Salmonella Derby, un sérovar emergent encore méconnu

Salmonella enterica subsp. *enterica* sérovar Derby a été isolé en 1922 lors d'une TIAC reliée à des tourtes à la viande de porc à Derby en Angleterre (Peckham and Savage, 1923). Dès 1937, *Salmonella* Derby est reconnu comme un sérovar de salmonelle par le sous-comité *Salmonella* de 1934 (The Salmonella Subcommittee of the Nomenclature Committee of the International Society for, 1934). Depuis *Salmonella* Derby est associé de façon sporadique à des TIAC dans les pays développés. Ainsi, en 1946, *S.* Derby cause en Australie une épidémie chez 68 nourrissons et enfants en bas âge causant la mort de 10 d'entre eux (Mushin, 1948). En 1963, une TIAC causée par une contamination d'œufs par *S.* Derby cause 822 cas dans 53 hôpitaux aux Etat Unis (Sanders et al., 1963). Plus récemment, *S.* Derby est responsable d'une TIAC en Allemagne touchant 145 patients, essentiellement des personnes âgées à Berlin et dans l'état du Brandebourg en Allemagne (Simon et al., 2017). Cependant, au-delà de quelques épidémies, *S.* Derby reste majoritairement associé à des cas sporadiques (EFSA and ECDC, 2018).

2.2 *Salmonella* Derby dans les filières alimentaires

Salmonella Derby est historiquement associé à la contamination asymptomatique des porcs et de la volaille, en particulier de la dinde (Felsenfeld and Young, 1947; Valdezate et al., 2005; Hayward et al., 2013). Au niveau international *S.* Derby est en effet en compétition avec *S.* Typhimurium pour la première place au niveau de la prévalence chez le porc (Hauser et al., 2011; Schmidt et al., 2012) avec, en Europe, des prévalences respectives de 22,9% et 20.6% (EFSA, 2016). Une étude menée sur les salmonelles rencontrées dans les œufs en Uruguay reporte également *Salmonella* Derby comme le sérovar le plus fréquemment isolé devant *S.* Enteritidis (Betancor et al., 2010).

En Chine *Salmonella* Derby est le sérovar le plus communément retrouvé dans les abattoirs du secteur porcin (Cai et al., 2016). Aux Etats Unis, *S.* Derby est le 2^{éme} sérovar le plus communément isolé chez le porc après *S.* Typhimurium (Schmidt et al., 2012). En Europe, *Salmonella* Derby est présent dans 28,5 % des élevages de production (EFSA, 2008) et

représente, en 2015, le cinquième sérovar le plus fréquemment isolé dans les aliments et les animaux (EFSA, 2016).

Parmi les souches de *S.* Derby reportées par l'EFSA en 2016, 64.4 % proviennent du porc, 21 % de la dinde et 11,3 % de poulets de chair (EFSA, 2017). Au-delà de sa prévalence importante dans la filière porcine, *S.* Derby a également été reporté chez la volaille. Il est en effet le sérovar le plus fréquemment isolé chez la dinde au Royaume Uni avec une multiplication par 5 de sa prévalence entre 2014 et 2015 (EFSA, 2016; Hayward et al., 2016).

En France, *S.* Derby est le 4^{éme} sérovar le plus fréquemment isolé dans les filières alimentaires après *S.* Typhimurium, son variant monophasique *S.* 1,4,[5],12:i:- and *S.* Enteritidis (Leclerc et al., 2016). Ce sérovar est isolé principalement dans la viande de porc et de volaille avec une prévalence respective de 1,4 % chez le porc (DGAl, 2015) et de 3,2 % chez le poulet de chair (DGAl, 2014).

2.2.1 Description des filières alimentaires critiques en France.

2.2.1.1 Secteur porcin

La viande porcine représente 33,1 % de la consommation de viande totale des français avec 2,1 millions TEC (Tonne Equivalent Carcasse) consommée en 2017. La TEC, tonne équivalent carcasse, correspond à une unité de poids employée pour mesurer des quantités industrielles de viande, elle constitue à l'application d'un coefficient dépendant du type de produit considéré afin de ramener son poids à celui des carcasses nécessaires à la production du produit fini.

La consommation de viande de porc se répartit pour trois-quarts vers la charcuterie, la viande fraîche représentant un quart de la consommation (Agrimer, 2017). Plus de 90 % de la production est commercialisé par une quarantaine d'organisation de producteurs. Ces organisations sont de plus en plus impliquées dans la sélection et la transformation de la viande en produit fini (Agrimer, 2017).

La France dispose d'un cheptel de 12,8 millions de têtes dont 1 million de truies reproductrices. Elle abat annuellement 23,1 millions de porcs, faisant d'elle le 3^{éme} pays producteur en Europe après l'Allemagne et l'Espagne (Agrimer, 2017). L'abattage est réalisé pour sa majorité (95 %) dans 23 abattoirs spécialisés répartis dans les principaux bassins de production. Neuf de ces abattoirs ont une capacité annuelle de plus d'un million de têtes, tous localisés dans les régions Bretagne et Pays de la Loire (Agrimer, 2017).

Le secteur porcin en France est concentré sur 4 zones de production comme montré sur la **figure 26:**

- Le bassin Ouest (75 % de la production)
- Le bassin Nord (3,8 % de la production)

- Le bassin du Sud-Ouest (2,9 % de la production)
- Le bassin du Centre-Est (2,5 % de la production)

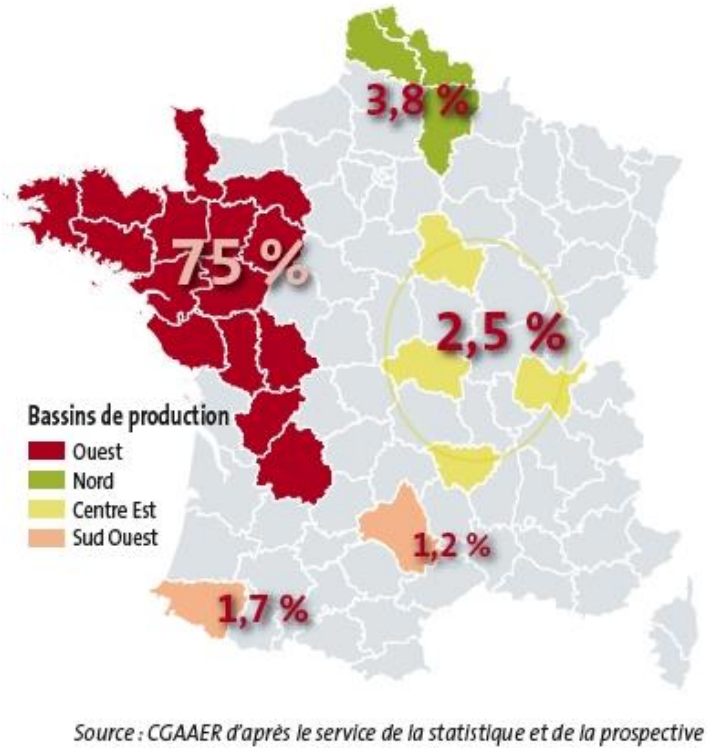
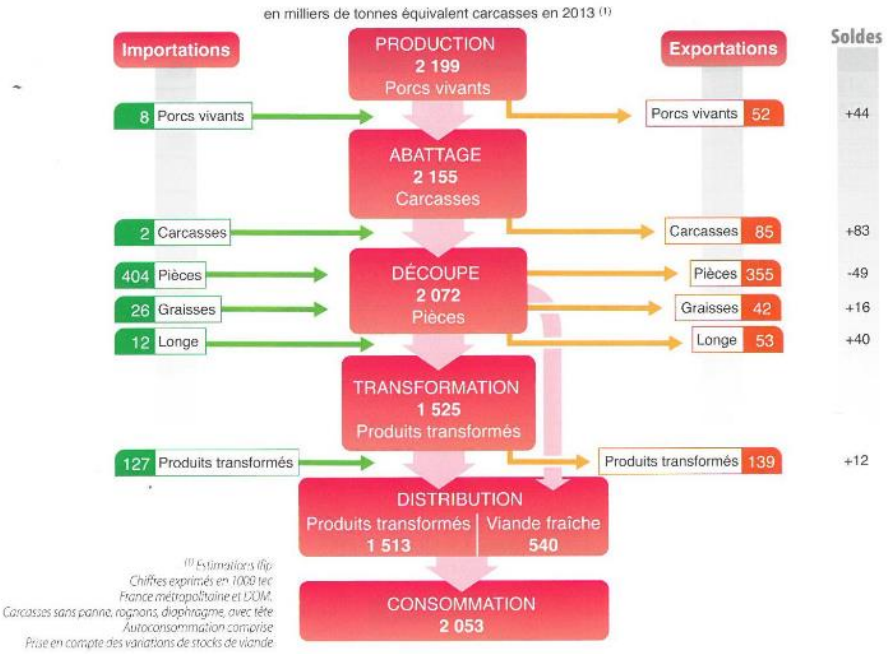


Figure 26: Répartition de la production porcine en France (CGAAER, 2012)**.**

La **figure 27** résume les chiffres clés de la filière :



Les échanges en France : les flux par maillon

Source : Le Porc par les Chiffres 2014-2015, Ifip-Institut du porc

Figure 27 : Chiffres clés de la filière Porc en France (IFIP, 2014)**.**

2.2.1.2 Secteur volailles

La consommation de volailles en France représente 28 % de la consommation annuelle avec 1,2 TEC consommé en 2017 et une consommation moyenne de viande rapportée à chaque habitant (Kgec/hab) de 27 dont 19,6 Kgec/hab (consommation de viande rapporté à chaque habitant) pour la viande de poulet et 4,3 Kgec/hab pour la dinde (Franceagrimer, 2018a).

La production de volailles française atteint les 1,81 millions TEC ce qui fait de la France le 3^{éme} producteur européen après la Pologne et le Royaume Uni. 34 % de cette production est destinée à l'export. Les abattoirs traitent 917 millions de volailles annuellement, des poulets de chair pour 82 % des abattages (Franceagrimer, 2018a). La **figure 28** résume les chiffres clés de la filière :

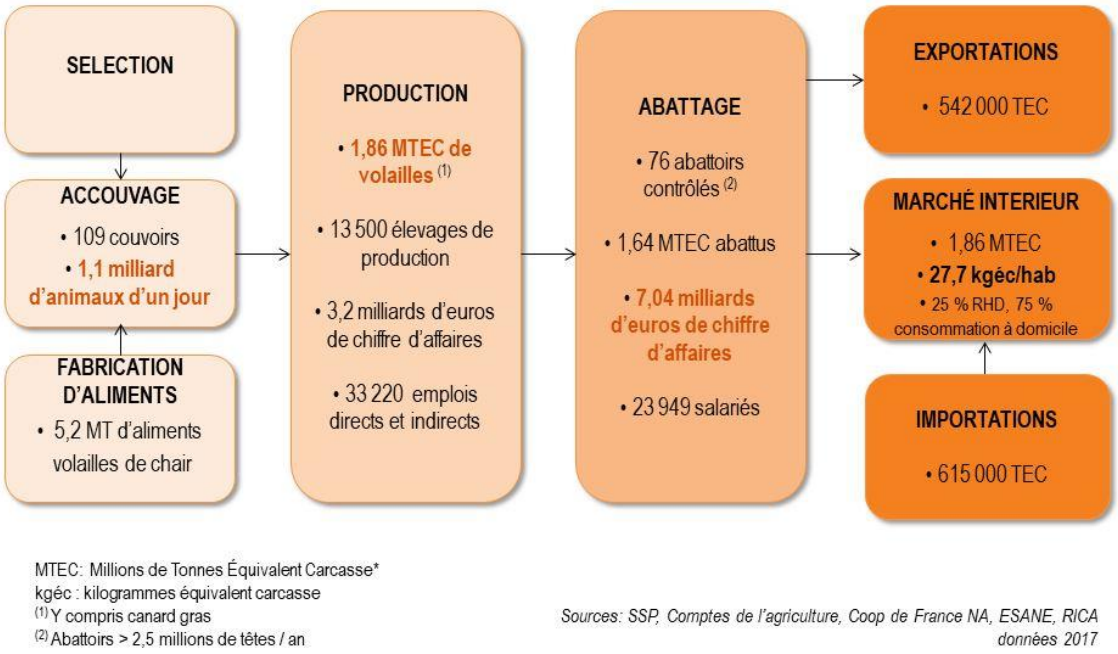


Figure 28 : Chiffres clés pour la filière volailles de chair (ITAVI, 2015)**.**

Le secteur volailles en France est concentré essentiellement dans l'Ouest du territoire :

- Le bassin Ouest (Bretagne, Pays de la Loire et Poitou-Charentes) représente 76 % de la production de volaille et compte 80 % des abattages.
- Le bassin Sud-Est (Rhône-Alpes, Bourgogne et Auvergne) représente 10 % de la production et concentre 11 % des abattages.
- Le bassin Sud-Ouest (Aquitaine et Midi-Pyrénées) représente 7 % de la production et concentre 6 % des abattages.
- Le bassin Nord (Nord-Pas-de-Calais, Picardie et Champagne-Ardenne) représente 6 % de la production et concentre 2, 5 % des abattages.

La répartition de la production de volailles en France est détaillée dans la **figure 29**.

