptidases in *S. sanguinis*.



SK36-A

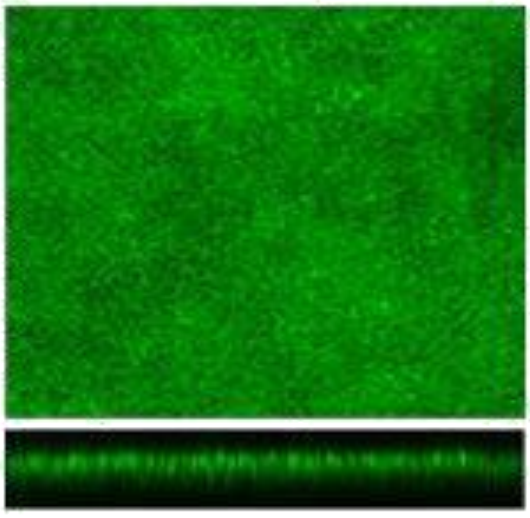


0351-A

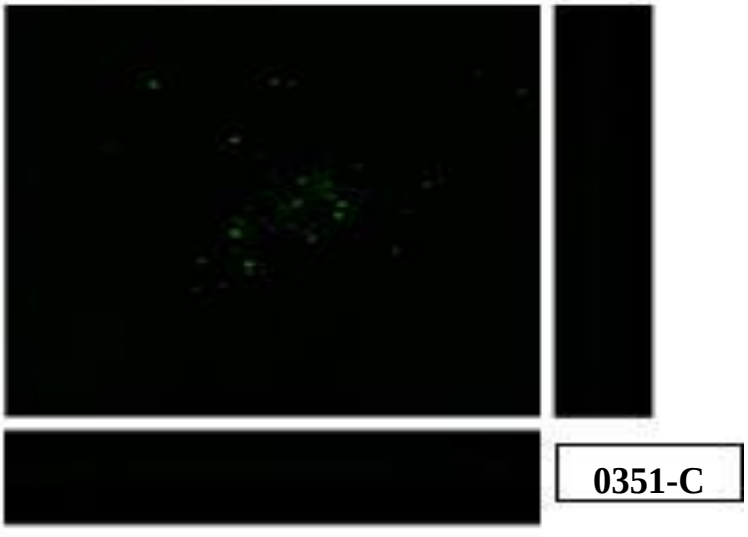


SK36-B

0351-B



SK36-C

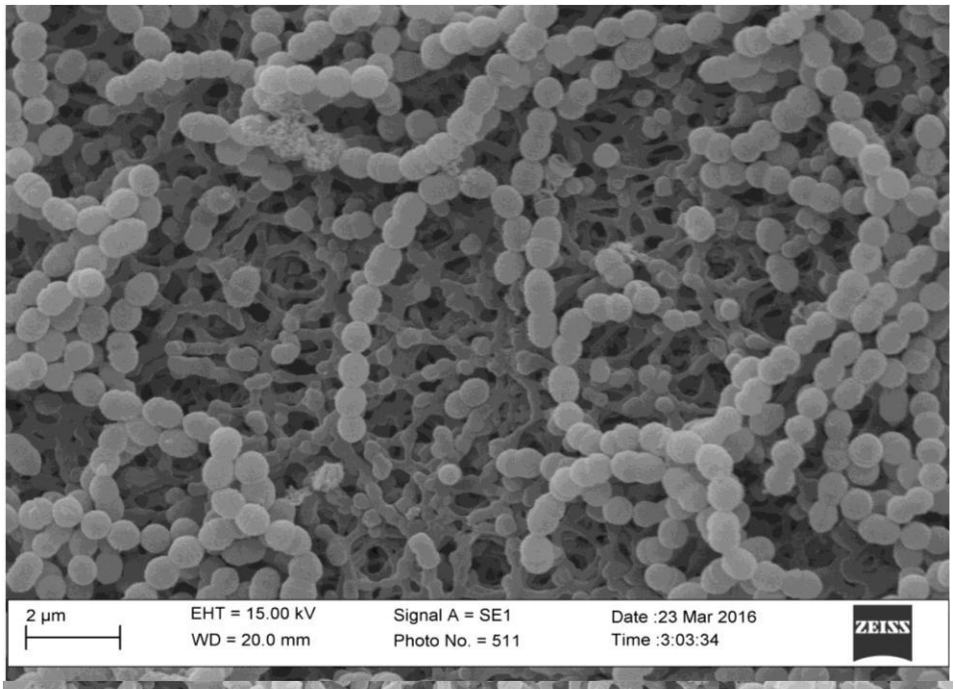


0351-C

Figure 5. Comparison of (A) SK36 and (B) XG2_0351 morphologies using scanning electron microscopy at 600 nm with 15 min intervals.

Figure 5.

A


2 µm
EHT = 15.00 kV
WD = 20.0 mm
Signal A = SE1
Photo No. = 511
Date :23 Mar 2016
Time :3:03:34
ZEISS

B

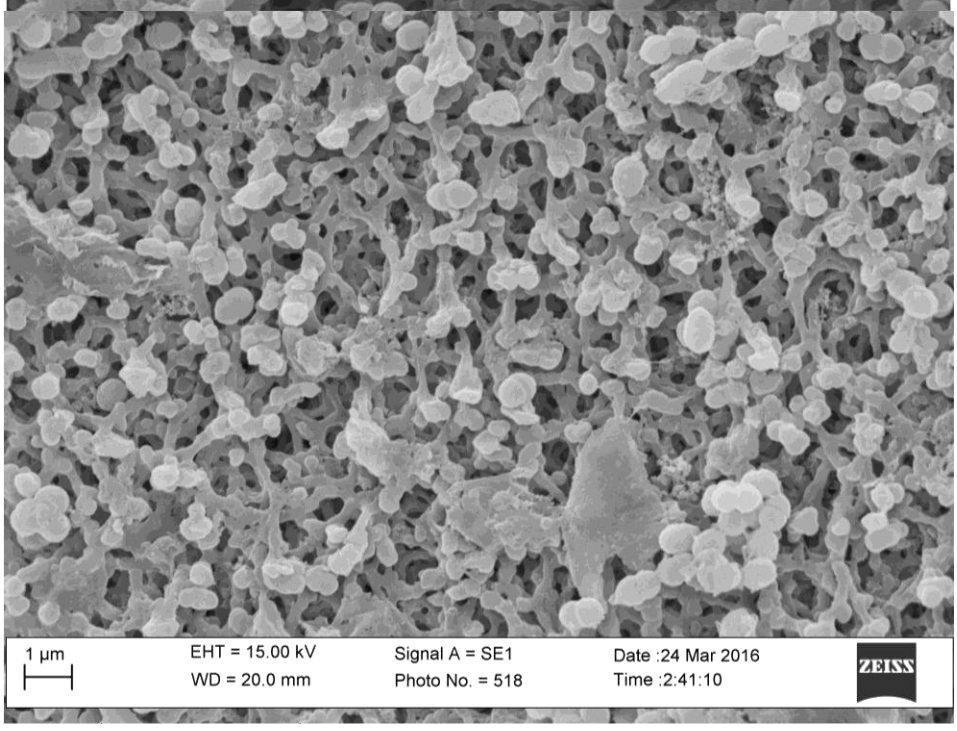
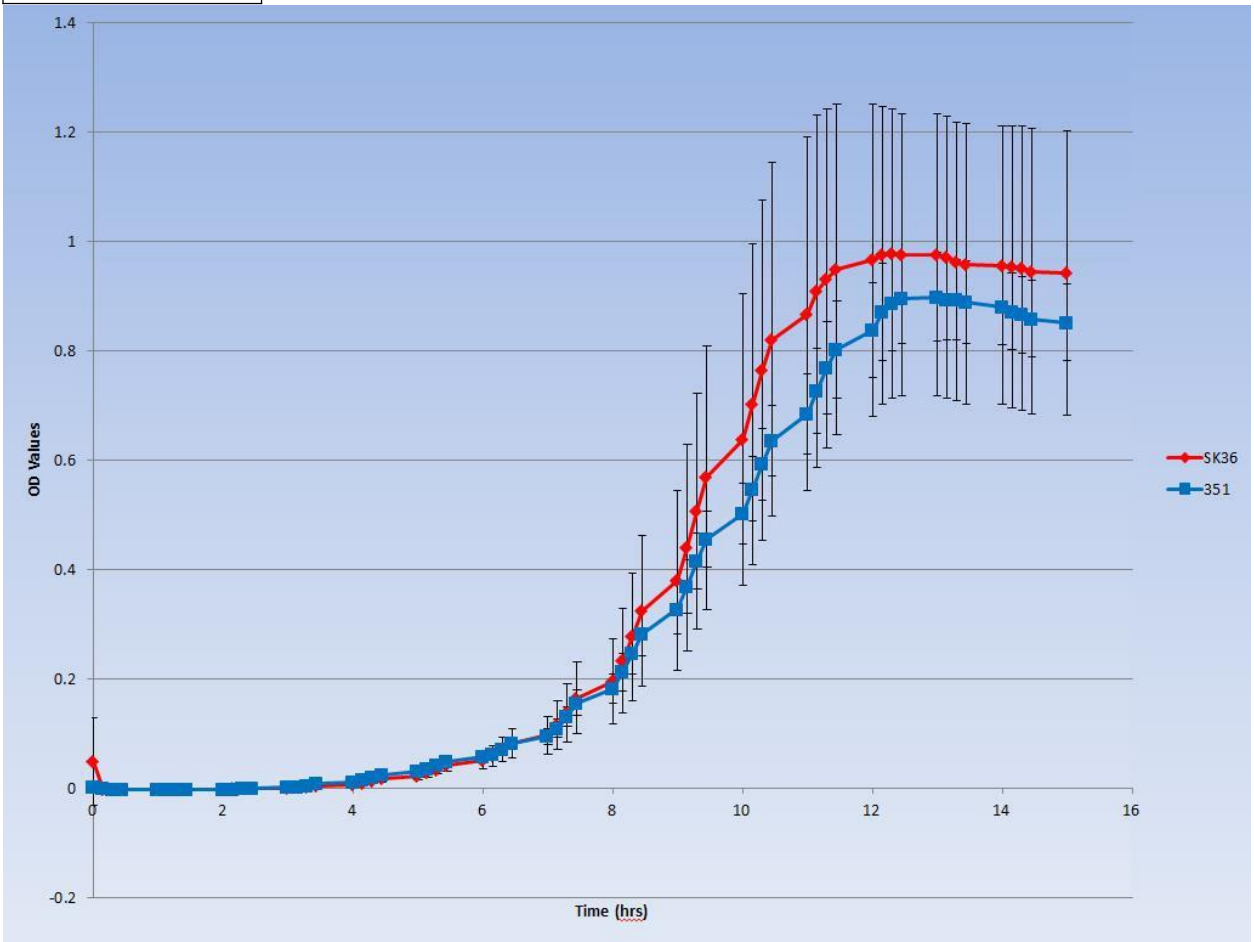