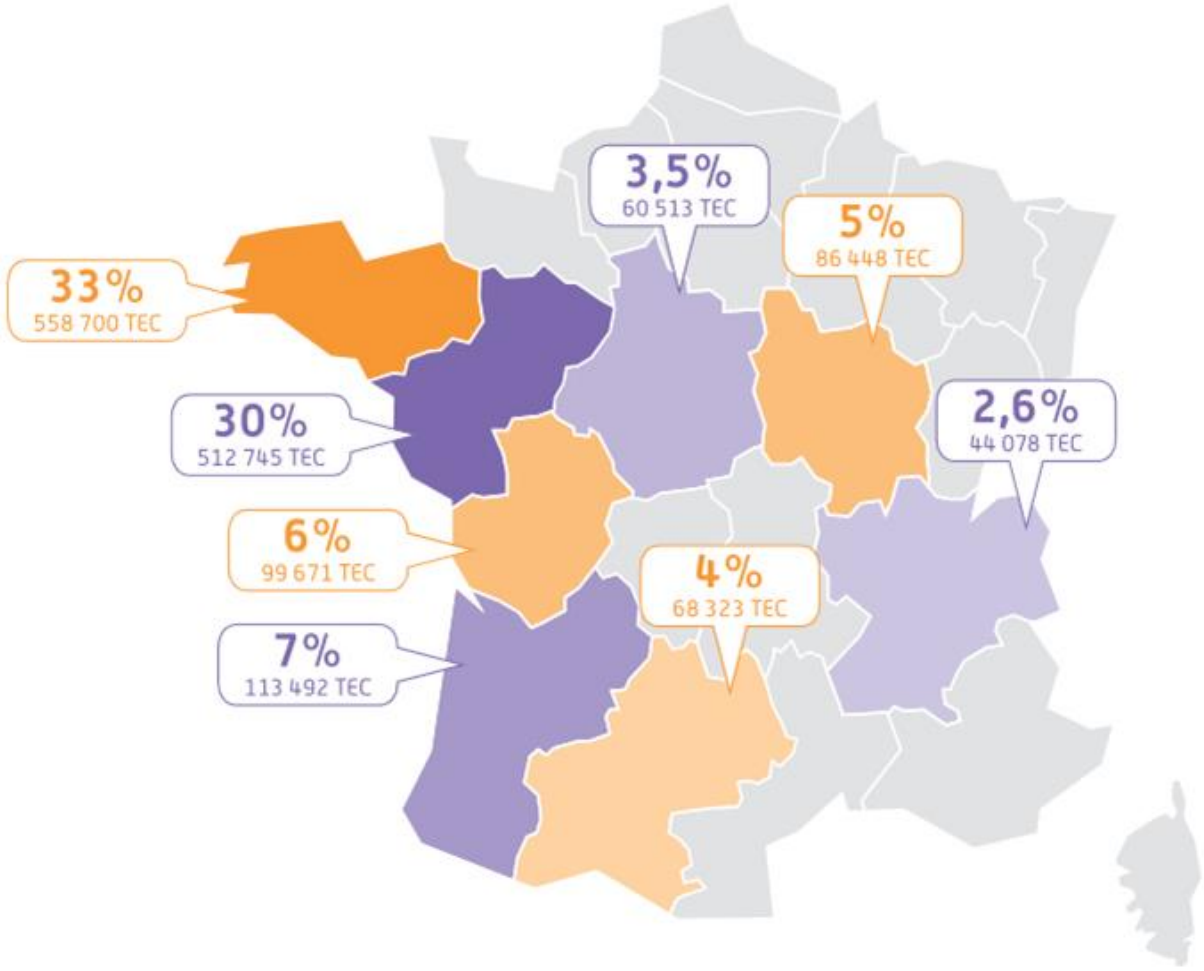


Figure 29: Répartition de la production avicole française (ITAVI, 2013)**.**

2.3 Impact sur la santé publique et prévalence

Au-delà de sa prévalence chez le porc et la volaille, plusieurs études ont démontré la capacité de *S.* Derby à passer du porc à l'humain via la consommation de viande de porc (Valdezate et al., 2005; Hauser et al., 2011; Kerouanton et al., 2013). *S.* Derby fait partie des 5 sérovars les plus fréquemment isolés chez l'humain (EFSA, 2015; 2016; 2017). En Chine, *S.* Derby est le 3ème sérovar le plus fréquemment isolé chez l'humain (Ran et al., 2011) et le plus fréquemment isolé chez les nourrissons et les jeunes enfants de 0 à 36 mois (Cui et al., 2009). Aux Etats-Unis par contre, *S.* Derby ne fait pas partie des sérovars les plus fréquemment rencontrés chez l'humain avec environ 120 cas par ans depuis 2005 (CDC, 2017).

En France, *Salmonella* Derby oscille entre la 5ème et la 8ème place des sérovars les plus fréquemment rencontrés chez l'humain depuis 2000 (avec un nombre de cas par an qui varie entre 164 et 178) (Weill, 2014).

2.4 Etudes de typage du sérovar Derby

Du fait de sa persistance dans les secteurs porcin et volailles, *S.* Derby pose un risque important pour la santé humaine. Il est donc nécessaire de pouvoir rapidement conduire des études épidémiologiques sur la source d'une contamination humaine à *S.* Derby.

Plusieurs études ont exploré la diversité du sérovar Derby par le passé :

En 2005, Valdezate et al. (Valdezate et al., 2005) étudient la diversité d'une collection de 110 souches espagnoles isolées entre 2000 et 2002. Ces souches, issues du porc ou isolées chez des patients, sont typées en utilisant la technique de typage de référence : la PFGE avec l'enzyme de restriction XbaI. Les résultats ne permettent pas de discrimination convenablement ces souches de *S.* Derby. En effet, la majorité des souches se répartissent sur un unique profil PFGE regroupant 53% des souches. La capacité d'antibio-résistance des souches a également été évaluée. La résistance à la streptomycine (STR) ou à la tétracycline (TET), ou au couple sulfaméthoxazole/tétracycline (SSS) était répandue dans plus de 50% des souches collectées. Parmi ces souches le profil de résistance majoritaire était le STR SSS TET.

En 2011, Hauser et al. (Hauser et al., 2011) conduisent une étude similaire en Allemagne sur 82 souches isolées chez le porc ou chez des patients entre 2006 et 2008. Le typage par PFGE (XbaI) donne 30 profils différents dont un profil majoritaire regroupant 31 souches et plusieurs profils minoritaire dont 26 profils ne comptant qu'une seule souche. Finalement, ces 30 profils différents ont été regroupés en 4 grands groupes (A, B, C et D) dont le groupe A qui rassemble 52 souches sur 82 (**figure 30**). L'étude comprend également une analyse MLST assignant les souches à 5 profils : le ST39 (48 souches), le ST40 (25 souches, le ST682 (5 souches) le ST71 (2 souches) et le ST774 (2 souches) très proche du ST39. La plupart des souches (72%) ne possédaient pas de résistance aux antibiotiques testés.

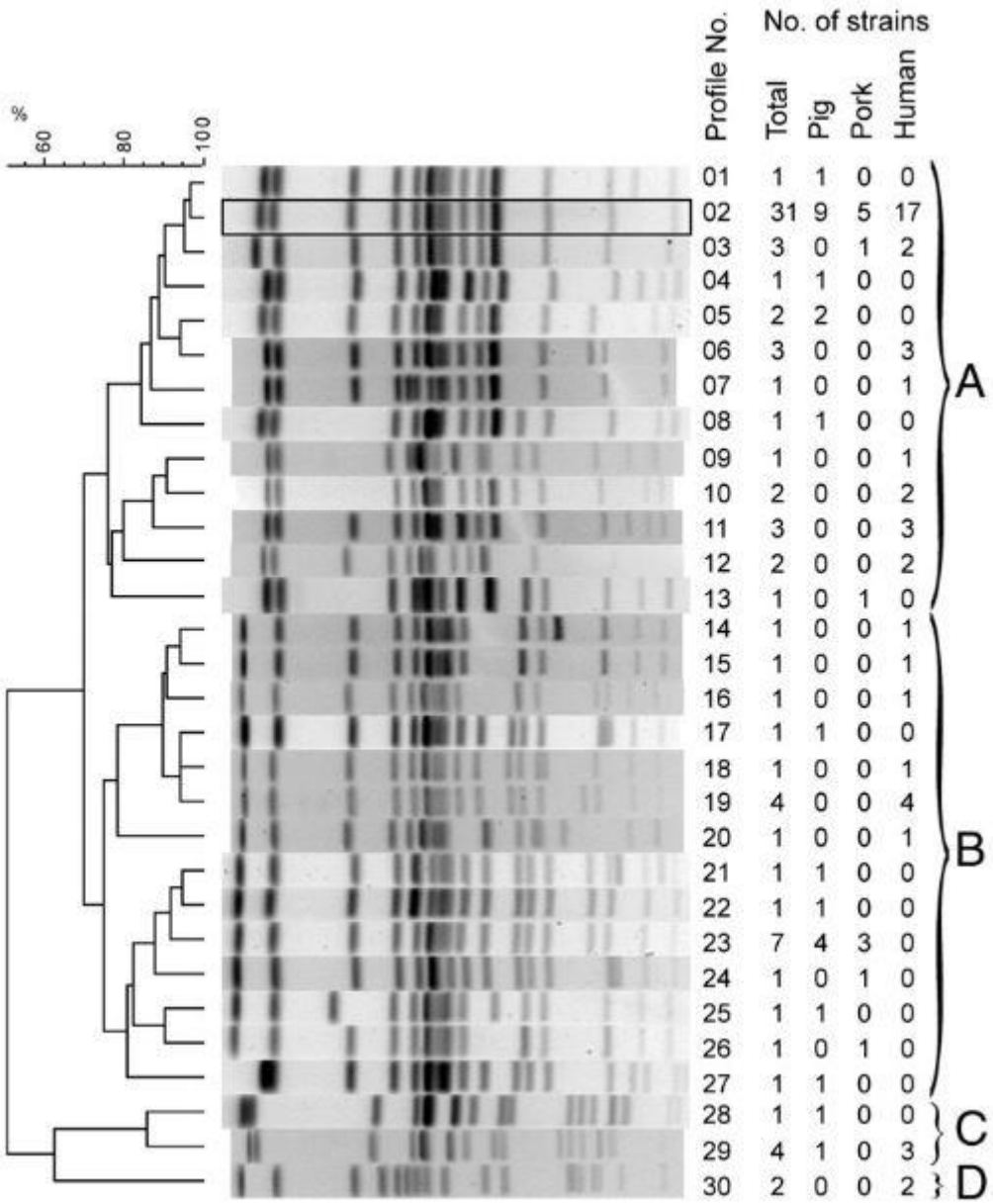
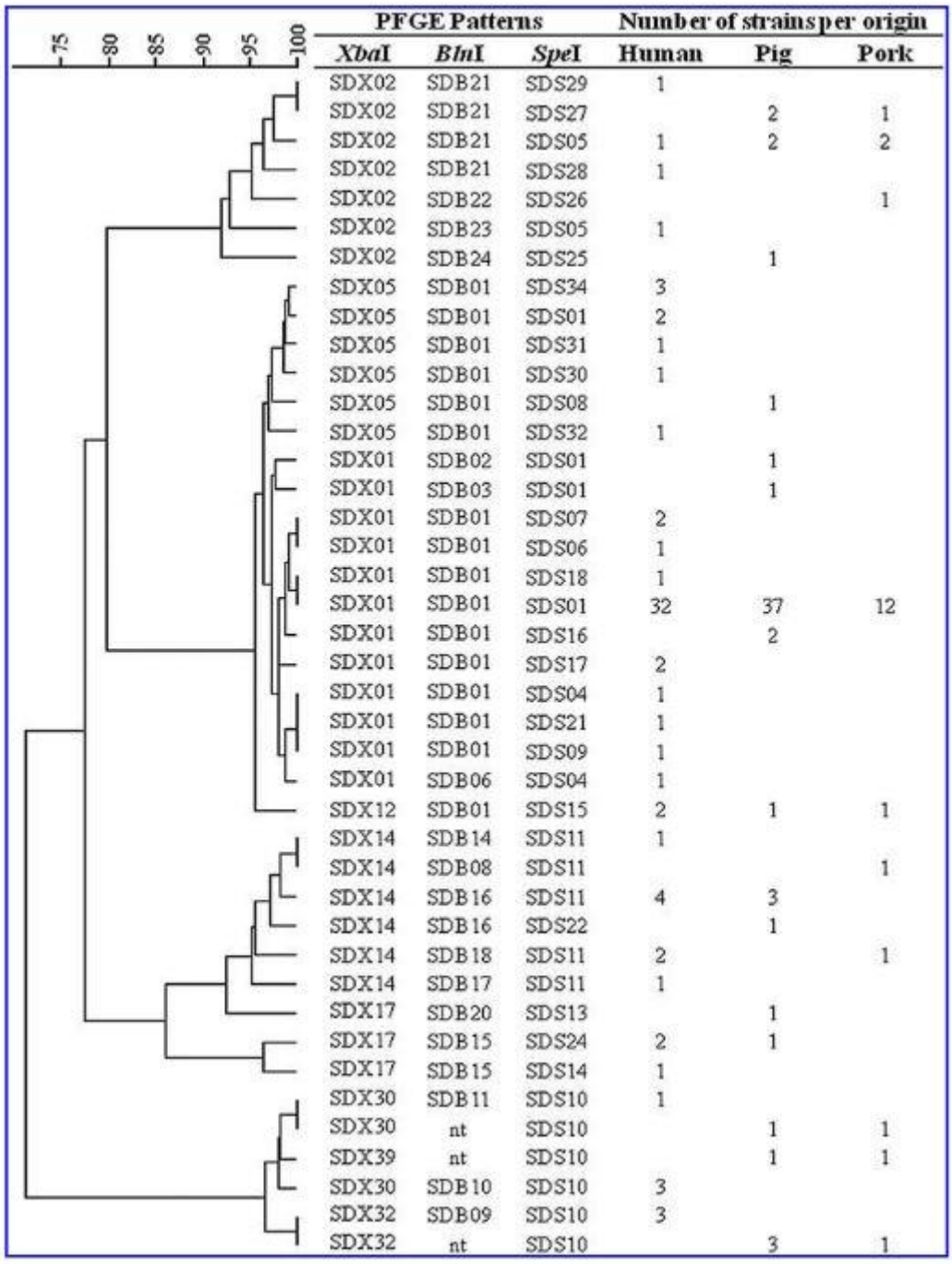


Figure 30 : Dendrogramme UPGMA des profils PFGE de 82 souches de *S.* Derby isolée en Allemagne entre 2006 et 2008 (Hauser et al., 2011)**.**

Le groupe A regroupe des souches des ST39, ST40 et ST774, le groupe B des souches du ST40, le groupe C du ST682 et le groupe D du ST71. Le profil majoritaire est encadré en noir.

En 2013, Kerouanton et al. (Kerouanton et al., 2013) conduisent une étude sur la diversité de *S.* Derby en France par PFGE sur une collection de 170 souches porcines et humaines isolées entre 2006 et 2007. Malgré l'utilisation de trois enzymes de restriction (XbaI, BlnI et SpeI) les résultats (**figure 31**) conduisent à l'identification d'un profil majoritaire regroupant 81 souches (48 %), les 41 autres profils regroupant moins de 7 souches. 84 % des souches étaient résistantes à au moins un des antibiotiques testés et 69 % des souches présentaient le profil de résistance STR SSS TET. Cette étude conclut que la PFGE n'a pas un pouvoir de discrimination suffisant pour le typage de *Salmonella* Derby.



PFGE Patterns			Number of strains per origin		
*Xba*I	*Bln*I	*Spe*I	Human	Pig	Pork
SDX02	SDB21	SDS29	1		
SDX02	SDB21	SDS27		2	1
SDX02	SDB21	SDS05	1	2	2
SDX02	SDB21	SDS28	1		
SDX02	SDB22	SDS26			1
SDX02	SDB23	SDS05	1		
SDX02	SDB24	SDS25		1	
SDX05	SDB01	SDS34	3		
SDX05	SDB01	SDS01	2		
SDX05	SDB01	SDS31	1		
SDX05	SDB01	SDS30	1		
SDX05	SDB01	SDS08		1	
SDX05	SDB01	SDS32	1		
SDX01	SDB02	SDS01		1	
SDX01	SDB03	SDS01		1	
SDX01	SDB01	SDS07	2		
SDX01	SDB01	SDS06	1		
SDX01	SDB01	SDS18	1		
SDX01	SDB01	SDS01	32	37	12
SDX01	SDB01	SDS16		2	
SDX01	SDB01	SDS17	2		
SDX01	SDB01	SDS04	1		
SDX01	SDB01	SDS21	1		
SDX01	SDB01	SDS09	1		
SDX01	SDB06	SDS04	1		
SDX12	SDB01	SDS15	2	1	1
SDX14	SDB14	SDS11	1		
SDX14	SDB08	SDS11			1
SDX14	SDB16	SDS11	4	3	
SDX14	SDB16	SDS22		1	
SDX14	SDB18	SDS11	2		1
SDX14	SDB17	SDS11	1		
SDX17	SDB20	SDS13		1	
SDX17	SDB15	SDS24	2	1	
SDX17	SDB15	SDS14	1		
SDX30	SDB11	SDS10	1		
SDX30	nt	SDS10		1	1
SDX39	nt	SDS10		1	1
SDX30	SDB10	SDS10	3		
SDX32	SDB09	SDS10	3		
SDX32	nt	SDS10		3	1

Figure 31 : Dendrogramme représentant les 42 profils PFGE obtenus sur 170 souches de *S.* Derby isolées en France entre 2006 et 2007 (Kerouanton et al., 2013)**.**
(nt: souches non typables)

Au Royaume Uni, Hayward et al. mène plusieurs études de génomique fonctionnelle sur *S.* Derby (Hayward et al., 2013; Hayward et al., 2014). Ce travail met en évidence la présence d'un nouvel îlot de pathogénicité le SPI-23, localisé dans le génome entre le tRNA-asn et le gène *docB*. Cet îlot de 37 kb comprend 42 cadres de lecture ouverts (ORF pour *Open reading Frame*) : des séquences susceptibles d'être traduites en protéine. La plupart de ces ORF (n=28) codent pour des protéines hypothétiques, les autres correspondent, entre autre, à des

protéines se fixant sur l'ADN (n=3), des effecteurs putatifs du T3SS (n=9) et de protéines associées à la production de pili (n=2) (Hayward et al., 2013). Des essais d'invasion cellulaire sur des souches de *S.* Derby démontrent que le SPI-23 joue un rôle encore mal élucidé dans l'invasion du jéjunum porcin (Hayward et al., 2014) ; Le SPI-23 subissant une régulation à la hausse dans le jéjunum porcin et participant ainsi à un tropisme vers les cellules épithéliales.

Une étude sur la structure de la population de 17 souches de *S.* Derby au Royaume Uni basée sur la comparaison des profils MLST et du SPI-23 (Hayward et al., 2016) démontre l'existence de deux lignées chez *S.* Derby, l'une correspondant au ST40, possédant le SPI-23, retrouvée chez le porc et la dinde, l'autre correspondant au ST71, ne possédant pas le SPI-23 retrouvée uniquement chez la dinde (Hayward et al., 2016).

En 2017 Zheng et al. (Zheng et al., 2017) étudie la diversité de *S.* Derby en combinant la méthode MLST et la technique CRISPR sur une collection de 100 souches chinoises isolées à partir du porc et de l'humain. Il parvient à la conclusion que le profil CRISPR est fortement corrélé avec le profil MLST et identifie deux groupes correspondant aux profils MLST ST40 et ST71 qui présentent des profils CRISPR très éloignés.

Pour conclure : Les études menées sur *S.* Derby concluent à un manque de pouvoir discriminant des méthodes de typage traditionnelles pour ce sérovar. Elles se concentrent essentiellement sur la filière porcine et sur les souches humaines avec un indice de discrimination qui reste faible (Kerouanton et al., 2013). Des études préliminaires menée au royaume Unis (Hayward et al., 2016) démontrent sur une petite collection (n=17 génomes) l'existence d'une lignée de *S.* Derby associé à la dinde présentant une distance génétique importante par rapport à la lignée majoritaire isolée chez le porc, la dinde et l'humain. Cependant, la petite taille de la collection étudiée ne permet pas de tirer de conclusions sur la diversité génétique de ce sérovar.

Problématique

Diversité génétique de S. Derby de la fourche à la fourchette et étude de sa pathogénicité.

En Europe, on considère que 10 à 20 % des cas de salmonelloses sont attribués au réservoir porcin (Pires et al., 2014). Les salmonelloses dues à la consommation de viande de porc constituent donc un problème majeur de santé publique à l'échelle européenne. Les sérovars les plus fréquemment rencontrés dans le secteur porcin sont : le variant monophasique de Typhimurium, S. Typhimurium et *S.* Derby avec des prévalences respectives de 22,3 %, 20,3 % et 22,9 % en 2015 (EFSA, 2016).

En France, l'incidence de *S.* Derby isolé de la viande de porc est passée de 20 à 40 % de 2005 à 2010, faisant de *S.* Derby le sérovar le plus fréquemment isolé dans l'alimentation humaine en 2014-2015 (Leclerc et al., 2016). En même temps, *S.* Derby s'est hissé depuis 2000 parmi les 10 sérovars les plus fréquemment isolés chez l'Homme (Weill, 2014). Alors que *S.* Typhimurium, par exemple, est soumis à plusieurs réglementations européennes, ce n'est pas le cas de *S.* Derby à l'heure actuelle. Pourtant, plusieurs études ont démontré la prédominance chez *S.* Derby de souches multirésistantes en France et en Espagne (Valdezate et al. 2005, Keouranton et al., 2013). Ce profil est fréquement isolé partout en Europe (EFSA, 2016).

Des souches de *S.* Derby porteuses de multirésistances aux antibiotiques sont donc capables d'échapper au système de surveillance européen et de circuler entre les états membres. L'absence de réglementation pour *S.* Derby est due pour partie au manque de techniques de typage fiables pour les souches de ce sérovar. L'attribution des cas humains à des sources d'infection (l'étude de source-attribution) est cruciale pour le choix de cibles d'intervention dans l'industrie agro-alimentaire (Pires et al., 2009). Pour être efficace, une étude de source-attribution doit s'appuyer sur une connaissance de la diversité de l'agent pathogène considéré afin de pouvoir retracer efficacement le réservoir des souches retrouvées chez l'Homme. Les techniques de typage de référence tel que la PFGE manquant de pouvoir de discrimination pour ce sérovar (Valdezate et al., 2005; Hauser et al., 2011; Kerouanton et al., 2013), l'attribution des cas sporadiques de *S.* Derby est problématique.

L'objectif de ma thèse était, d'une part, de mener une étude globale de la diversité génétique de *Salmonella* Derby de l'abattage au produit finis entre 2014 et 2015 en France ainsi que, d'autre part, d'effectuer une étude de source-attribution incluant les souches humaines isolées à cette même période. Les techniques d'analyse du génome complet étant considérés comme les plus discriminantes pour le sous typage de *Salmonella* (Ruppitsch, 2016), j'ai mené cette étude en utilisant ces approches.

S. Derby étant considéré comme caractéristique de la lignée porcine, mais étant également très présent chez la volaille, j'ai voulu :

- identifier les clones circulant en France dans les secteurs porcin et avicole et analyser les différences génétiques entre ces réservoirs ;
- évaluer la virulence des différents clones de *S.* Derby pour l'humain ainsi que leur adaptation à leur hôte respectif.
- A partir de ces informations : mener une étude d'attribution des sources sur les cas sporadiques attribués à *S.* Derby en France.
- Identifier des marqueurs génétiques permettant d'identifier les clones circulants (voir le clone majoritaire de *S.* Derby).

Pour mener à bien ce projet de thèse, j'ai sélectionné une collection de 140 souches isolées en 2014 et 2015 issues des secteurs porcin et avicole français tel que décrit dans le chapitre 4.1. J'ai conduit une analyse phylogénique basée sur la prédiction des SNP par *variant calling* afin de déterminer la diversité génétique du sérovar Derby sur ma collection de souches. J'ai également étudié les propriétés génétiques des lignées rencontrées en me focalisant sur les facteurs de virulences, particulièrement les Ilots de Pathogénicité de *Salmonella* (SPI-1 à 5), le SPI-23 spécifique de *S.* Derby, et les gènes de résistance aux antibiotiques (Premiere Publication : "*Polyphyletic nature of Salmonella enterica serotype Derby and lineage-specific host-association revealed by genome-wide analysis"* – Frontiers May 2018 | Volume 9 | Article 891 DOI : 10.1128/MRA.01027-18 *: Résultats partie 1*).

Pour étudier les caractéristiques des différentes lignées de *S.* Derby identifiées précédemment de façon plus fine il était nécessaire de produire des génomes de références. A l'heure actuelle il n'existe pas de génomes circulaires pour ce pathogène dans les bases de données en libre accès. J'ai donc choisi de produire deux génomes circulaires de *S*. Derby, l'un issu d'une souche de la filière porcine et l'autre des filières volailles. J'ai sélectionné ces deux génomes au sein de ma collection en fonction de la qualité des extractions et des séquençages obtenues par la technologie Illumina. Un séquençage par PacBio a ensuite été réalisé pour produire des données de séquençage de grande taille, facilitant l'assemblage complet des génomes. Cela m'a permis de produire deux séquences de génomes de référence pour *S.* Derby.

Le génome issu de la filière porcine a été valorisé par une deuxième publication : « *Complete genome for Salmonella enterica subsp. enterica serotype Derby associated with the pork sector in France*" – Microbiology Resource Anouncement, Volume 7 | Issue 12 | e01027-18, décrit dans les Résultats partie 2)

Le génome de référence issu de la lignée volailles m'a permis d'effectuer une analyse plus fine de la diversité génétique des souches de *S*. Derby issues de la volaille et de l'humain, isolées principalement en Europe. Cette étude fait l'objet d'une troisième publication actuellement soumise à la revue PlosOne : « *Phylogenomic analysis of Salmonella enterica subsp. enterica*

serovar Derby circulating in Europe, Asia and the United States and associated with the poultry sector », (Résultats partie 3)

Muni de génomes de référence de *S.* Derby, j'ai entrepris une étude de source-attributionen incluant l'ensemble des souches humaines de *S.* Derby collectée par l'Institut Pasteur (n=302) en 2014 et 2015, cette étude fait l'objet d'un article en préparation (résultats partie 4). Enfin, j'ai étudié le pan-génome (*core* génome + génome accessoire) des différentes lignées de *S.* Derby identifiées en effectuant une analyse comparative poussée sur la diversité des séquences des facteurs de virulences. Enfin, la capacité d'invasion des différentes lignées de *S.* Derby a été estimée via plusieurs essais d'invasion cellulaire sur des tapis de cellules épithéliales humaine, porcine et de poulet, étude menée en collaboration avec une équipe de l'Unité mixte de recherche Infectiologie et santé publique de l'INRA de Nouzilly. Les résultats de cette étude font l'objet d'une publication en préparation (résultats partie 5).

Résultats

1 Sélection de la collection.

J'ai sélectionné 140 souches alimentaires de la collection du Réseau *Salmonella* (LSAl) isolées en 2014 et 2015. L'intégralité des 302 souches humaines issues de la collection du CNR (et isolées dans le même laps de temps) a été obtenue de l'Institut Pasteur (collaboration CNR-LSAl).

L'ensemble de souches alimentaires (n=140) ont été séquencées grâce au financement ANSES/INAPORC - AAP 2015 (collaboration avec l'Institut du Porc (Ifip)).