1 µm
EHT = 15.00 kV
WD = 20.0 mm
Signal A = SE1
Photo No. = 518
Date :24 Mar 2016
Time :2:41:10
ZEISS

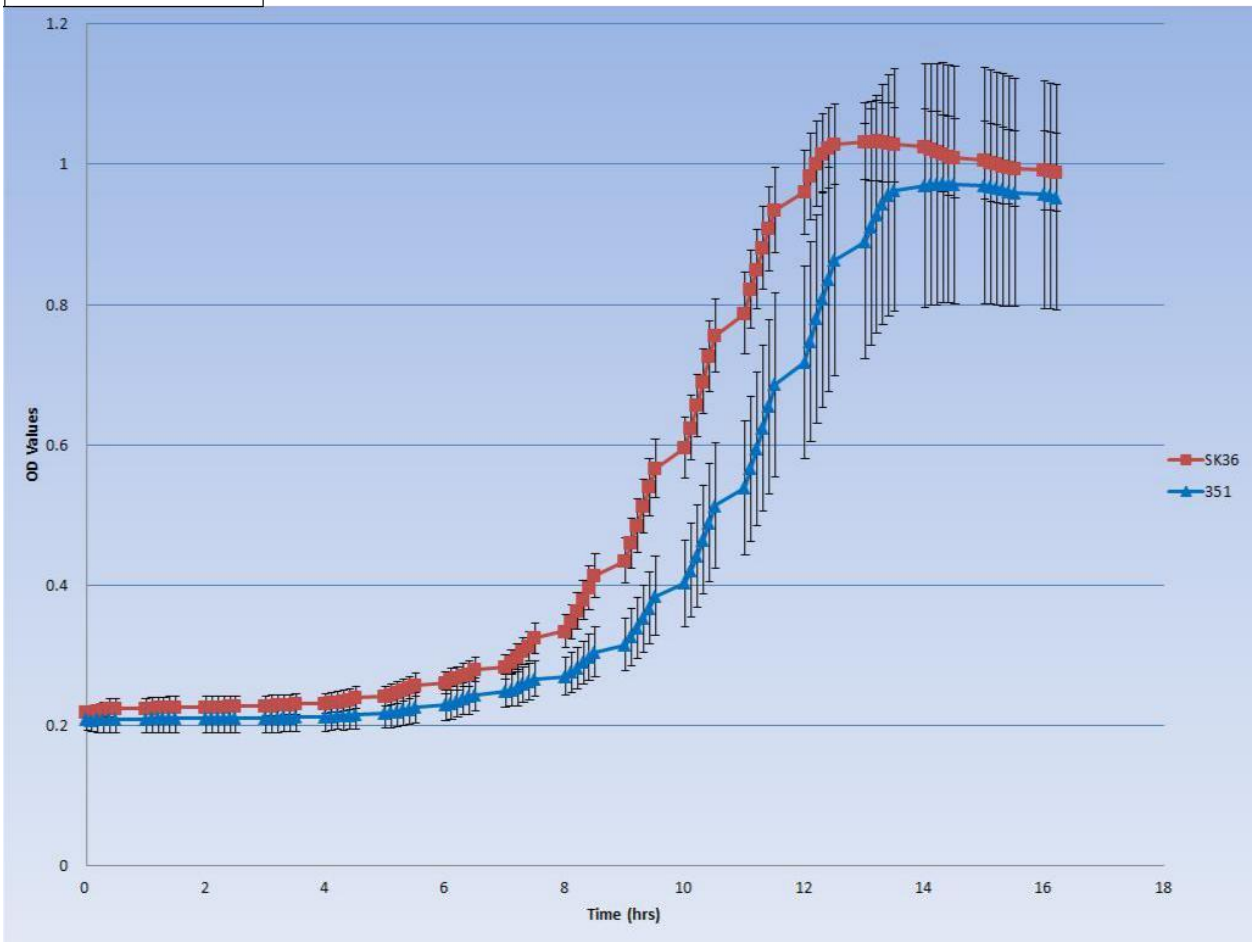
Figure 6. Three trials comparing the growth of SK36 with XG2_0351 using the plate reader.

Figure 6.

Trial 1



Trial 2



Trial 3

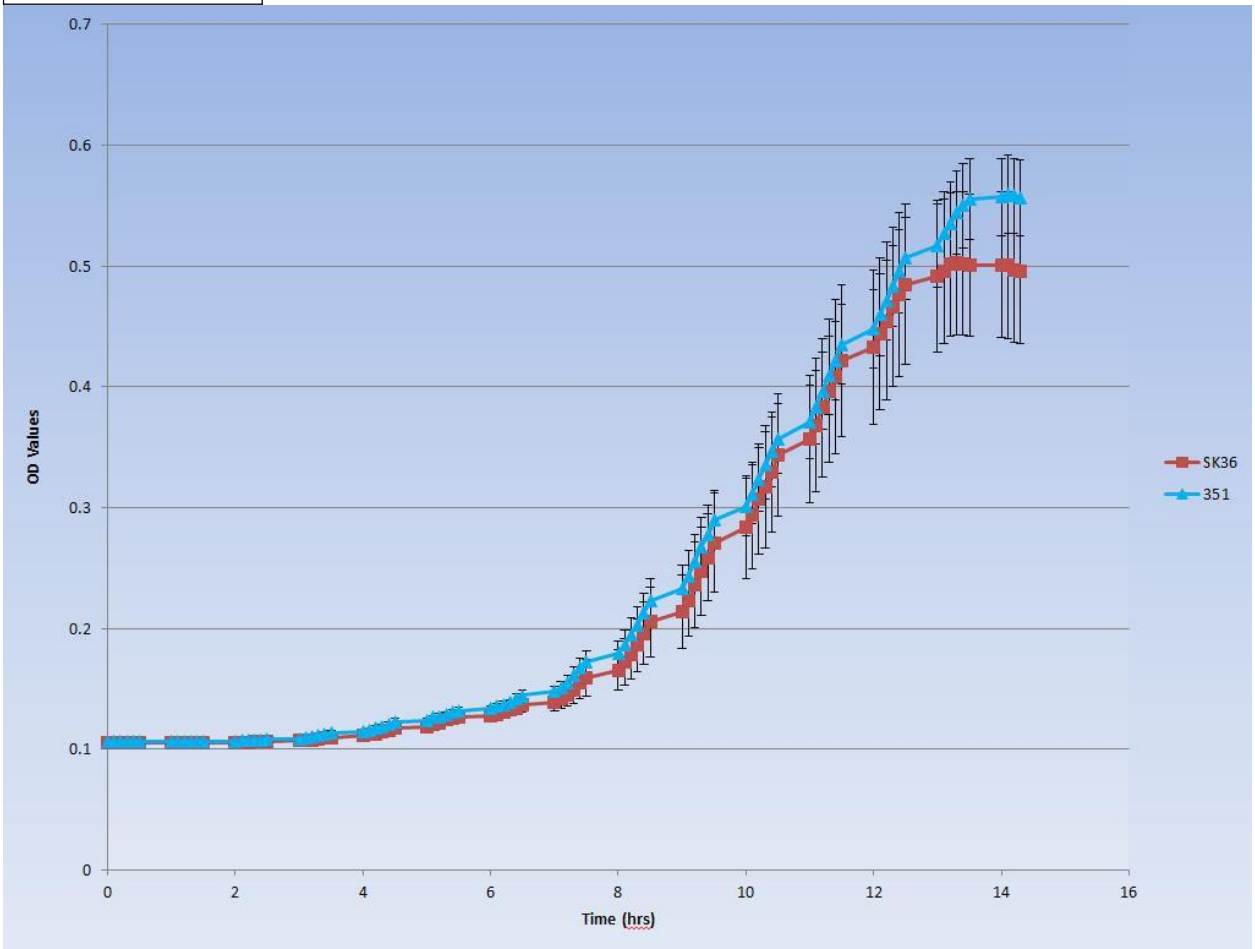
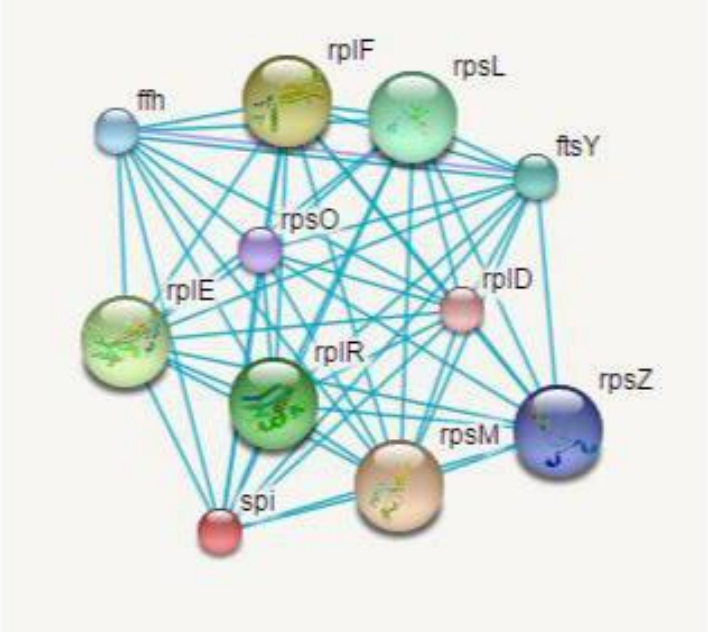