1.1 Sélection des souches alimentaires

Les souches alimentaires ont été sélectionnées parmi celles isolées des secteurs porcin et volailles. Les souches issues des secteurs bovin et ovin ont été exclues de l'analyse car la prévalence de *S.* Derby a été estimée à moins d'1 % pour ces deux secteurs, et du fait de la faible part du secteur ovin dans la consommation de viandes en France (3 %) (Franceagrimer, 2018b). Pour les secteurs porcin et volailles la prévalence de *S.* Derby est respectivement supérieure à 10 % pour la viande fraiche (DGAl, 2015) et 12 % (dont 12.6 % pour *Gallus gallus* et 14.6 % pour la dinde) (DGAl, 2014).

Dans le cadre de ma thèse, j'ai sélectionné principalement des souches isolées de matière première (carcasses à l'abattage et viandes dans les ateliers de découpe ou préparation) considérant que cet échelon est le lien le plus probable entre l'animal et l'humain et la première source de contamination pour les produits élaborés et transformés.

La sélection de l'effectif des souches pour chaque secteur (porcin et volailles) a été déterminée sur la base de la proportion de la consommation de viande de porc et de volailles en France. Le **tableau 4** illustre cette proportion.

Tableau 4 : Plan de sélection des souches alimentaires et répartition de l'effectif pour chaque secteur (porcin et volailles) par rapport à la consommation en France.

Secteur	% de la consommation	Nombre de souches
Total	67,7	140
Porc	38,3	84
Volailles	29,4	56

Pour ces deux filières, les souches ont été sélectionnées suivant deux critères :
i/ favoriser les souches de la collection issues des plans de surveillance ;
ii/ sélectionner, pour chaque région, un nombre de souches proportionnel à la part de cette région dans la production nationale de la filière respective, porcine ou volailles.

1.1.1 Souches du secteur porcin

La **figure 32** représente les souches issues du secteur porcin collectées par le réseau *Salmonella* en 2014 et 2015 (n=181). La majorité des souches provient des bassins Ouest et Sud-Ouest. La répartition des souches est concordante avec la répartition géographique du secteur porcin. La plupart des souches sont issues des étapes d'abattage (69 souches), de transformation industrielle (69) et de la vente au détail (41).

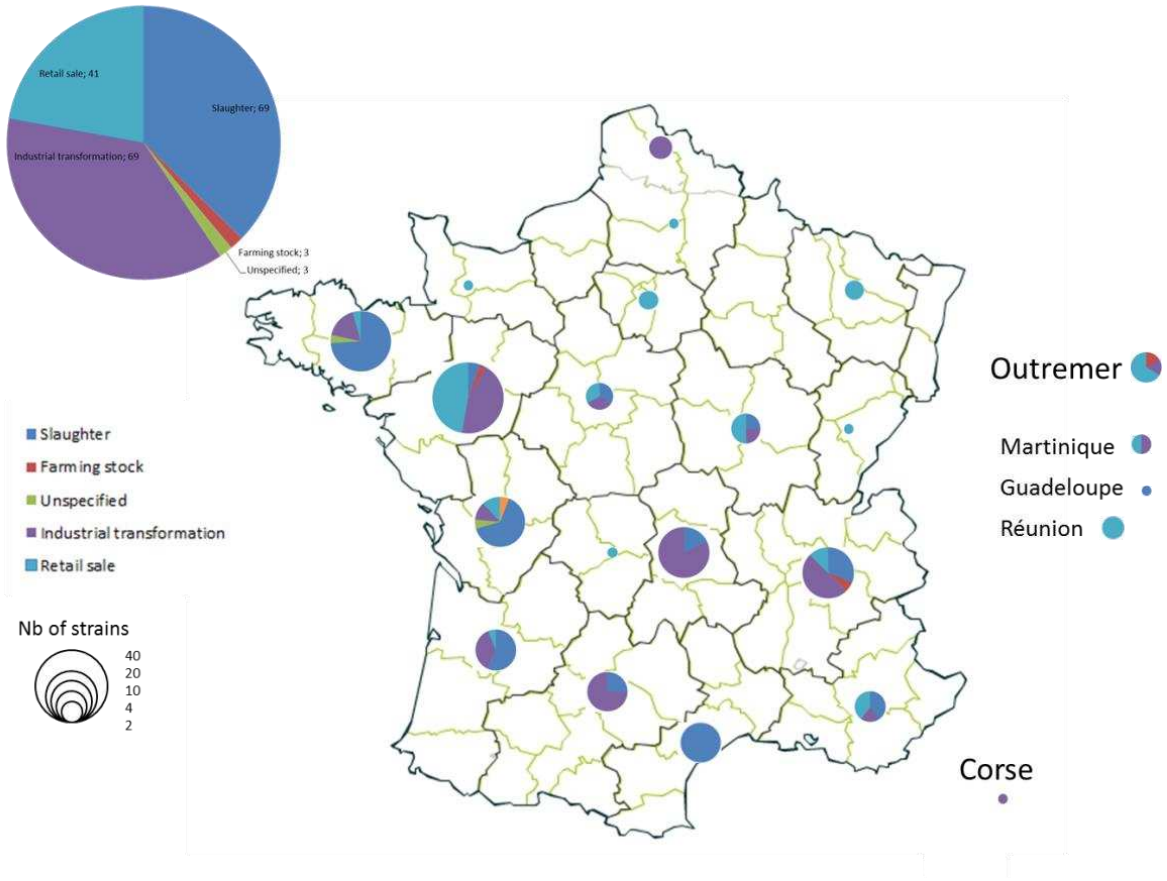


Figure 32: Répartition des souches de *S.* Derby collectées par le réseau *Salmonella* pour le secteur porcin en 2014-2015.

Le **tableau 5** résume la sélection finale des souches issues de la collection du Réseau *Salmonella* (secteur porcin).

Tableau 5 : Plan de sélection pour les souches issue du secteur porcin.

(note : régions suivant la nomemclature à l'œuvre en 2014)

Région	Effectifs (milliers de têtes)	Part dans la production nationale %	Souches disponibles	Souches dédoublonnées	Plans de surveillance	Sélection finale
Île de France	8	0,06	2	2	0	1
Champagne-Ardennes	176	1,31	0	0	0	0
Picardie	121	0,90	1	1	0	0
Haute Normandie	138	1,03	0	0	0	0
Centre	333	2,48	3	3	0	2
Basse Normandie	566	4,22	1	1	0	0
Bourgogne	126	0,94	4	4	0	2
Nord-Pas-de-Calais	472	3,52	2	2	0	2
Lorraine	97	0,72	2	2	0	1
Alsace	101	0,75	0	0	0	0
Franche-Comté	118	0,88	1	1	0	0
Pays de la Loire	1573	11,74	36	34	1	13
Bretagne	7635	56,97	27	17	7	15
Poitou-Charentes	366	2,73	17	14	0	4
Aquitaine	388	2,90	17	15	1	12
Midi-Pyrénées	435	3,25	8	7	0	2
Limousin	124	0,93	1	1	0	1
Rhône-Alpes	294	2,19	16	16	3	8
Auvergne	231	1,72	18	16	2	11
Languedoc Roussillon	28	0,21	8	7	3	8
PACA	22	0,16	5	4	0	1
Corse	50	0,37	1	1	0	0
Total	**13402**	**100**	**179**	**148**	**17**	**84**

Quand le nombre de souches dans une région n'est pas suffisant pour correspondre à cette proportion idéale, des souches additionnelles issues du même bassin de production ont été ajoutées. Si cela est impossible, les souches additionnelles ont été choisies sur des critères de proximité géographique.

1.1.2 Souches du secteur avicole

La **figure 33** correspond à la réparation géographique des souches collectées par le Réseau *Salmonella* en 2014-2015 pour le secteur avicole.

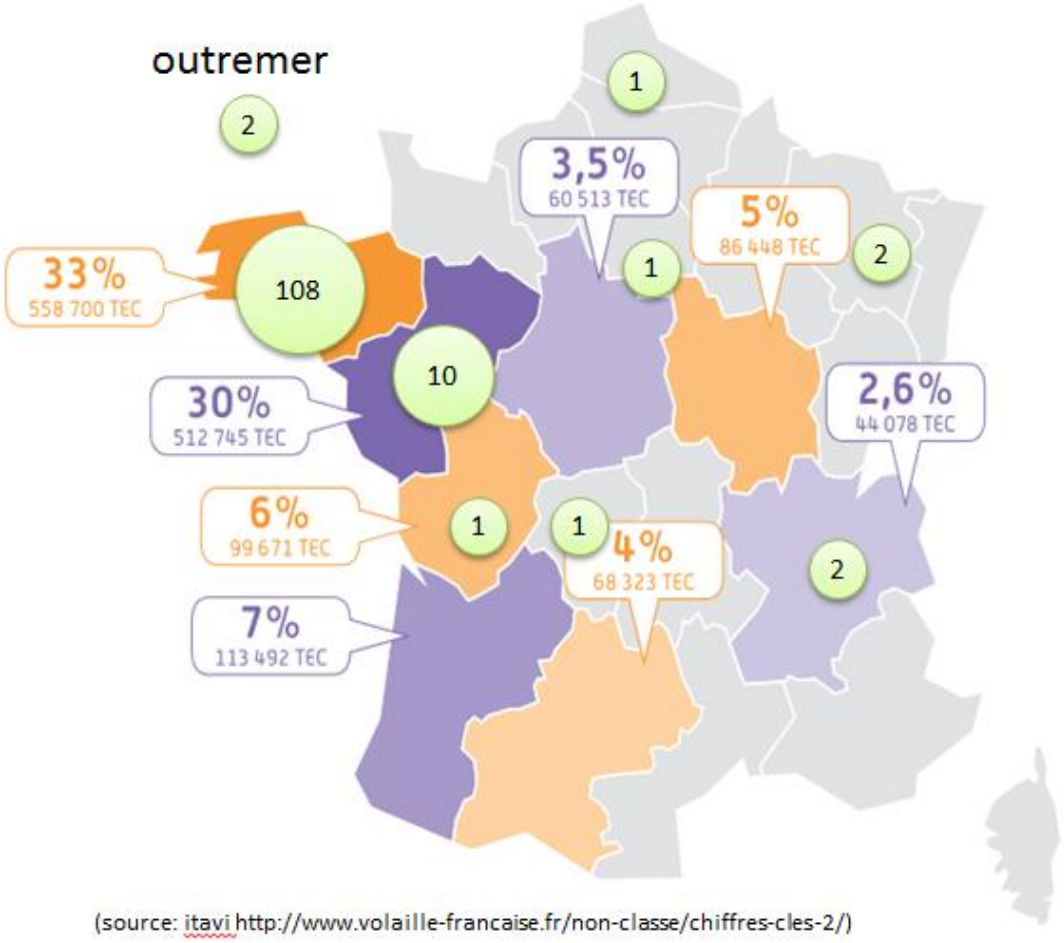


Figure 33: Correspondance entre les souches de *S.* Derby sélectionnées et les chiffres de production du secteur avicole en France en 2014-2015 (TEC = tonne équivalent carcasse) (ITAVI, 2016).

La majorité des souches collectées correspondent au plan de surveillance N°2013-9926. La **figure 34** détaille la répartition de ces souches. La majorité des souches a été isolée à l'étape d'abattage (97 souches sur 127 au total).

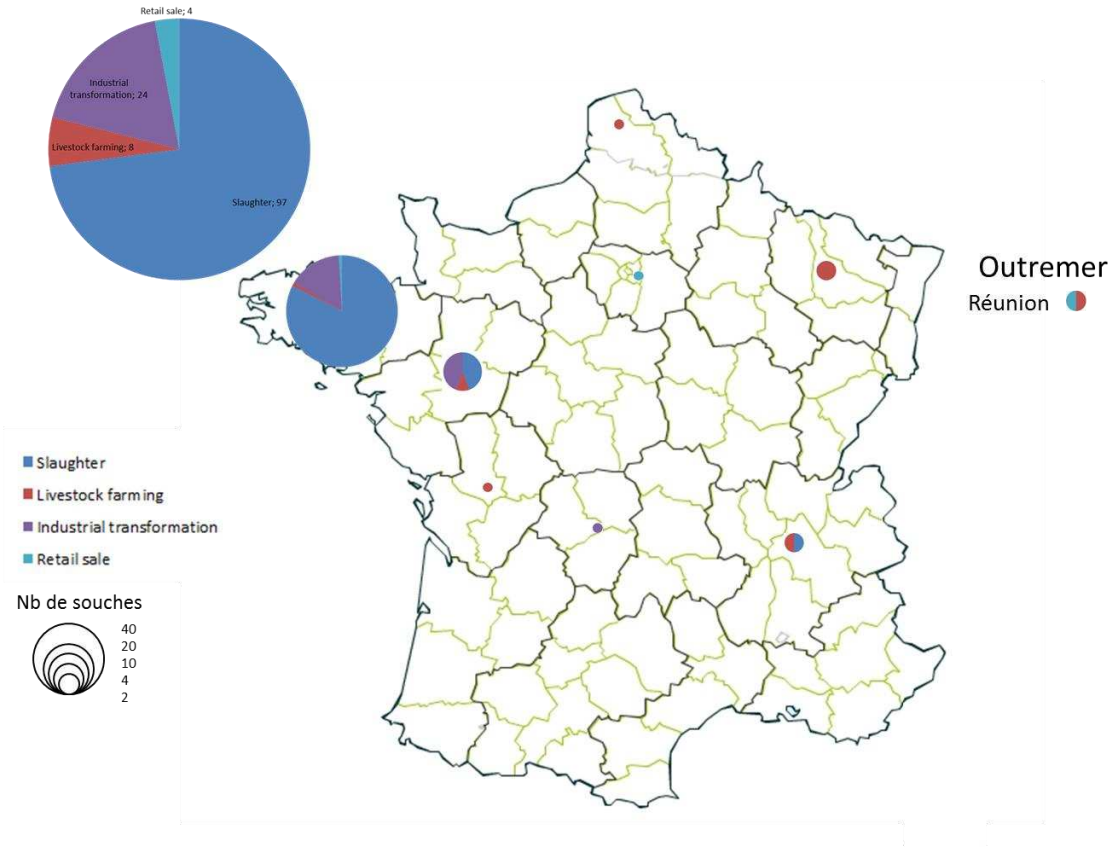


Figure 34 : Répartition des souches de *S.* Derby collectées par le réseau *Salmonella* pour le secteur avicole.

Le **tableau 6** résume la répartition des souches sélectionnées pour les filières volailles. Le manque de souches hors de Bretagne nous a conduit à inclure l'ensemble des souches isolées dans les autres régions.

Tableau 6: Plan de sélection pour les souches du secteur volaille.

(note : régions seuivant la nomemclature à l'œuvre en 2014)

Région	Production (en milliers de têtes)	Pourcentage de la production nationale	Souches disponibles	Souches dédoublonnées	Plans de surveillance	Sélection finale
Île de France	6	0	1	1	0	0
Champagne-Ardennes	30	2	0	0	0	0
Picardie	35	2	0	0	0	0
Haute Normandie	7	0	0	0	0	0
Centre	109	7	0	0	0	0
Basse Normandie	42	3	0	0	0	0
Bourgogne	49	3	0	0	0	0
Nord pas de calais	56	3	1	1	0	0
Lorraine	5	0	2	0	0	1
Alsace	6	0	0	0	0	0
Franche-Comté	3	0	0	0	0	0
Pays de la Loire	407	24	10	9	1	6
Bretagne	571	34	108	83	55	46
Poitou-Charentes	100	6	1	1	0	1
Aquitaine	68	4	0	0	0	0
Midi-Pyrénées	38	2	0	0	0	0
Limousin	7	0	1	1	0	0
Rhône-Alpes	69	4	2	2	1	2
Auvergne	46	3	0	0	0	0
Languedoc-Roussillon	10	1	0	0	0	0
PACA	7	0	0	0	0	0
Corse	0,8	0	0	0	0	0
Total	**1671,8**	**100**	**127**	**98**	**57**	**57**

L'annexe 1 regroupe la liste et les données épidémiologiques des souches incluses dans mon projet de thèse.

1.2 Souches humaines

La totalité des 302 souches cliniques de *S.* Derby isolées chez l'humain entre 2014-2015 ont été incluses dans mon projet de thèse. La **figure 35** détaille la répartition de ces souches.

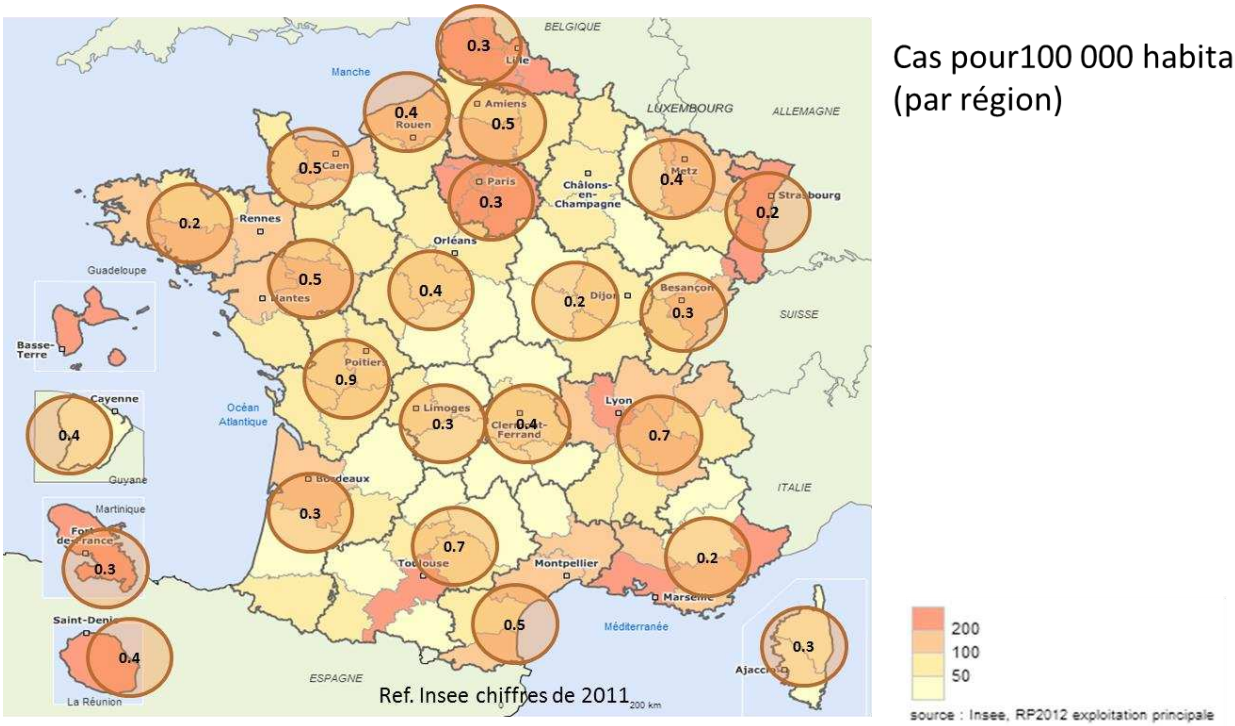


Figure 35 : Répartition des cas cliniques recensés par le CNR de *S.* Derby par région (période 2014-2015) au regard de la densité de population française.

2 Publication I : Polyphyletic nature of *Salmonella* enterica serotype Derby and lineage-specific host-association revealed by genome-wide analysis

2.1 Résumé

J'ai commencé par étudier la diversité génétique des souches alimentaires.

Afin de disposer d'un élément de comparaison avec les études précédentes (Hauser et al., 2011; Achtman et al., 2012), j'ai identifié les profils MLST des génomes des souches sélectionnées. J'ai également entrepris l'étude comparative des profils MLST identifiés pour *Salmonella* Derby avec les 24 autres sérovars de *Salmonella* les plus fréquemment isolés en France et en Europe (n=504 génomes au total) afin d'apprécier leur distribution (**tableau 8** et **figure 36**).

Tableau 8. Liste des sérovars utilisés pour l'analyse MLST et nombre des génomes correspondants.

Sérovars	Nombre de génomes
Agona	20
Anatum	19
Derby	112
Dublin	19
Enteriditis	20
Hadar	20
Havana	20
Indiana	8
Infantis	17
Kentucky	20
Livingstone	13
Mbandaka	18
Montevideo	18
Muenster	15
Newport	20
Oranienburg	17
Panama	5
Poona	15
Rissen	4
Saintpaul	17
Senftenberg	20
Tennessee	19
Typhi	16
Typhimurium	19
Virchow	13
Total	**504**

J'ai collecté 405 génomes dans des bases de données accessibles au public (NCBI: https://www.ncbi.nlm.nih.gov/ et Enterobase: https://Enterobase.warwick.ac.uk/), 61

génomes de la collection du réseau *Salmonella* (obtenus grâce au projet 100k Genome (http://100kgenome.vetmed.ucdavis.edu/) et au projet GenoSalmo (Anses Projet 2015-2016) et 38 génomes de *S.* Derby qui faisaient partie de cette étude de thèse.

L'ensemble des 504 génomes ont été vérifiés in silico via SISTR (https://lfz.corefacility.ca/sistr-app/) (Yoshida et al., 2016). Les résultats montrent que la population de *S.* Derby peut-être divisée en 3 lignées génétiquement distinctes (figure 36). L'analyse SNP montre d'ailleurs qu'il y plus de 15 000 SNP en moyenne de distance entre ces trois lignées génétiques. Les ST40 et 39 bien qu'ils forment une unique lignée, diffèrent entre eux de 3962 SNP en moyenne. Ces résultats démontrent le caractère polyphylétique de *S.* Derby.

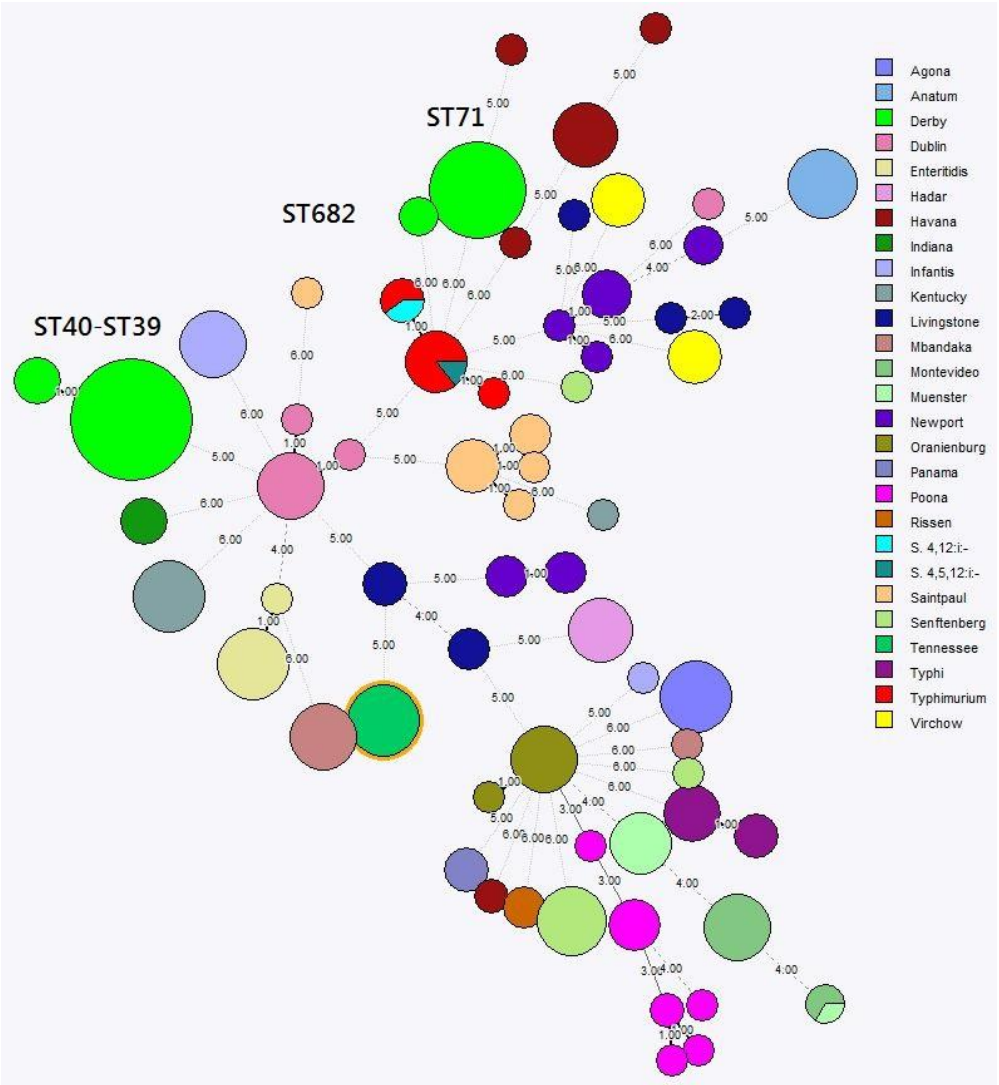


Figure 36 : Comparaison des profils MLST des 25 sérovars les plus fréquemment isolés en Europe.

Les souches du sérovar Derby se regroupent en 3 clusters de souches correspondants aux profils MLST ST40-ST39, ST71 et ST682. Les chiffres et l'épaisseur des traits correspondent au nombre de loci MLST différents entre les profils.

Les souches appartenant aux lignées ST39-ST40 et ST682 sont associées au secteur porcin avec le ST40 représentant la lignée majoritaire chez le porc. La lignée ST71 regroupe quant à elle les souches issues du secteur avicole.

La résistance des souches aux antibiotiques a été évaluée à la fois par prédiction des gènes de résistances et par diffusion sur disque et montrent l'existence d'un profil de résistance majoritaire aux aminosides, sulphonamides et tétracyclines chez un sous-groupe des souches du ST40. Les gènes codant pour cette résistance se trouvent sur un îlot génomique de *Salmonella* (SGI-1) de type 1 dans lequel les gènes codant pour la résistance au mercure sont aussi présents.

Les caractéristiques génomiques des différentes lignées ont été étudiées sur la base d'une comparaison de séquences des SPI-1 à -5, connus pour être présents chez toutes les salmonelles de la subsp*. enterica,* et du SPI-23 isolé spécifiquement chez *S.* Derby. Les séquences codante des SPI-1 à -5 présentaient une forte identité (supérieure à 98%) vis-à-vis de la référence *S*. Typhimurium chez l'ensemble de ces lignées à l'exception du SPI-3 pour lequel les ST39, ST40 et ST71 présentent une délétion des gènes *sugR* et *rhuM* en faveur d'un opéron fimbriale. Enfin, le SPI-23, caractéristique de S. Derby s'est avéré n'être présent que dans les lignées porcines ST39 et ST40, et totalement absent des souches des lignées ST71 et ST682.

Enfin, la variation allélique d'un marqueur de la spécificité à l'hôte : le gène *fimH*, codant pour le facteur d'adhésion des fimbriae de type 1, a été exploré. Le gène *fimH* de la lignée ST682 présente la même séquence que celui des sérovars *S.* Typhisuis et *S*. Choleraesuis spécifiques du porc, ce qui pourrait expliquer le tropisme de cette lignée pour la filière porcine malgré l'absence de SPI-23.