0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
OD Values
0
2
4
6
8
10
12
14
16
Time (hrs)
SK36
351

Figure 7. (A) STRING analysis of protein-protein interactions of XG2_0351 and (B) gene co-occurrence network.

Figure 7.

A



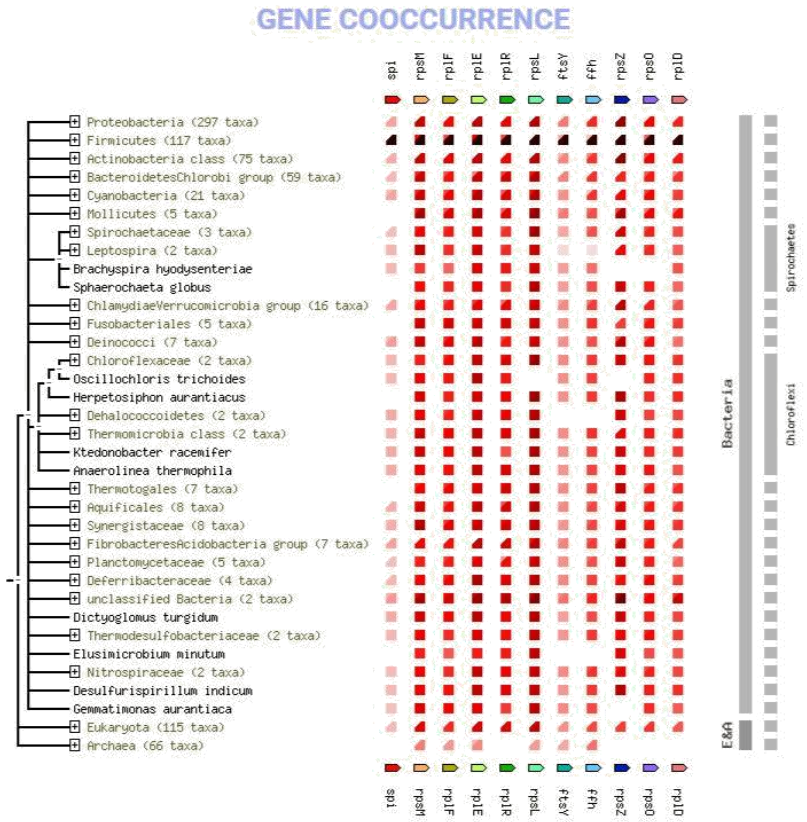
Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
spi	*Signal peptidase I (209 aa)*									

Predicted Functional Partners:

rpsM	*30S ribosomal protein S13; Located at the top of the head of the 30S subunit, it contacts several helices of the 16S rRNA. In the 7...*	x	x	x	x	x	•	x		0.855
rplF	*50S ribosomal protein L6; This protein binds to the 23S rRNA, and is important in its secondary structure. It is located near the su...*	x	x	x	x	x	•	x		0.855
rplE	*50S ribosomal protein L5; This is 1 of the proteins that binds and probably mediates the attachment of the 5S RNA into the large r...*	x	x	x	x	x	•	x		0.851
rplR	*50S ribosomal protein L18; This is one of the proteins that binds and probably mediates the attachment of the 5S RNA into the lar...*	x	x	x	x	x	•	x		0.849
rpsL	*30S ribosomal protein S12; With S4 and S5 plays an important role in translational accuracy (By similarity) (137 aa)*	x	x	x	x	x	•	x		0.823
ftsY	*SRPR, signal recognition particle-docking protein; Involved in targeting and insertion of nascent membrane proteins into the cytop...*	x	x	x	x	x	•	x		0.823
ffh	*SRP54, signal recognition particle GTPase protein; Involved in targeting and insertion of nascent membrane proteins into the cyto...*	x	x	x	x	x	•	x		0.823
rpsZ	*30S ribosomal protein S14; Binds 16S rRNA, required for the assembly of 30S particles and may also be responsible for determini...*	x	x	x	x	x	•	x		0.821
rpsO	*30S ribosomal protein S15; One of the primary rRNA binding proteins, it binds directly to 16S rRNA where it helps nucleate assem...*	x	x	x	x	x	•	x		0.815
rplD	*50S ribosomal protein L4; One of the primary rRNA binding proteins, this protein initially binds near the 5'-end of the 23S rRNA. It i...*	x	x	x	x	x	•	x		0.815

B



Similarity Scale

no similarity detectable — *100% sequence conservation.*

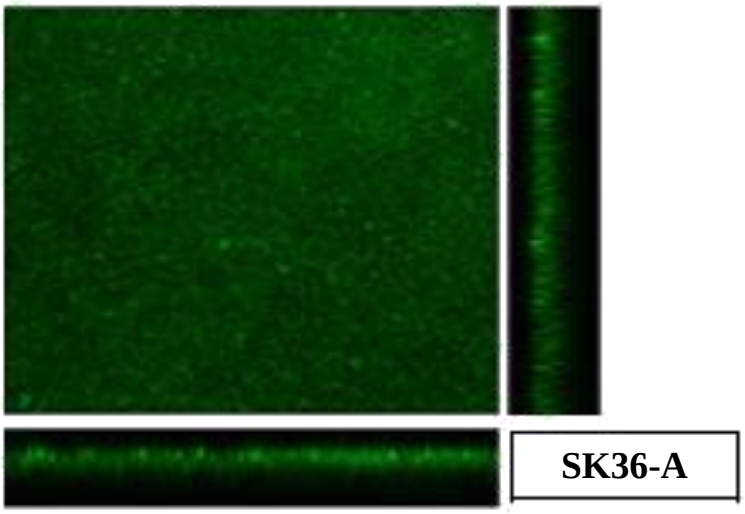
the color denotes, for each gene of interest, the similarity of its best hit in a given STRING genome. Similarities in these presence/absence profiles can predict interactions.

Clade Coverage

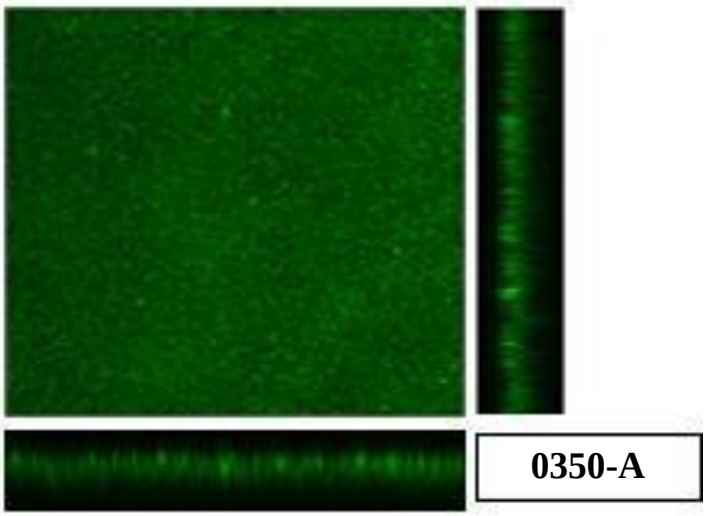
for groups of genomes that are collapsed in the phylogenetic tree, two distinct colors indicate the lowest and highest similarity observed within that clade.

lowest similarity — *highest similarity*

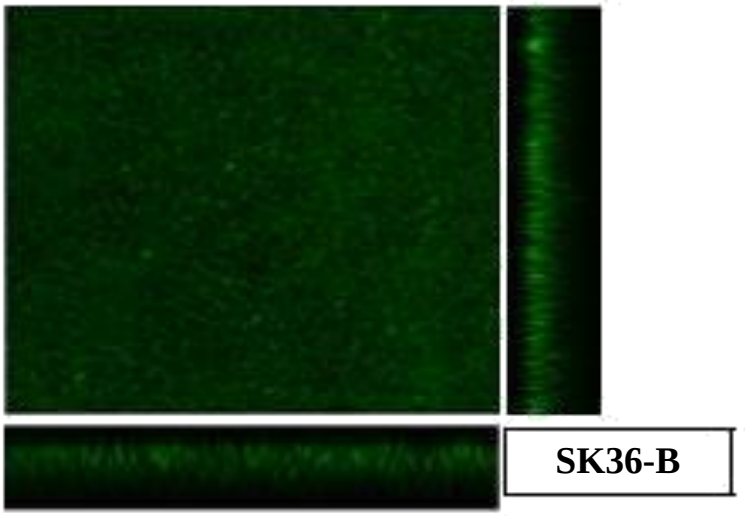
Figure 8. Biofilm imaging using confocal laser scanning microscopy (CLSM) of downstream mutant, XG2_0350.



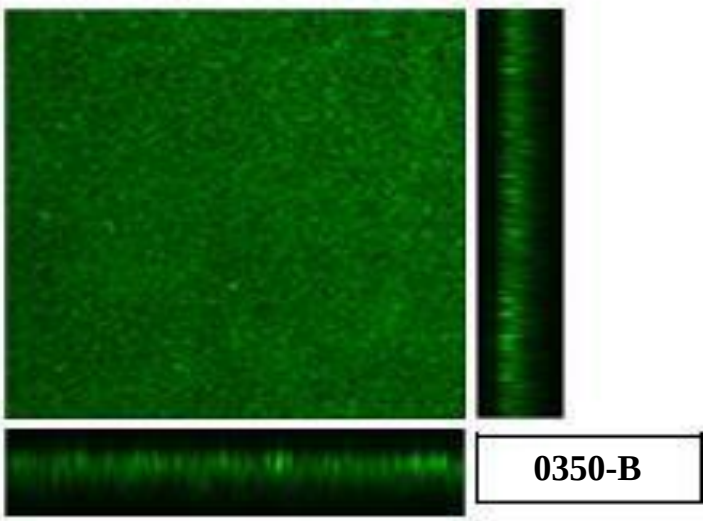
SK36-A



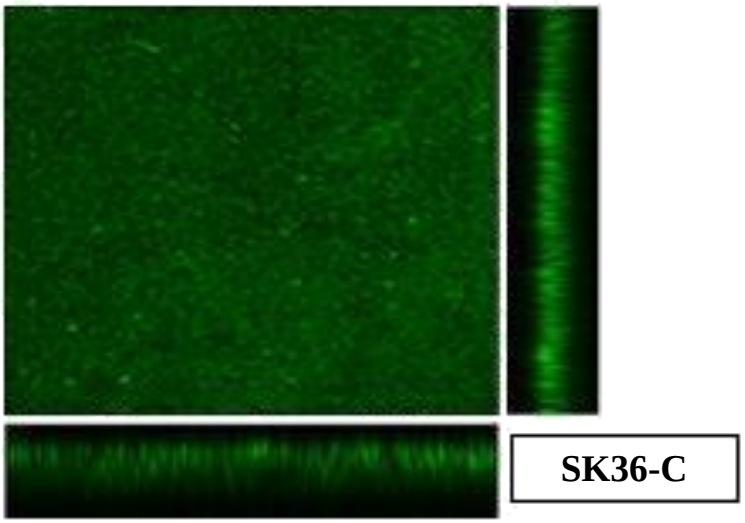
0350-A



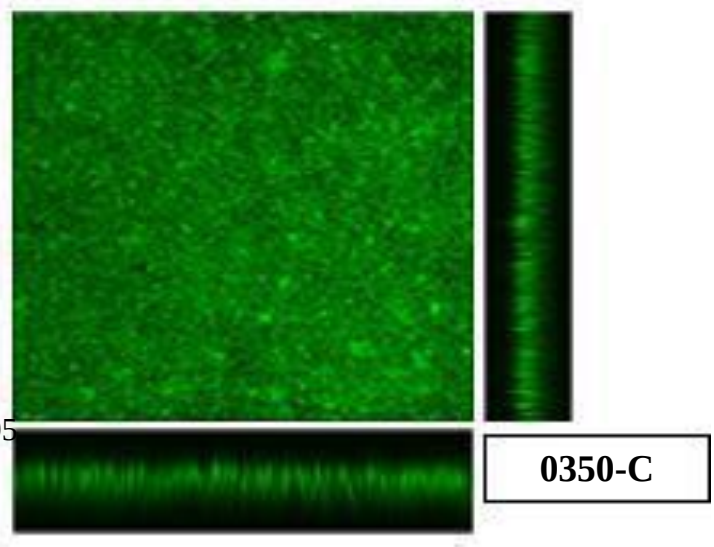
SK36-B



0350-B

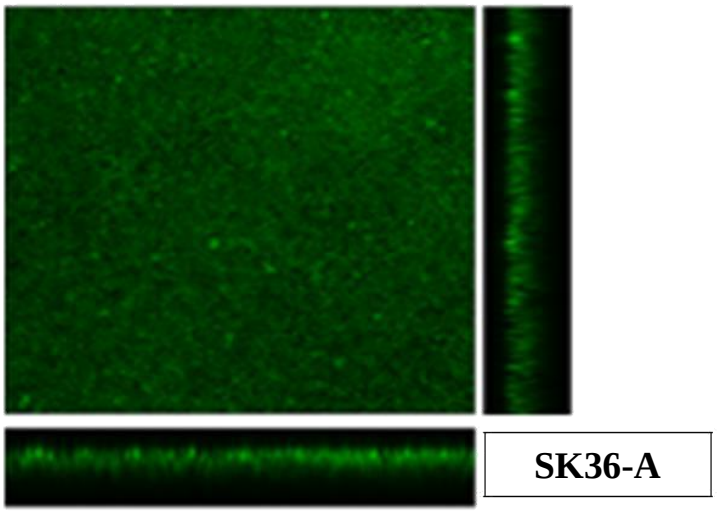


SK36-C

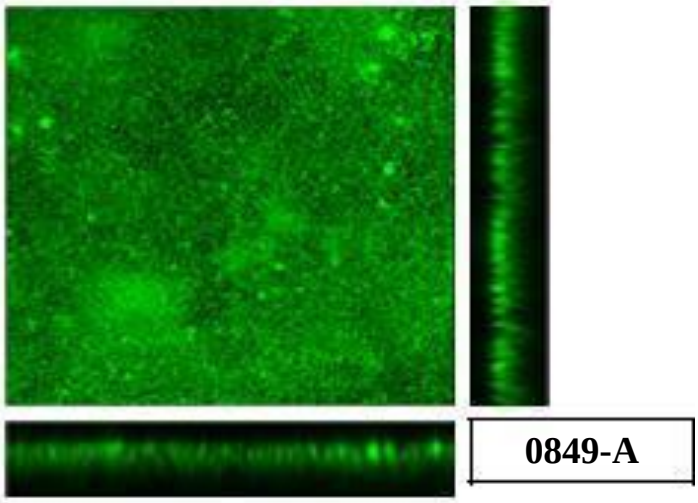


0350-C

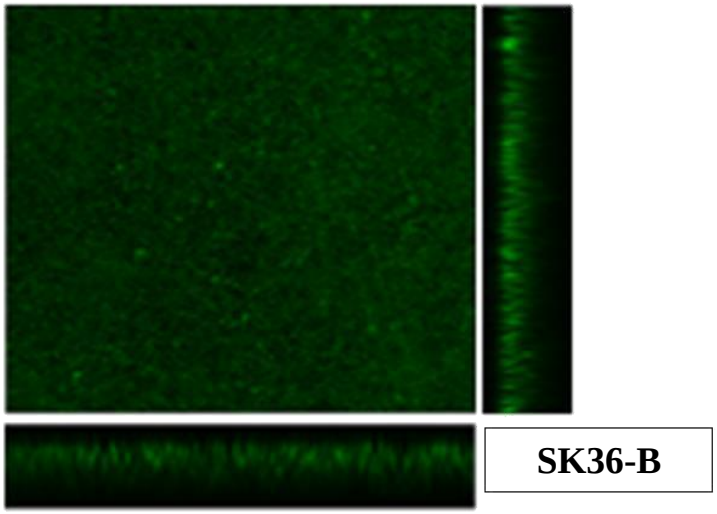
Figure 9. Biofilm imaging using confocal laser scanning microscopy (CLSM) of SPase I paralog, XG2_0849.