2.2 Publication

frontiers in Microbiology

ORIGINAL RESEARCH
published: 17 May 2018
doi: 10.3389/fmicb.2018.00891

Polyphyletic Nature of *Salmonella enterica* Serotype Derby and Lineage-Specific Host-Association Revealed by Genome-Wide Analysis

*Yann Sévellec[1], Marie-Léone Vignaud[1], Sophie A. Granier[1], Renaud Lailler[1], Carole Feurer[2], Simon Le Hello[3], Michel-Yves Mistou[1] and Sabrina Cadel-Six[1]**

[1] Université PARIS-EST, Agence Nationale de Sécurité Sanitaire de L'Alimentation, de L'Environnement et du Travail (ANSES), Laboratory for Food Safety, Maisons-Alfort, France, [2] French Institute for Pig and Pork Industry, Le Rheu, France, [3] Centre National de Référence des Salmonella, Unité des Bactéries Pathogènes Entériques, Institut Pasteur, Paris, France

In France, *Salmonella* Derby is one of the most prevalent serotypes in pork and poultry meat. Since 2006, it has ranked among the 10 most frequent *Salmonella* serotypes isolated in humans. In previous publications, *Salmonella* Derby isolates have been characterized by pulsed field gel electrophoresis (PFGE) and antimicrobial resistance (AMR) profiles revealing the existence of different pulsotypes and AMR phenotypic groups. However, these results suffer from the low discriminatory power of these typing methods. In the present study, we built a collection of 140 strains of *S.* Derby collected in France from 2014 to 2015 representative of the pork and poultry food sectors. The whole collection was characterized using whole genome sequencing (WGS), providing a significant contribution to the knowledge of this underrepresented serotype, with few genomes available in public databases. The genetic diversity of the *S.* Derby strains was analyzed by single-nucleotide polymorphism (SNP). We also investigated AMR by both genome and phenotype, the main *Salmonella* pathogenicity island (SPI) and the *fimH* gene sequences. Our results show that this *S.* Derby collection is spread across four different lineages genetically distant by an average of 15k SNPs. These lineages correspond to four multilocus sequence typing (MLST) types (ST39, ST40, ST71, and ST682), which were found to be associated with specific animal hosts: pork and poultry. While the ST71 and ST682 strains are pansusceptible, ST40 isolates are characterized by the multidrug resistant profile STR-SSS-TET. Considering virulence determinants, only ST39 and ST40 present the SPI-23, which has previously been associated with pork enterocyte invasion. Furthermore, the pork ST682 isolates were found to carry mutations in the *fimH* sequence that could participate in the host tropism of this group. Our phylogenetic analysis demonstrates the polyphyletic nature of the *Salmonella* serotype Derby and provides an opportunity to identify genetic factors associated with host adaptation and markers for the monitoring of these different lineages within the corresponding animal sectors. The recognition of these four lineages is of primary importance for epidemiological surveillance throughout the food production chains and constitutes the first step toward refining monitoring and preventing dispersal of this pathogen.

Keywords: *Salmonella* Derby, SNP analysis, AMR analysis, *Salmonella* pathogenicity island, FimH adhesin, host-association's genetic markers

OPEN ACCESS

Edited by:
Giovanna Suzzi,
Università degli Studi di Teramo, Italy

Reviewed by:
Craig T. Parker,
Agricultural Research Service (USDA), United States
Haijian Zhou,
National Institute for Communicable Disease Control and Prevention (China CDC), China

****Correspondence:***
Sabrina Cadel-Six
sabrina.cadelsix@anses.fr

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: *02 February 2018*
Accepted: *18 April 2018*
Published: *17 May 2018*

Citation:
Sévellec Y, Vignaud M-L, Granier SA, Lailler R, Feurer C, Le Hello S, Mistou M-Y and Cadel-Six S (2018) Polyphyletic Nature of Salmonella enterica Serotype Derby and Lineage-Specific Host-Association Revealed by Genome-Wide Analysis. Front. Microbiol. 9:891. doi: 10.3389/fmicb.2018.00891

INTRODUCTION

In the European Union, *Salmonella enterica* subspecies *enterica* serotype Derby (*S.* Derby) is the most abundant serotype isolated from pork. It accounts for 22.9% of all isolates, followed by monophasic strains of *S.* Typhimurium (22.3%) and *S.* Typhimurium (20.6%) (EFSA, 2016). In France, between 2000 and 2015 *S.* Derby ranked between 5th and 8th position (n = 164 to 178 clinical isolates) of the most frequently isolated serotypes in humans (Weill and Le Hello, 2014). In the entire food sector, the data of the ANSES *Salmonella* Network (jointly with the National Reference Laboratory) show that this serotype is the 4th most frequently isolated after *S.* Typhimurium, its monophasic variant *S.* 1,4,[5],12:i:- and *S.* Enteritidis (Leclerc et al., 2016). *S.* Derby is principally isolated from pork and poultry meat in France with a prevalence reaching 1.4% for pork and 3.2% for *Gallus gallus* (compared with <1% for turkey) (DGAl, 2014, 2015). All together, these data indicate that *S.* Derby is a significant threat to human health, mainly associated with the pork and poultry sectors.

This serotype is not exclusively adapted to pigs but most often associated with this source (Valdezate et al., 2005; EFSA, 2009). *S.* Derby was recently reported as the most common serotype from turkey flocks in the United Kingdom (∼50% of isolates from pigs and ∼40% from turkeys) (Hayward et al., 2016). In the United Kingdom, between 2014 and 2015 the number of notifications increased more than fivefold in turkey flocks (from 38 to 217 isolates, respectively), showing how well this serotype is adapted to this host (EFSA, 2016).

It is notable that two distinct lineages of *S.* Derby have been identified in the United Kingdom. They differ genotypically and phenotypically by the presence and absence of the *Salmonella* pathogenicity island 23 (SPI-23) and by the higher ability of strains possessing the SPI-23 to invade the porcine jejunum-derived cell line IPEC-J2 (Hayward et al., 2014). These two lineages seem to be adapted to distinct animal sources, probably pig and turkey, but the hypothesis cannot be confirmed because of the limited number of *S.* Derby isolates analyzed (n = 16) (Hayward et al., 2016). The presence of different *S.* Derby clonal groups prominent within the food chain was revealed previously in Spain, Germany, and France (Valdezate et al., 2005; Hauser et al., 2011; Kerouanton et al., 2013). These studies suffer, however, from the low discrimination potential of the technique used, pulsed field gel electrophoresis (PFGE) (Valdezate et al., 2005; Kerouanton et al., 2013). A study conducted by the ANSES *Salmonella* Network on a large panel of *S.* Derby strains showed that 42% of the *S.* Derby strains collected since 2005 were assigned to the same profile (SDBYXB0001 for 52/123 strains), highlighting the discrimination limits of PFGE for this serotype with an overall discrimination index of 0.75 (Kerouanton et al., 2007). Contrasting with PFGE studies, recent investigations based on multilocus sequence typing (MLST), clustered regularly interspaced short palindromic repeats (CRISPRs) and whole genome sequencing (WGS) analysis on a small selection of strains suggested that *S.* Derby should be considered as a polyphyletic serotype (Hayward et al., 2016; Zheng et al., 2017). Those studies, however, do not represent the total diversity of the serotype, either because of the limited number of strains analyzed or the low resolution given by the method used.

Considering the prevalence of *S.* Derby in the pork and poultry food sectors, we decided to thoroughly investigate the genetic diversity of this serotype using a WGS approach on a large dataset (n = 140) representative of the geographical and source origins in France. In contrast to conventional molecular typing methods, WGS has the potential to compare whole genomes at a single-nucleotide resolution. Methods based on single-nucleotide polymorphisms (SNPs) allow for a detailed, targeted analysis of variations among related bacterial isolates. WGS has recently been postulated to be an ultimate subtyping technique (Gilchrist et al., 2015; Dunn, 2016), and SNP-based cluster analysis was already successfully used to explore the genomic diversity of *Salmonella* isolates across serotypes as well as among and within specific food sources (Wilson et al., 2016; Ferrari et al., 2017).

The strains of this collection were isolated in 2014 and 2015 from the pork and poultry sectors, which are the main sources of this serotype, to obtain a comprehensive view of their distribution. The collection was investigated by MLST and SNP analysis. Because resistance of *Salmonella* to antimicrobial agents is a worldwide problem, and antimicrobial resistance (AMR) has already been described in *S.* Derby isolates (Valdezate et al., 2005; Hauser et al., 2011; Keelara et al., 2013; Kerouanton et al., 2013), susceptibility tests were performed and identification of acquired AMR genes was also investigated. The detection and characterization of *Salmonella* pathogenicity islands (SPI-1 to 5 and SPI-23), coding for virulence factors implied in adhesion and invasion of the host, and of the *fimH* gene, known to be a marker of the host specificity within *Salmonella*, were also investigated to identify potential genome signatures responsible for host specificity of the Derby serotype for pig and poultry.

MATERIALS AND METHODS

Strain Selection

A panel of 140 *S.* Derby strains was selected from all the isolates received by the *Salmonella* Network in 2014 and 2015 (n = 598) (Leclerc et al., 2016). The epidemiological data of the 140 *S.* Derby strains are listed in **Supplementary Table S1**. Duplicates (isolates with the same sampling date, geo-localization, and isolation matrix) and isolates from animals and feed were excluded. The strains selected came principally from the food sector (from slaughterhouses to the retail market). Only pork and poultry meat categories were considered, because of the low prevalence of *S.* Derby (<1%) in beef and cattle (Leclerc et al., 2016). Within the 140 strains, the proportion of strains from pork and poultry meat (85 and 60, respectively) was chosen in line with French production data. The production of pork and poultry meat accounts for 38.3 and 29.4% of French meat production, respectively (Menard et al., 2015). For each sector, the number of strains was selected, in each region, proportionally to its production compared to the total

French national production of meat products (**Supplementary Table S1** and **Supplementary Figure S1**). For poultry, since the *S.* Derby strains were concentrated in Brittany, all strains belonging to other regions were incorporated into the collection.

All strains were identified as belonging to the Derby serotype by glass slide agglutination, according to the White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007).

Genomic DNA Preparation and Sequencing

DNA was prepared from 10 ml of BHI overnight cultures with the Wizard® Genomic DNA Purification Kit (Promega, France) according to the manufacturer's instructions for gram-negative organisms. Gels of 1.5% agarose were used to assess the quality of the extraction (and an eventual degradation of the DNA). The DNA concentration was measured with a Qbit® fluorimeter and the purity ratio was assessed with a Nanodrop® Spectrophotometer. Libraries were prepared and the NGS sequencing were performed by the *Institut du Cerveau et de la Moelle épinière* (ICM)[1] (Hôpital de la Pitié-Salpêtrière, Paris). Each individual library (batch of 96 DNA) was prepared with the Nextera XT technology (Illumina). The indexing of the DNA was carried out during the construction of the library. The libraries were purified with the Agencourt AMPure XP system (Beckman Coulter) and quantification was performed using the Microfluidic LabChip GX (PerkinElmer).

The sequencing of the final library (DNA 96) by NextSeq 500 was carried out using a 300 cycle High Output kit v2 cartridge (400 million paired reads and 800 million single reads in 150 bases). Each Illumina paired-end sequence contained 300 base pairs (bp) (reads are 150 bases). The minimum theoretical coverage is of 30×–50×.

Multilocus Sequence Typing (MLST)

The seven housekeeping gene sequences (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*) for each strain were detected using the MLST service of the Center for Genomic Epidemiology (CGE)[2], which enabled us to determine the sequence type (ST) directly from the read files.

Single Nucleotide Polymorphism (SNP) Analysis

The SNP analysis was conducted using the VARCall workflow (Felten et al., 2017). In the absence of a complete reference *S.* Derby genome sequence and since VARCall requires a closed reference genome, sequence reads were mapped to *Salmonella* Typhimurium LT2 (NCBI NC_003197.1). The VCF files computed by VARCall were combined into a merged VCF which was filtered using Samtools and Picard tools to eliminate duplicated regions and variants solely linked to the reference genome. A pseudogenome obtained on the reference was generated using the Genome Analysis ToolKit (GATK) (McKenna et al., 2010), SNP and inDEL were predicted, and the distance matrix between each pair of genomes was calculated.

Phylogenetic analyses on the dataset were computed using RaXml software (Stamatakis, 2014). The phylogenetic trees were constructed under the maximum likelihood criterion using the GTR-gamma model of nucleotide evolution. The phylogenetic analyses were based on the pseudogenome obtained using the GATK. The phylogenetic data were visualized using interactive Tree Of Life (iTOL[3]) (Letunic and Bork, 2016).

Statistical Analyses

The non-normality of the data (number of paired SNP differences) was assessed using the Shapiro test (Royston, 1995) on R from the pairwise matrix generated by the VARCall workflow described above. The comparison between the paired SNP differences was tested by a Kolmogorov–Smirnov test (KS-test) (Huang et al., 2016), to find the variance of the distribution by paired SNP differences, which had been proven significantly unequal by the Fisher test (Markowski and Markowski, 1990).

Identification of Acquired Resistance Genes

The whole panel of genomes was analyzed using the ResFinder 2.1 application (Zankari et al., 2012) on the CGE server. The threshold for reporting a match between a gene in the ResFinder database and the input *S.* Derby genome was set at 90% identity over at least 3/5 of the length of the resistance gene. BioNumerics software version 7.6.1 (Applied Maths, Sint-Martens-Latem, Belgium) was used to perform a BLAST to localize each resistance gene inside the assembled genome. In order to investigate the implantation of the AMR gene within the genome of *S.* Derby, SGI-1 coding sequences (NCBI:AF261825.2) were extracted and BLASTed against the dataset with the BioNumerics BLAST tool. The complete genomic sequence of the SGI-1 was investigated using the BioNumerics alignment and sequence visualization tools.