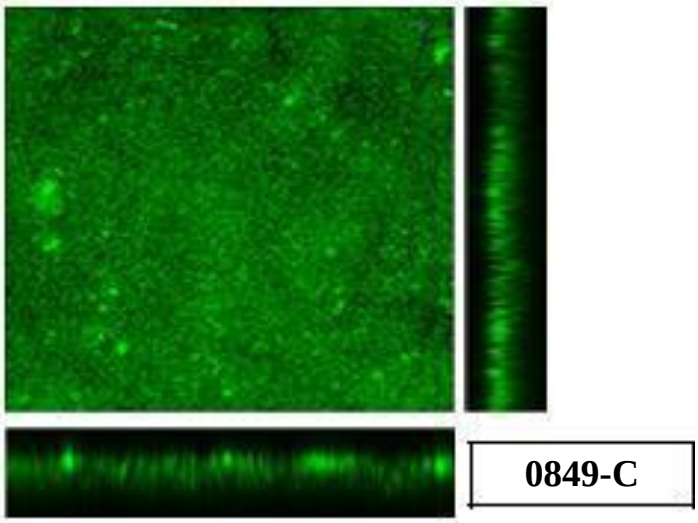
SK36-A



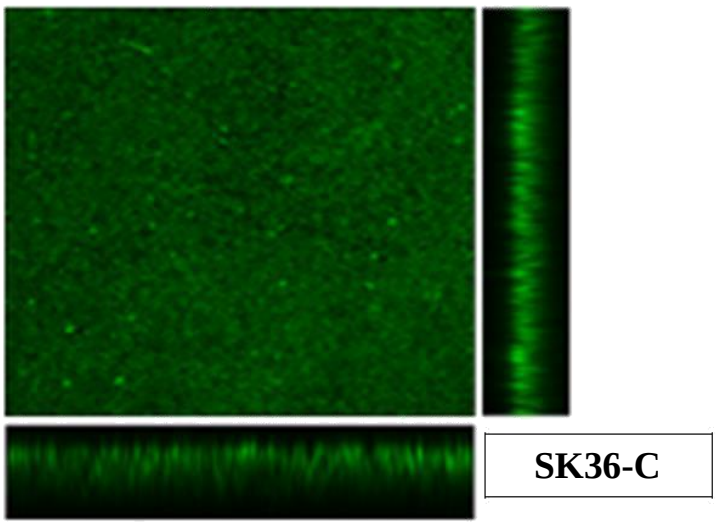
0849-A



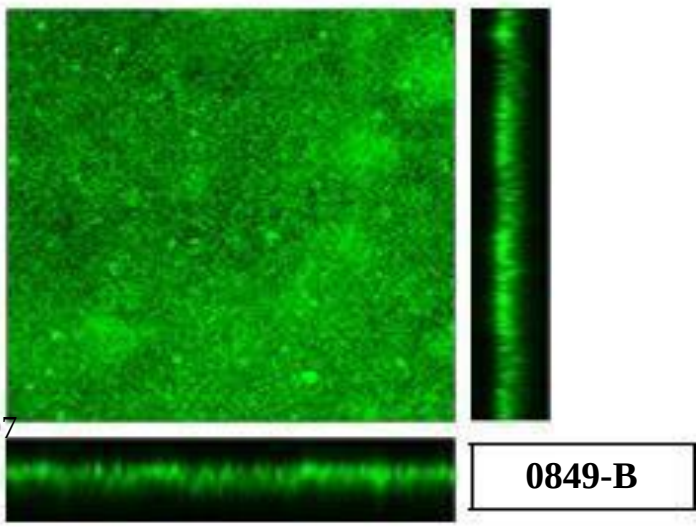
SK36-B



0849-C



SK36-C



0849-B

Figure 10. Quantitative data for confocal images. Relative intensity of biofilm formation was measured for SK36, XG2_0351, XG2_0350, and XG2_0849.

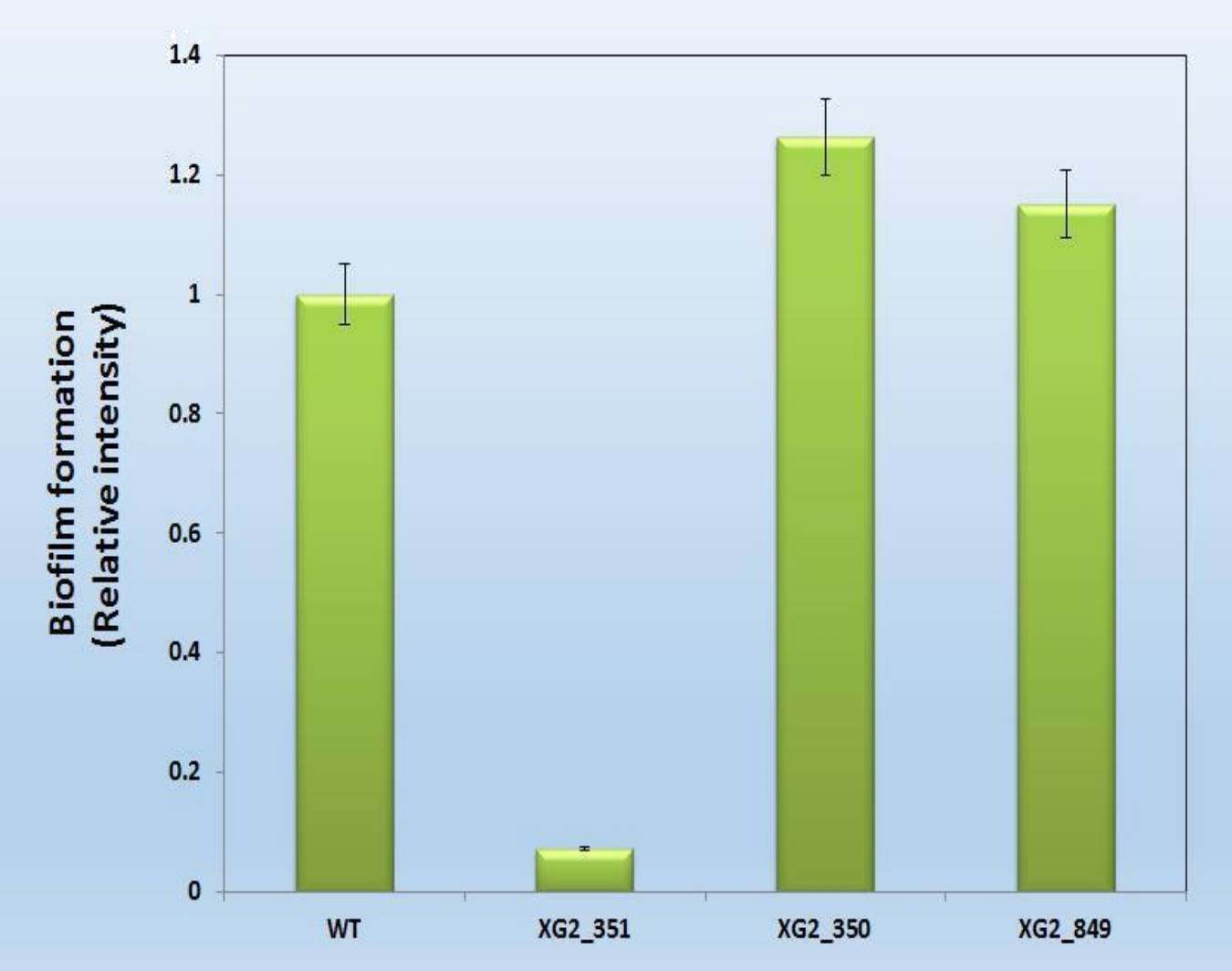
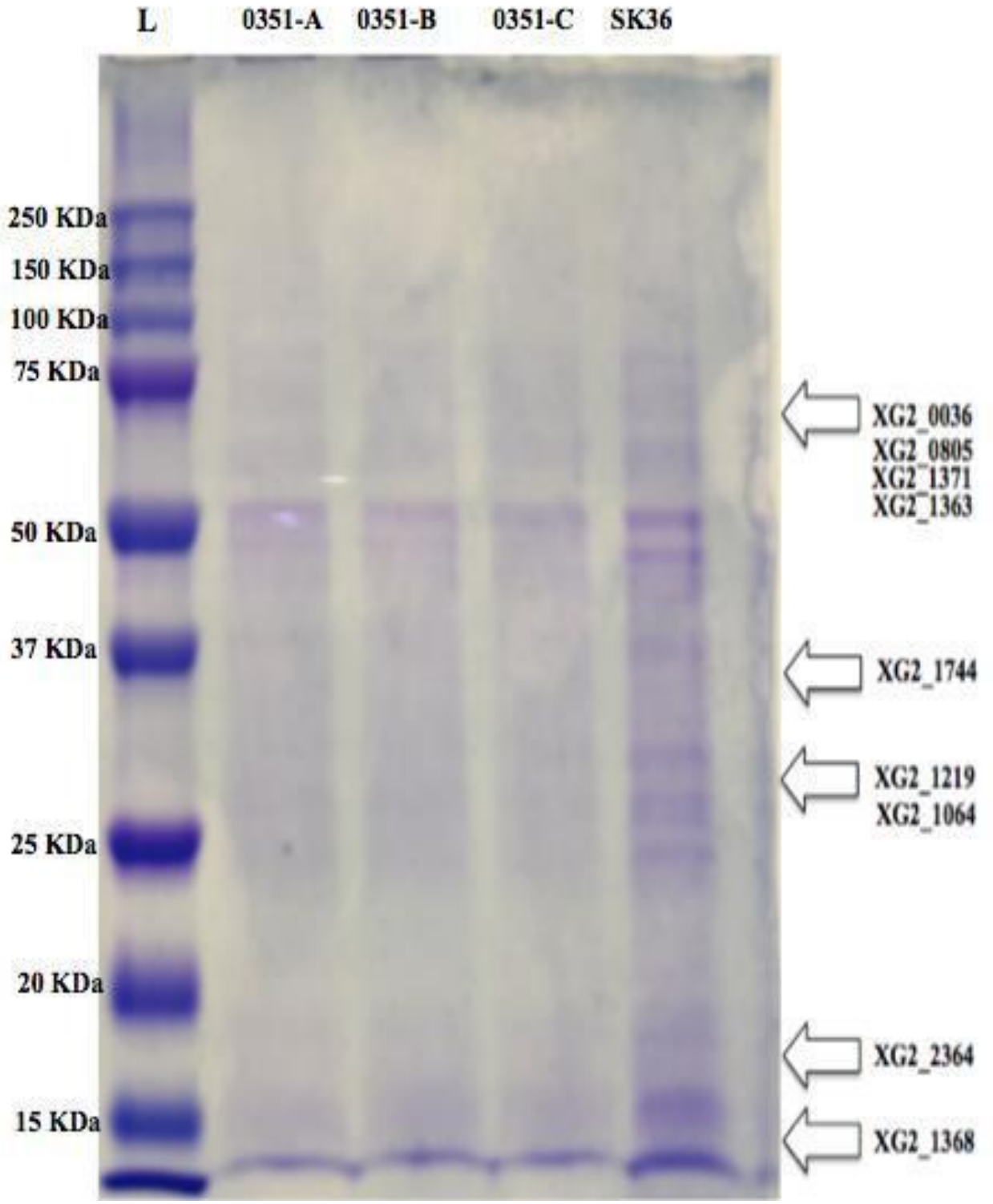

Biofilm formation
(Relative intensity)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
WT
XG2_351
XG2_350
XG2_849

Figure 11. Whole protein extraction from bacterial strains SK36 and mutant sample were separated by SDS-PAGE and stained with coomassie blue. The arrows indicate possible proteins that have been shown to reduce biofilm formation by confocal imaging.

Figure 11.


L
0351-A
0351-B
0351-C
SK36
250 KDa
150 KDa
100 KDa
75 KDa
50 KDa
37 KDa
25 KDa
20 KDa
15 KDa
XG2_0036
XG2_0805
XG2_1371
XG2_1363
XG2_1744
XG2_1219
XG2_1064
XG2_2364
XG2_1368

DISCUSSION

Oral biofilm formation in streptococci has shown to be involved in a variety of microbial infections in the human body, through recruiting diverse bacterial species to site of infection and displaying an effective defense system against host immune defenses (Hall, McGillicuddy et al. 2014). Biofilm formation involves numerous stages, namely attachment, maturation, and dispersion (Lister, Horswill 2014, Foster, Geoghegan et al. 2014). *Streptococcus sanguinis* has been shown to be involved in biofilm formation (Xu, Alves et al. 2007, Kolenbrander, London 1993). Investigating the *S. sanguinis* genes involved in biofilm formation will be indispensable in uncovering potential drug targets against diverse bacterial infections that involve biofilm formation in the oral cavity.

A set of 51 non-essential genes (Table 2) was screened previously in our lab for the ability to affect biofilm formation using microtiter assay as described by O'Toole (O'Toole 2011). This constituted the starting point for the project of studying *S. sanguinis* genes involved in biofilm formation. Bioinformatically, we identified for every biofilm related gene functions using clusters of orthologous groups (COGs) annotations (Table 3), as described by Uniprot, as well as operon (and genetic neighbors) and the presence/absence of paralogues (using BLAST with identity >70% and E-value cutoff 10^{-7}) (Table 3). On the wet lab level, we ran a microtiter assay with crystal violet (CV) staining to confirm the previous findings, which concurred to a high degree to the previous findings. Out of 51 mutants, 25 mutants that exhibited a reduced biofilm formation as shown by CV staining (Figure 2) were further investigated using confocal laser scanning microscopy (CLSM). The CLSM images, obtained in triplicates from every mutant, showed variant patterns in biofilm reduction (Figure 3).

Mutant XG2_0351 was selected for further study in this project based upon results that were obtained through crystal violet staining and CLSM images (Figure 4). The biofilm formation of XG2_0351 showed significant reduction in biofilm formation, as compared to SK36 through CV staining, which was later

conclusively ascertained through CLSM. Contrasting greatly with the biofilm formed by SK36, there was essentially no biofilm formation observed in either experiment when the XG2_0351 gene was knocked out. Downstream mutant, XG2_0350, was visualized using CLSM to eliminate the possibility of the polar effect (Figure 8). Consequently, XG2_0351, one of two signal peptidase I (SPase I) enzymes in *S. sanguinis*, was selected for investigation in this project because of considerable phenotypic differences from SK36, in terms of biofilm formation, and also because little information is known about the role that type I signal peptidase plays in biofilm formation in this bacterium. Furthermore, even less is known about how SPase I affects biofilm in a bacterium with multiple SPase I enzymes.

To further investigate the phenotypic role of XG2_0351 gene on biofilm formation, we used SEM to visualize the morphological differences in growth between the SK36 and SPase I mutant, XG2_0351. The images (Figure 5) show that the XG2_0351 chains are shorter in length and remain stunted in growth in comparison to the long chains formed by the wild type. In bacteria, the cell wall bears the stress and helps maintain the shape, and is important for cell viability. The scaffold of the cell wall consists of the cross-linker polymer peptidoglycan. Studies have demonstrated that there is a relationship between peptidoglycan synthesis, bacterial growth, and cell shape (Scheffers, Pinho 2005). Mutants that lack one or several enzymes involved in the synthesis of peptidoglycan or other cell wall components display changes in cell shape. Molecular analyses of another member of viridans, *S. gordonii*, showed that some genes required for biofilm formation are involved in peptidoglycan biosynthesis (Loo, Corliss et al. 2000). Therefore, we can conclude that in SK36 the SPase I is necessary to cleave proteins which are crucial to maintain cell shape and cell wall, possibly through peptidoglycan synthesis.

To gain additional information about the growth disparities between SK36 and XG2_0351, comparison of growth curves between wild type and XG2_0351 using a plate reader (Figure 6). The exponential (logarithmic) portions of the resulting growth curves are useful for determining growth rates. Although the SEM

pictures depict phenotypic differences in morphology between SK36 and mutant, the growth curves show no substantial difference in rate of growth. The OD values remain comparable for both which means that this SPase I has demonstrated no substantial role in cell growth. In many bacteria that have been analyzed so far, type I signal peptidase has been proven to be essential for cell viability (Sharma, Pradhan et al. 2005, Paetzel, Karla et al. 2002, Date 1983, Klug, Jager et al. 1997). For example, *E. coli* strain IT41 possesses a mutated leader peptidase gene, which has a drastically reduced growth rate. The growth rate was reduced because *E. coli* only has one SPase I (Sharma, Pradhan et al. 2005, Inada, Court et al. 1989). In contrast to these findings, *S. sanguinis* does not show a diminished growth rate when this SPase I is knocked out.

Defining the link between XG2_0351 and biofilm formation on a molecular level demands characterizing the mechanism of action in details, including substrates of XG2_0351. This is a challenging task given the fact that *S. sanguinis* possesses two type I signal peptidases, namely XG2_0351 and XG2_0849, which may share the same pool of substrates. This led to further examination of bacteria that have multiple SPases I, a common feature of gram-positive bacteria (Bonnemain, Raynaud et al. 2004). In *E. coli* the SPase I is essential for cell viability but *S. sanguinis* has proven to sustain life even without this enzyme as seen in knockout experiments and growth curves (Figure 6). XG2_0351 is not essential because the other SPase I in *S. sanguinis*, XG2_0849, functionally compensates with respect to cell viability when XG2_0351 is absent. Some other types of bacteria that share this characteristic with *S. sanguinis* are *Streptomyces, S. lividans, L. monocytogenes, and B. japonicum* (Bonnemain, Raynaud et al. 2004).