Antimicrobial Susceptibility Tests

Strains were selected by their genetic distance: for the pork sector, all strains above a cut-off defined as the median genetic distance (110 SNPs) were selected (n = 40). The same cut-off (36 SNPs) was set to select ST71 strains from the poultry sector (n = 21).

Antibiotic susceptibility was determined using the disk diffusion method as recommended by the Clinical and Laboratory Standards Institute (CLSI) (CLSI, 2015, 2016). Fifteen antimicrobials (Bio-Rad, Marne-la-Coquette, France) were tested: amoxicillin/clavulanic acid (AMC; 30 µg), ampicillin (AMP; 10 µg), cephalothin (CEF; 30 µg), cefotaxime (CTX; 30 µg), ceftazidime (CAZ; 30 µg), chloramphenicol (CHL, 30 µg), sulfonamides (SSS; 300 µg), trimethoprim-sulfamethoxazole (SXT; 1.25+23.75 µg), streptomycin (STR; 10 U), gentamicin (GEN; 10 µg), kanamycin (KAN; 30 UI), tetracycline (TET; 30 UI), nalidixic acid (NAL; 30 µg), ciprofloxacin (CIP; 5 µg), pefloxacin (PEF; 5 µg). A Colistin disk

[1] www.icm-institute.org

[2] https://cge.cbs.dtu.dk/services/MLST/

[3] https://itol.embl.de/

(CST; 10 μg) was used on each plate for quality management purposes to ensure the absence of contamination. Automatic readings were performed using the BIOMIC® V3 system (Giles Scientific Inc., Santa Barbara, CA, United States). Isolates were classified as susceptible, intermediate, or resistant according to the clinical interpretive criteria recommended by CLSI (2016).

Salmonella Pathogenicity Islands (SPI) – Identification

SPI-1, the two segments of SPI-2, SPI-3, SPI-4, SPI-5, and SPI-23 were tested for within the set of 140 genomes of *S.* Derby using BLASTn[4] with a cut-off of 90% identity. The SPI reference sequences used for the BLASTs were collected by the NCBI database (accession numbers KP279311.1, AJ224978.1, KP258194.1, AF106566.1, KP234070.1, AY144492.1, and LAZB00000000-project PRJNA270707, respectively). All these sequences correspond to complete SPIs from *S.* Typhimurium strains with the exception of SPI-23, which corresponds to *S.* Derby strain 07CR553 (Kerouanton et al., 2015).

In the absence of a reference sequence for SPI-23, the primers defined by Hayward et al. (2016) were used to generate the *in silico* PCR and to extract a complete SPI-23 sequence from the genome of the *S.* Derby strain 07CR553 (accession number LAZB00000000), contig 5 (1894692..1931302) (Kerouanton et al., 2015). Both the complete sequence of SPI-23 obtained as described above and the 41 Coding DNA Sequences (CDS) described by Hayward were BLASTed separately against our dataset to verify the results of the *in silico* PCR.

Fimbriae FimH Allele Characterization

The fimbriae FimH allele, identified as host specificity marker by previous studies (Yue et al., 2015), was also sought and characterized within our panel of *S.* Derby genomes. The *fimH* gene sequence was isolated from the project PRJNA297164 (NCBI) and compared with the sequence of the dataset using BLASTn.

As an element of comparison, 25 *fimH* sequences from 25 different serotypes of *Salmonella enterica* subsp. *enterica* selected among the most frequently isolated in humans, animal, and food (Weill and Le Hello, 2014; EFSA, 2016) were added to this study (**Supplementary Table S2**). Whole FimH alleles were extracted by dataset, annotated using the BioNumerics 7.6.1 annotation plugin (Applied Maths, Sint-Martens-Latem, Belgium) and aligned on the reference *fimH* sequence of the *S.* Typhimurium strain SL1344 (accession NC_016810.1). The alignment was carried out using the BioNumerics 7.6.1 alignment tool and the mutations were identified using the mutation prediction tool of the same software.

RESULTS

MLST Profiles

Four different MLST profiles, ST40, ST39, ST71, and ST682, were identified among the 140 studied genomes. ST40 and ST39 were distinct by only eight mutations in the *sucA* locus and could therefore be included in the same eBURST complex (Achtman et al., 2012). The most frequent profile within the collection was ST40 ($n = 64$ genomes) followed by ST71 ($n = 58$), ST39 ($n = 13$), and ST682 ($n = 5$) (**Figure 1**). ST40 grouped together 61 strains isolated from pork meat (95.3%) and 3 from poultry (4.7%). The ST39 and ST682 strains were isolated from pork (13 and 5 strains, respectively). ST71 was represented by 54 strains isolated from poultry meat (92%) and 4 strains from pork meat (8%).

Phylogenetic Analysis of *S.* Derby Strains Panel

Four clusters were obtained by the SNP analysis of the 140 genomes of *S.* Derby analyzed (**Figure 1**). As shown in **Figure 1**, the SNP analysis clustered the 140 genomes in four groups. These four groups were fully consistent with the ones identified by MLST analysis (100% identity). To simplify the comprehension of the results, these 4 lineages were named with their ST profiles (ST39, ST40, ST71, and ST682).

Within each of these groups, genomes were found to differ by less than 300 SNPs. The genomes belonging to ST39 were most closely related to ST40 genomes with an average of 3,962 SNPs and a standard deviation (SD) of 20 SNPs. The strains belonging to the ST71 cluster were distant from the ones belonging to ST40 by 26,957 SNPs, with an SD of 1,583. The ST profile 682 was the most genetically distant from ST39, ST40, and ST71 with an average of 33,961 SNPs and an SD of 4,102. Considering epidemiological information, there was no evidence of a relationship between the genomic proximity

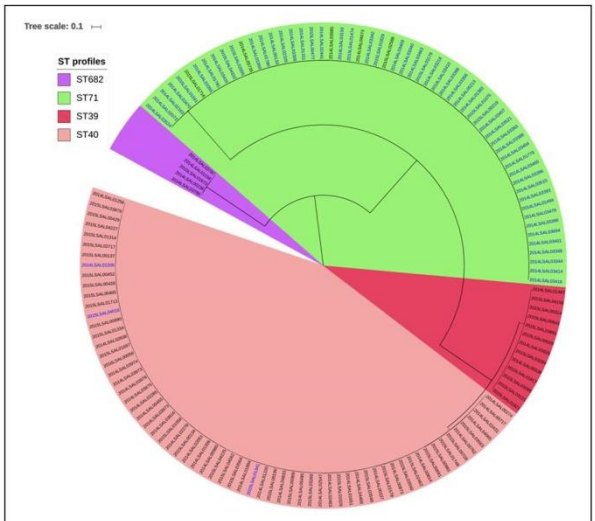


FIGURE 1 | Phylogenetic SNP tree for 140 strains of *S.* Derby constructed under the maximum likelihood criterion using the GTR-gamma model. The scale bar indicates the number of substitutions per site. Four clusters were identified and named ST39, ST40, ST71, and ST682 with reference to the MLST profiles with which the correspondence is perfect. The labels in blue correspond to the strains of the poultry sector. The black labels correspond to the strains of the pork sector.

[4]https://blast.ncbi.nlm.nih.gov/Blast.cgi

and geographical localization for the different lineages. Closely related strains could have very different geographical origins and vice versa. It is also the case for the ST39 and ST682 lineages, which grouped together only 13 and 5 genomes, respectively. These lineages presented a very wide geographic distribution, clustering together strains from Pays-de-la-Loire, Bourgogne, Languedoc-Roussillon and Centre. The Pays-de-la-Loire region, however, presented considerable diversity in terms of ST profiles with strains from the four different lineages (**Figures 2**, **3** and **Supplementary Figure S2**). Consistent with the distribution of the pork and poultry food production chains in France, ST40 was geographically more widespread than ST71. Our results clearly associated three lineages with the pork sector (ST39, ST40, and ST682) and one with the poultry sector (ST71).

Phylogenetic Analysis of the ST40 Cluster

As shown in **Figure 2**, the strains forming the ST40 cluster could be divided into two clades. Clade 1 contained 16 strains with an average of 192 SNPs and an SD of 83 SNPs while Clade 2 contained 48 strains with an average of 68 SNPs and an SD of 16. These two clades were separated by an average of 135 SNPs and an SD of 64. A Kolmogorov–Smirnov test between these two groups statistically confirmed the distinction between those two clades (p-value $\leq 2.2.10^{-16}$). There is no apparent correlation between geographical and phylogenetic distribution of the different strains belonging to these two clades.

SNP Inference for the ST71 Cluster

The ST71 cluster contained genomes presenting only an average of 28 SNPs and an SD of 10 (**Figure 3**). The statistical analysis showed that two clades could be identified (p-value = 6.10^{-13}) in the ST71 cluster. Between two genomes, in Clade 1 there was an average of 19 SNPs (SD of 5) while in Clade 2 there was an average of 34 SNPs (SD of 14). There was no apparent chronological relationship between the different strains belonging to these two clades. Considering the geographical distribution of the sampling (**Supplementary Figure S2**), the

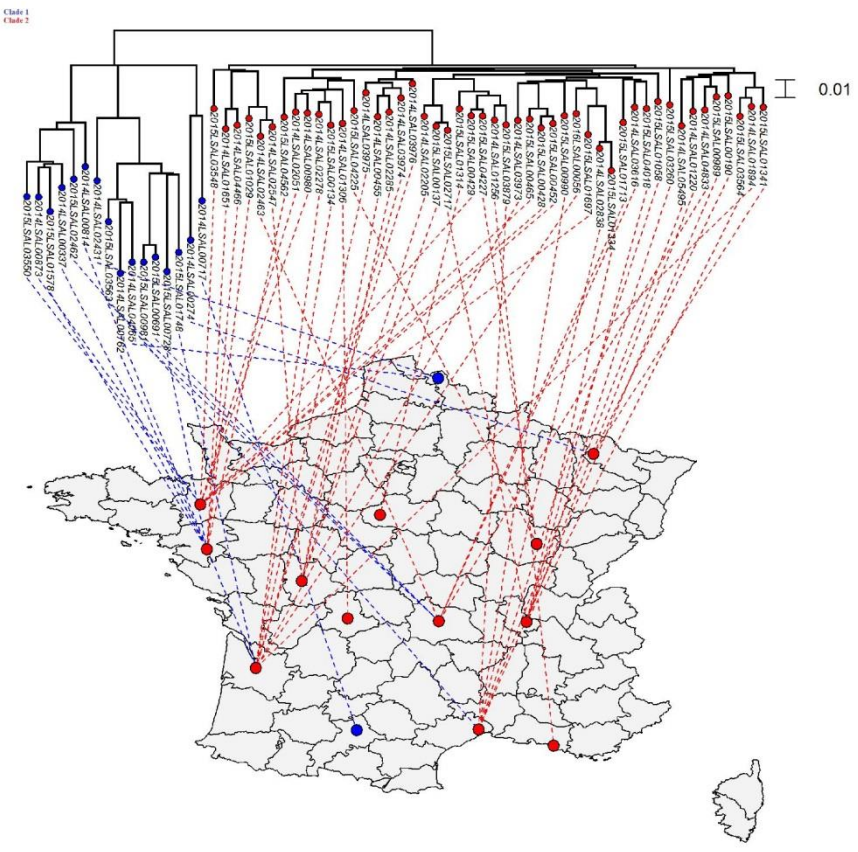


FIGURE 2 | Phylogenetic SNP distribution of the strains from the cluster ST40 in relation with their geographical repartition. Tree was constructed under the maximum likelihood criterion using the GTR-gamma model. The scale bar indicates number of substitutions per site. Two distinct clades were identified. There is no apparent correlation between geographical and phylogenetic distribution of the different strains belonging to these two clades.