The largest number of type I signal peptidases in one single species thus far have been found in gram-positive eubacterium *Bacillus subtilis*. Five genes that specify type I signal peptidases present on the *B. subtilis* chromosome. Studies have shown that these enzymes, denoted as SipS, SipT, SipU, SipV, and SipW, have different but overlapping substrate specificities (Sharma, Pradhan et al. 2005). There are two main advantages of having multiple SPase I encoding genes in *B. subtilis*. First there is a broader substrate specificity or

preference and secondly, a modulation of activity in response to high demands on the secretion machinery (Bron, Bolhuis et al. 1998, Bolhuis, Sorokin et al. 1996). Unlike the SPase I in *E. coli*, SipS was not essential for viability of the cell nor for protein secretion. Although in the absence of SipS, the rate of processing of several preproteins was reduced (Bolhuis, Sorokin et al. 1996). These sip genes are not essential individually but a specific combination of mutations in these genes is lethal (Bron, Bolhuis et al. 1998). It will be intriguing to discover if the same scenario will occur in *S. sanguinis* once both SPase I genes are knocked out.

To gain further clues regarding the molecular mechanism of action of XG2_0351, we searched XG2_0351 in the STRING database for protein-protein interaction networks from known metabolic pathways, protein complexes, signal transduction pathways, and other carefully selected databases (Figure 7A). The information obtained is from experimental data, computational prediction methods, and public text collections. Knowledge of protein-protein interactions is essential to understand cellular processes at the system-level. The protein-protein interaction network of XG2_0351 visualized by STRING revealed that this SPase I protein interacts with several ribosomal proteins. Ribosomal subunits that are involved in the cellular process of translation are composed of these proteins and rRNA. This SPase I also interacts with SRPR and SRP54 proteins which are signal recognition particles involved in targeting and inserting nascent membrane proteins into the cytoplasm. One or more SRP protein in conjunction with SRP RNA contributes to the binding and release of signal peptide. Then the SPase I proteolytically cleaves them from translocated precursor proteins from the extracytoplasmic site of the membrane (Auclair, Bhanu et al. 2012, du Plessis, Nouwen et al. 2011).

Gene co-occurrence visually displayed the gene families whose occurrence patterns across genomes show similarities (Figure 7B). For each gene of interest, the color indicates the similarity of its best hit in a given STRING genome. The similarities in these presence/absence profiles can predict interactions. Two distinct colors indicate the lowest and highest similarity observed within that clade. The highest similarities are in Firmicutes, which mostly have gram-positive cell wall structure. *Listeria monocytogenes*, a member of

Firmicutes, has three contiguous SPase I genes called SipX, SipY and SipZ. The major SPase I of *L. monocytogenes* is SipZ because the amounts of extracellular virulence factors such as listeriolysin O, phosphatidylcholine C, and zinc metalloproteinase were significantly decreased upon inactivation. For the majority of Sec-secreted exoproteins identified, the three SPases I were found to function redundantly. This became clear when protein secretion was not affected by the inactivation of only one or two of the SPases I. Since the SipZ of *L. monocytogenes* applies only to a small subset of the secreted exoproteins, the concept of minor and major SPases appears to be relative, not absolute (Bonnemain, Raynaud et al. 2004, Renier, Chafsey et al. 2015). In order to compare the type I signal peptidases in *S. sanguinis,* CLSM images of XG2_0351 paralogue, XG2_0849, were compared to images of SK36.

In order to investigate the potential role of the other SPase I, XG2_0849, in biofilm formation, we compared the biofilm formation between *S. sanguinis* wild type and XG2_0849 using confocal microscopy (Figure 9). The CLSM image of XG2_0849 showed a slight difference when assessed against SK36 but not as drastically as XG2_0351. There is a possibility that XG2_0351 is responsible for cleaving more biofilm related proteins than XG2_0849. The quantitative data obtained by measuring the relative intensities of confocal images (Figure 10) clearly illustrated that XG2_0351 biofilm is ten-fold decreased when compared to SK36, whereas the downstream mutant (XG2_0350) and parlogue (XG2_0849) were not significantly different from the wild type. In *S. sanguinis*, XG2_0351 appears to be the major SPase I when biofilm formation is concerned. This scenario was shown to occur in *P. aeruginosa*, which has two noncontiguous SPases I. PA1302 is involved with quorum-sensing cascade and includes the suppression of virulence factor secretion and virulence-associated phenotypes, while LepB is the primary SPase (Waite, Rose et al. 2012).

Finally, to narrow the list of potential substrates of XG2_0351 involved in biofilm formation, we extracted whole cell proteins from *S. sanguinis* wild type and compared it to that from XG2_0351 (Figure 11). The amount of protein extracted from wild type was almost two-fold the amount extracted from XG2_0351. We

further separated the extracted proteins by SDS-PAGE and stained with Coomassie-Brilliant Blue. Nine substrates of SPase I enzyme, XG2_0351, which were previously shown through CV staining (Figure 2) and confocal imaging (Figure 3) to exhibit reduced biofilm formation, were absent in XG2_0351 mutant in comparison to wild type. These may provide clues about molecular mechanism adopted by XG2_0351 to affect biofilm formation and provide a potential drug target with promising impact on reducing biofilm formation.

Further study is necessary to claim that XG2_0351 is to *S. sanguinis* what SipZ is to *L. monocytogenes*. The predominance of one SPase I over another is a bacterium is dependent on more than biofilm formation factors. The major SPase I is essential for efficient protein secretion which is contingent upon the availability of SPases, the production levels of secreted proteins, and substrate specificity or substrate preference of the different type I SPases (Bolhuis, Sorokin et al. 1996). Coomassie blue staining was done to measure the levels of proteins in SK36 and XG2_0351.

This study indicates that type I signal peptidase mutant, XG2_0351, causes a decrease in biofilm formation when compared to SK36. This SPase I performs a more critical role in biofilm formation than XG2_0849. Gene 0351 is possibly necessary for functions that include but are not limited to cell-wall metabolism, collagen-binding, iron transportation, and antibiotic resistance (Table 5 & Figure 11). These contribute to the successful formation of biofilm in *S. sanguinis*. Future studies may further the investigation by creating a double knockout mutant of both XG2_0351 and XG2_0849 and measuring cell viability. Pulse-chase protein radiolabeling would give a deeper understanding of the activity of proteins over a prolonged period of time. Mass spectrometry could be utilized to reveal which proteins are missing in XG2_0351 and therefore, which proteins are affecting biofilm formation in *S. sanguinis*.

LITERATURE CITED

Auclair SM, Bhanu MK, Kendall DA. Signal peptidase I: cleaving the way to mature proteins. *Protein Sci*. 2012;21:13-25.

Barnard JP, Stinson MW. The alpha-hemolysin of *Streptococcus gordonii* is hydrogen peroxide. *Infect Immun*. 1996;64:3853-3857.

Benito N, Miro JM, de Lazzari E, et al. Health care-associated native valve endocarditis: importance of non-nosocomial acquisition. *Ann Intern Med*. 2009;150:586-594.

Beynon RP, Bahl VK, Prendergast BD. Infective endocarditis. *BMJ*. 2006;333:334-339.

Bochud PY, Calandra T, Francioli P. Bacteremia due to viridans streptococci in neutropenic patients: a review. *Am J Med*. 1994;97:256-264.

Bolhuis A, Sorokin A, Azevedo V, et al. *Bacillus subtilis* can modulate its capacity and specificity for protein secretion through temporally controlled expression of the sipS gene for signal peptidase I.*Mol Microbiol*. 1996;22:605-618.

Bonnemain C, Raynaud C, Reglier-Poupet H, et al. Differential roles of multiple signal peptidases in the virulence of Listeria monocytogenes. *Mol Microbiol*. 2004;51:1251-1266.

Bron S, Bolhuis A, Tjalsma H, Holsappel S, Venema G, van Dijl JM. Protein secretion and possible roles for multiple signal peptidases for precursor processing in bacilli. *J Biotechnol*. 1998;64:3-13.

Cahill TJ, Prendergast BD. Infective endocarditis. *Lancet*. 2016;387:882-893.

Cahill TJ, Prendergast BD. Current controversies in infective endocarditis. *F1000Res*. 2015;4:10.12688/f1000research.6949.1. eCollection 2015.

Caufield PW, Dasanayake AP, Li Y, Pan Y, Hsu J, Hardin JM. Natural history of Streptococcus sanguinis in the oral cavity of infants: evidence for a discrete window of infectivity. *Infect Immun*. 2000;68:4018-4023.

Chen L, Ge X, Dou Y, Wang X, Patel JR, Xu P. Identification of hydrogen peroxide production-related genes in Streptococcus sanguinis and their functional relationship with pyruvate oxidase.*Microbiology*. 2011;157:13-20.

Costerton JW, Geesey GG, Cheng KJ. How bacteria stick. *Sci Am*. 1978;238:86-95.

Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: a common cause of persistent infections. *Science*. 1999;284:1318-1322.

Dalbey RE, Wickner W. Leader peptidase catalyzes the release of exported proteins from the outer surface of the Escherichia coli plasma membrane. *J Biol Chem*. 1985;260:15925-15931.

Date T. Demonstration by a novel genetic technique that leader peptidase is an essential enzyme of Escherichia coli. *J Bacteriol*. 1983;154:76-83.

de la Fuente-Nunez C, Reffuveille F, Fernandez L, Hancock RE. Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. *Curr Opin Microbiol*. 2013;16:580-589.

Donlan RM. Biofilms: microbial life on surfaces. *Emerg Infect Dis*. 2002;8:881-890.

Donlan RM. Role of biofilms in antimicrobial resistance. *ASAIO J*. 2000;46:S47-52.

Donlan RM, Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev*. 2002;15:167-193.

du Plessis DJ, Nouwen N, Driessen AJ. The Sec translocase. *Biochim Biophys Acta*. 2011;1808:851-865.

Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. *Nat Rev Microbiol*. 2014;12:49-62.

Ge X, Kitten T, Chen Z, Lee SP, Munro CL, Xu P. Identification of Streptococcus sanguinis genes required for biofilm formation and examination of their role in endocarditis virulence. *Infect Immun*. 2008;76:2551-2559.

Ge X, Shi X, Shi L, et al. Involvement of NADH Oxidase in Biofilm Formation in Streptococcus sanguinis. *PLoS One*. 2016;11:e0151142.

Hall MR, McGillicuddy E, Kaplan LJ. Biofilm: basic principles, pathophysiology, and implications for clinicians. *Surg Infect (Larchmt)*. 2014;15:1-7.

Hoen B, Duval X. Infective Endocarditis. *N Engl J Med*. 2013;368:1425-1433.

Inada T, Court DL, Ito K, Nakamura Y. Conditionally lethal amber mutations in the leader peptidase gene of Escherichia coli. *J Bacteriol*. 1989;171:585-587.

Jenkinson HF, Lamont RJ. Oral microbial communities in sickness and in health. *Trends Microbiol*. 2005;13:589-595.

Kilian M, Mikkelsen L, Henrichsen J. Taxonomic study of viridans streptococci: description of *Streptococcus gordonii* sp. nov. and emended descriptions of *Streptococcus sanguis* (White and Niven 1946), *Streptococcus oralis* (Bridge and Sneath 1982), and *Streptococcus mitis* (Andrewes and Horder 1906). *International Journal of Systemic Bacteriology*. 1989;39:471-484.

Klug G, Jager A, Heck C, Rauhut R. Identification, sequence analysis, and expression of the lepB gene for a leader peptidase in Rhodobacter capsulatus. *Mol Gen Genet*. 1997;253:666-673.

Kolenbrander PE, London J. Adhere today, here tomorrow: oral bacterial adherence. *J Bacteriol*. 1993;175:3247-3252.

Kreth J, Merritt J, Qi F. Bacterial and host interactions of oral streptococci. *DNA Cell Biol*. 2009;28:397-403.

Kreth J, Merritt J, Shi W, Qi F. Competition and coexistence between Streptococcus mutans and Streptococcus sanguinis in the dental biofilm. *J Bacteriol*. 2005;187:7193-7203.

Leriche V, Sibille P, Carpentier B. Use of an enzyme-linked lectinsorbent assay to monitor the shift in polysaccharide composition in bacterial biofilms. *Appl Environ Microbiol*. 2000;66:1851-1856.

Lewis K. Current topics in microbiology and immunology. In: Romeo T, ed. *Multidrug Tolerance of Biofilms and Persister Cells.* Vol 322. Tony Romeo; 2008:107.

Lister JL, Horswill AR. Staphylococcus aureus biofilms: recent developments in biofilm dispersal. *Front Cell Infect Microbiol*. 2014;4:178.

Loesche WJ. Role of Streptococcus mutans in human dental decay. *Microbiol Rev*. 1986;50:353-380.

Loo CY, Corliss DA, Ganeshkumar N. Streptococcus gordonii biofilm formation: identification of genes that code for biofilm phenotypes. *J Bacteriol*. 2000;182:1374-1382.

Luke I, Handford JI, Palmer T, Sargent F. Proteolytic processing of Escherichia coli twin-arginine signal peptides by LepB. *Arch Microbiol*. 2009;191:919-925.

Meijer WJ, de Jong A, Bea G, et al. The endogenous Bacillus subtilis (natto) plasmids pTA1015 and pTA1040 contain signal peptidase-encoding genes: identification of a new structural module on cryptic plasmids. *Mol Microbiol*. 1995;17:621-631.

Moreno-Hagelsieb G. The power of operon rearrangements for predicting functional associations. *Comput Struct Biotechnol J*. 2015;13:402-406.

Mylonakis E, Calderwood SB. Infective Endocarditis in Adults. *N Engl J Med*. 2001;345:1318-1330.

O'Toole GA. Microtiter dish biofilm formation assay. *J Vis Exp*. 2011;(47). pii: 2437. doi:10.3791/2437.

Paetzel M, Karla A, Strynadka NC, Dalbey RE. Signal peptidases. *Chem Rev*. 2002;102:4549-4580.

Parahitiyawa NB, Jin LJ, Leung WK, Yam WC, Samaranayake LP. Microbiology of odontogenic bacteremia: beyond endocarditis. *Clin Microbiol Rev*. 2009;22:46-64.

Pobre V, Arraiano CM. Next generation sequencing analysis reveals that the ribonucleases RNase II, RNase R and PNPase affect bacterial motility and biofilm formation in E. coli. *BMC Genomics*. 2015;16:72-015-1237-6.

Prendergast BD, Tornos P. Surgery for infective endocarditis: who and when? *Circulation*. 2010;121:1141-1152.

Rawlings ND, Barrett AJ. Evolutionary families of peptidases. *Biochem J*. 1993;290 (Pt 1):205-218.

Renier S, Chafsey I, Chambon C, et al. Contribution of the multiple Type I signal peptidases to the secretome of Listeria monocytogenes: deciphering their specificity for secreted exoproteins by exoproteomic analysis. *J Proteomics*. 2015;117:95-105.

Richet H, Casalta JP, Thuny F, et al. Development and assessment of a new early scoring system using non-specific clinical signs and biological results to identify children and adult patients with a high probability of infective endocarditis on admission. *J Antimicrob Chemother*. 2008;62:1434-1440.

Scheffers DJ, Pinho MG. Bacterial cell wall synthesis: new insights from localization studies. *Microbiol Mol Biol Rev*. 2005;69:585-607.

Scherr TD, Heim CE, Morrison JM, Kielian T. Hiding in Plain Sight: Interplay between Staphylococcal Biofilms and Host Immunity. *Front Immunol*. 2014;5:37.

Selton-Suty C, Celard M, Le Moing V, et al. Preeminence of Staphylococcus aureus in infective endocarditis: a 1-year population-based survey. *Clin Infect Dis*. 2012;54:1230-1239.

Seymour GJ, Ford PJ, Cullinan MP, Leishman S, Yamazaki K. Relationship between periodontal infections and systemic disease. *Clin Microbiol Infect*. 2007;13 Suppl 4:3-10.

Sharma S, Pradhan A, Chauhan VS, Tuteja R. Isolation and characterization of type I signal peptidase of different malaria parasites. *J Biomed Biotechnol*. 2005;2005:301-309.

Singh R, Ray P, Das A, Sharma M. Role of persisters and small-colony variants in antibiotic resistance of planktonic and biofilm-associated Staphylococcus aureus: an in vitro study. *J Med Microbiol*. 2009;58:1067-1073.

Sonneville R, Mirabel M, Hajage D, et al. Neurologic complications and outcomes of infective endocarditis in critically ill patients: the ENDOcardite en REAnimation prospective multicenter study.*Crit Care Med*. 2011;39:1474-1481.

Tan X, Qin N, Wu C, et al. Transcriptome analysis of the biofilm formed by methicillin-susceptible Staphylococcus aureus. *Sci Rep*. 2015;5:11997.

Thuny F, Avierinos JF, Tribouilloy C, et al. Impact of cerebrovascular complications on mortality and neurologic outcome during infective endocarditis: a prospective multicentre study. *Eur Heart J*. 2007;28:1155-1161.

Tjalsma H, Bolhuis A, van Roosmalen ML, et al. Functional analysis of the secretory precursor processing machinery of Bacillus subtilis: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. *Genes Dev*. 1998;12:2318-2331.

Tjalsma H, Noback MA, Bron S, Venema G, Yamane K, van Dijl JM. Bacillus subtilis contains four closely related type I signal peptidases with overlapping substrate specificities. Constitutive and temporally controlled expression of different sip genes. *J Biol Chem*. 1997;272:25983-25992.

Turner LS, Das S, Kanamoto T, Munro CL, Kitten T. Development of genetic tools for in vivo virulence analysis of Streptococcus sanguinis. *Microbiology*. 2009;155:2573-2582.

Waite RD, Rose RS, Rangarajan M, Aduse-Opoku J, Hashim A, Curtis MA. Pseudomonas aeruginosa possesses two putative type I signal peptidases, LepB and PA1303, each with distinct roles in physiology and virulence. *J Bacteriol*. 2012;194:4521-4536.

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57-63.

Xu P, Alves JM, Kitten T, et al. Genome of the opportunistic pathogen Streptococcus sanguinis. *J Bacteriol*. 2007;189:3166-3175.

Xu P, Ge X, Chen L, et al. Genome-wide essential gene identification in Streptococcus sanguinis. *Sci Rep*. 2011;1:125.

Yew HS, Murdoch DR. Global trends in infective endocarditis epidemiology. *Curr Infect Dis Rep*. 2012;14:367-372.