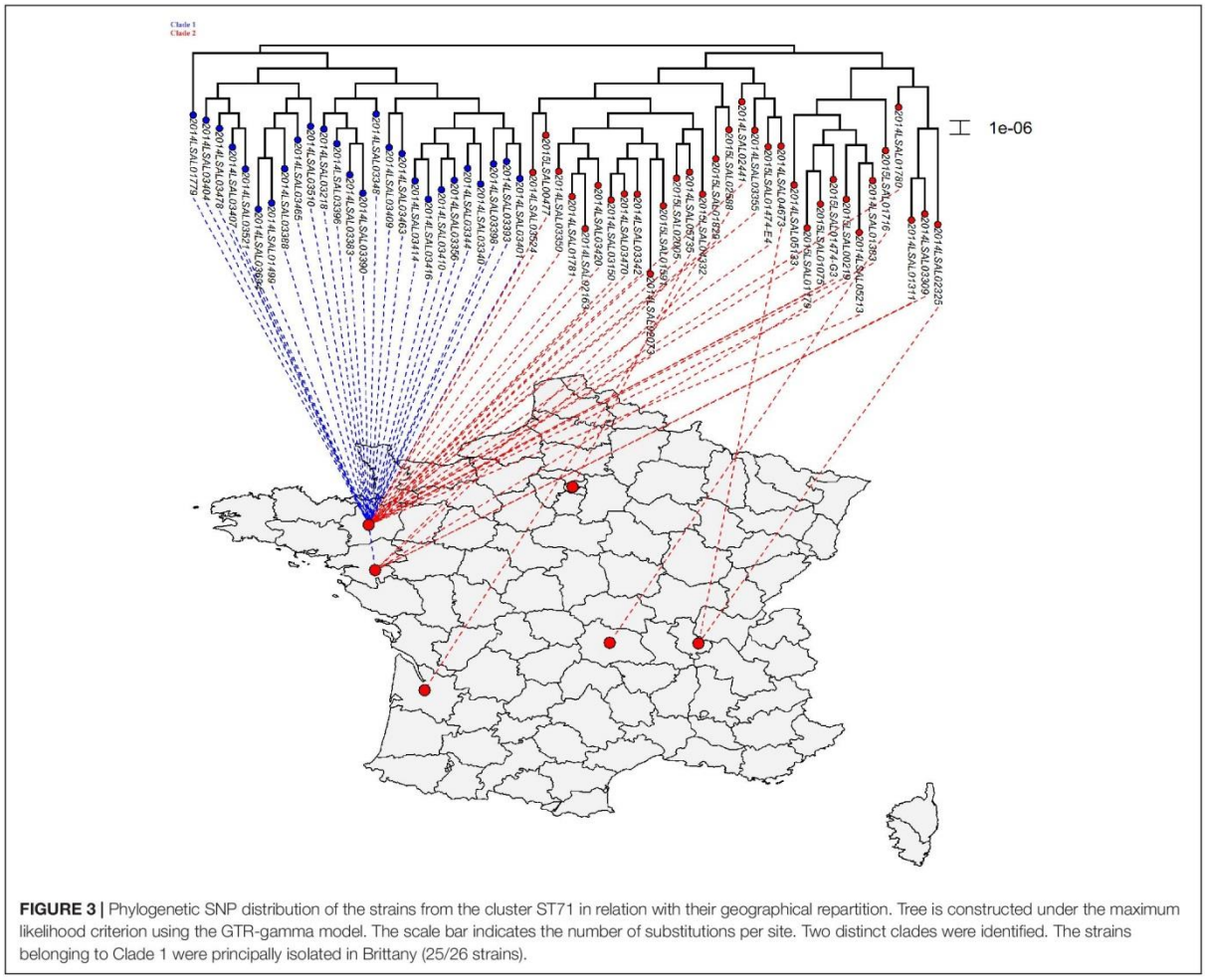


FIGURE 3 | Phylogenetic SNP distribution of the strains from the cluster ST71 in relation with their geographical repartition. Tree is constructed under the maximum likelihood criterion using the GTR-gamma model. The scale bar indicates the number of substitutions per site. Two distinct clades were identified. The strains belonging to Clade 1 were principally isolated in Brittany (25/26 strains).

strains belonging to Clade 1 were principally isolated in Brittany (25/26 strains).

Antimicrobial Susceptibility Tests and Detection of Acquired Resistance Genes

Complete and partial antibiotic resistance genes identified in the 140 *S.* Derby genomes studied are listed in **Supplementary Table S3**. No resistance gene was detected in 59% (83/140) of the studied genomes. Among the 57 genomes harboring AMR genes, 81% were simultaneously carrying *aadA2, sul1*, and *tetA* (**Figure 4** and **Supplementary Table S3**). These genes mediate resistance to aminoglycosides, sulfonamides, and tetracyclines, respectively. This expected phenotype had been confirmed phenotypically for 46 "STR-SSS-TET" strains belonging to Clade 2 of ST40 (**Figure 4**). The resistance genes carried by the strains of Clade 1 of ST40 were more diverse. Most of the ST71, ST39, and ST682 strains were found to be carrying no AMR genes, 97% (57/58), 100% (13/13), and 80% (4/5), respectively.

Distribution and Diversity of SPI-1, -2, -3, -4, -5

Table 1 compiles the information about the length, number of genes, and percentage of GC content in the different SPIs. SPIs 1 to 5 are present in the 140 genomes studied. The alleles are identical within each lineage (100% of identity) but can be different between lineages. ST40 and ST39 share the same alleles (100% identity) with the exception of 6 genes in SPI-5 (*pipB*, *pipC*, *pipD*, *sopB*, *copS*, and *copR*). The global conservation of SPIs 1 to 5 is high between the four Derby lineages, with an overall sequence identity above 97% and allelic difference that does not exceed four substitutions. The genes *sseB*, *sseC*, and *sseD* of the second region of SPI-2 present the highest allelic difference between the lineages, with a sequence identity between 94.91 and 97.80%. The gene *rhuM* at the left-hand end of SPI-3 is only present in ST682. This gene (coding for a hypothetical protein) and *sugR* (coding for a putative ATP binding protein) are absent in ST40 and ST39 and only present as a fragment in ST71 (corresponding to 1..1246 out of a 1560 bp long gene) and

FIGURE 4 | Phylogenetic SNP tree of the strain from the cluster ST40. The scale bar indicates the number of substitutions per site. The box in blue corresponds to Clade 1 and the box in red to Clade 2. The predicted antibiotic resistance genes are indicated on the right. The red cubes correspond to the resistance gene to aminoglycosides antibiotic, the blue circle to beta-lactam, the green triangle to sulfonamides, the purple triangle to tetracycline, and the orange circle to trimethoprim.

ST682 (one fragment of 520 bases corresponding to the beginning of the gene *sugR* and a second fragment of 535 corresponding to the end of the gene *sugR*). An integron previously described by Amavisit et al. (2003), containing seven genes related to the adhesion structures, pili and fimbriae (*ecpD2, ecpD1, htrE, yadM, yadL, yadK,* and *yadC*), flanked by two truncated transposase genes (*tpase2* and *tpase1*), was detected in the SPI3 of ST71, ST40, and ST39.

Distribution and Diversity of the SPI-23 Region

SPI-23 was only present in the sequences of *S.* Derby belonging to ST40 and ST39 with a sequence identity of 100% or above to 99.9% (for a 100% coverage). This *Salmonella* pathogenicity island was missing in ST682 strains and in most ST71 strains with the exception of strain 2014LSAL05133, which only contains fragments of SPI-23. The *docB* gene, reported to end SPI-23, was present in all strains. We do not believe that the assembly process could alter the reconstruction of SPI-23, because all ST39 and ST40 strains were found to carry the 41 CDS (with 100% of identity) constitutive of SPI-23 and used as references for BLAST analysis. The results obtained by *in silico* PCR with the primers described by Hayward et al. (2016) are shown in **Supplementary Figure S2**.

FimH Gene Alleles

FimH amino acidic sequences of the four *S.* Derby lineages obtained in the present study were compared with those of strains previously isolated from swine, poultry, and cattle and

TABLE 1 | Length, genes, and GC content of the main SPIs for *Salmonella*.

Salmonella pathogenicity island	Length (Kb)	GC%	Total number of genes	ST40 and ST39	ST71	ST682	Range of identities (%) between the four lineages
SPI-1	38	45.77	39	39/39	39/39	39/39	98.01 to 100
SPI-2 region 1	12	53.05	10	10/10	10/10	10/10	98.10 to 100
SPI-2 region 2	25	40.00	31	31/31	31/31	31/31	94.91 to 100
SPI-3	16	51.42	10	8/10[a]	9/10[b]	9/10[c]	95.43 to 100
SPI-4	27	45.18	10	10/10	10/10	10/10	97.70 to 100
SPI-5	9	45.50	8	8/8	8/8	8/8	97.27 to 100
SPI-23	36	37.91	41	41/41	0/41[d]	0/41	99.90 to 100

[a] *Missing sugR and rhuM;* [b] *Truncated sugR and missing rhuM;* [c] *Missing sugR (fragment only);* [d] *Except the strain 2014LSAL05133 which present fragments of SPI-23.*

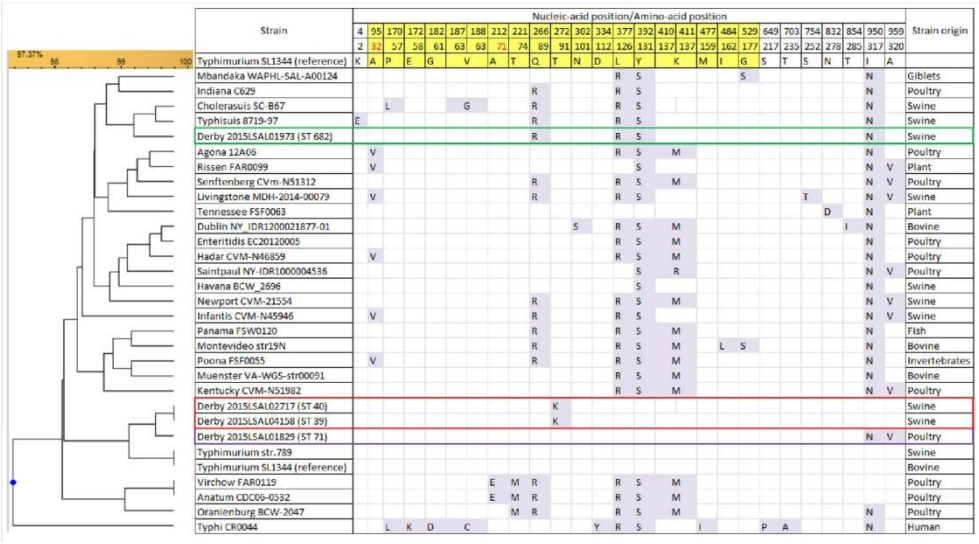


FIGURE 5 | Variant residue positions in the FimH protein sequence for 25 serotypes and some *S.* Derby strains of this study. The results show a 100% identity of the *fimH* sequences inside each ST lineage. Only three strains of each group are represented in the comparison, as examples. The residues belonging to the lectin domain (23–196, yellow background) and the pilin domain (199–335, white background) are framed, for a total of 27 missense SNPs (23 already described and 4 new ones) for a total of 17 mutations of the lectin domain and 9 of the pilin domain. The residues in red are known to be part of the mannose binding pocket.

belonging to the most frequent serotypes isolated in Europe (Weill and Le Hello, 2014; EFSA, 2016). The results show that there are three different *fimH* alleles in our collection, grouping together ST39 and ST40 isolates, while ST71 and ST682 display a peculiar *fimH* allele-type. FimH protein is characterized by two domains, the lectin and pilin domains. The sequence of the lectin domain of the strains belonging to the poultry-associated ST71 shows 100% identity with that of the *Salmonella* Typhimurium reference strain SL1344. The sequences of strains belonging to ST39–ST40 and ST682 differ by one and two substitutions, respectively, from the SL1344 sequence. The sequence of the pilin domain of T39-ST40 isolates differs by one substitution at position 91 from the SL1344 sequence. The ST71 FimH sequence differs by two missense substitutions from ST39–ST40 and three substitutions from ST682 (**Figure 5**). Considering FimH sequence variations among the 25 most prevalent serotypes, *S.* Derby ST39–ST40 and ST71 are closely related to *S.* Typhimurium, while ST682 is closely related to *S.* Choleraesuis and *S.* Typhisuis with (4/6 and 4/5, respectively) missense mutations and (15/19) silent mutations in common. The residues predicted to be part of the mannose-binding pocket (positions 32 and 71) are conserved in the genomes of all *S.* Derby lineages. The unique substitution found in ST40 and ST39 and not encountered in the sequences of the other serotypes is located at position 272 and is predicted to change amino acid residue 91 from Threonine to Isoleucine. The ST71 has two missense mutations in position 950 (changing residue 317 from Threonine to Asparagine) and 959 (changing residue 320 from Alanine to Valine). Otherwise, the three ST groups ST40, ST39, and ST71 share 6 identical silent mutations. The ST682 clade is more distant from the *fimH* sequence of *S.* Typhimurium SL1344 and displays 16 silent mutations and

4 missense mutations at positions 266 (changing residue 89 from Quinine to Arginine), 377 (changing residue 126 from Leucine to Arginine), 392 (changing residue 131 from Tyrosine to Serine), and 950 (changing residue 317 from Threonine to Asparagine).

DISCUSSION

Genetic Diversity of *S.* Derby

The collection of *Salmonella* Derby strains used in this study was chosen to be representative of the diversity in the French food sector. The strains were essentially collected in the west/south-west parts of France corresponding to the most intensive pork and poultry meat production regions.

The phylogenetic analysis reveals that four major *Salmonella* Derby genetic lineages cohabit in France, corresponding to the MLST profiles ST40, ST39, ST682, and ST71. Two of these lineages, ST39 and ST40, differ by 3,962 SNPs (SD of 20 SNPs). This genetic proximity is confirmed by considering their MLST alleles, which can be grouped into the same MLST eBURST complex. The other lineages are much more distant from each other and display an overall inter-group difference of 14,866 SNPs with an SD of 12,961. ST682 is the most genetically different, with an average distance of 33,961 SNPs (SD of 4,102) from the three other groups.

These results highlight that strains displaying the same antigenic pattern *S.* 1,4,[5],12: f,g: [1,2] according to the White-Kauffmann-Le Minor scheme (Grimont and Weill, 2007) and consequently grouped under the same *Salmonella* Derby serotype include at least four genomic lineages. Like studies on *Salmonella* Newport (Timme et al., 2013), our results obtained on a large number of strains provide another example of a polyphyletic serotype.

The four MLST profiles, ST39, ST40, ST71, and ST682, have already been described (Hauser et al., 2011; Wang et al., 2011; Cai et al., 2016; Hayward et al., 2016; Zheng et al., 2017). Hauser et al. (2011) reported ST682 strains in Germany that were involved in a 2013–2014 outbreak associated with the consumption of pork (Simon et al., 2017). Isolation of ST40 strains has been previously described in England, Germany, Midwestern United States, and China, making this ST the most frequently reported in the literature (Hauser et al., 2011; Wang et al., 2011; Cai et al., 2016; Hayward et al., 2016; Zheng et al., 2017). Recently, Zheng et al. (2017) described a collection of 92 strains isolated from the swine and pork production chain in China, almost one third of which belong to ST71. This result contrasts with ours where all ST71 strains were associated with the poultry sector and none was isolated from the pork sector. However, Chinese isolates came from slaughterhouses and retail markets and never from the swine and farm environment, leading to the hypothesis that cross-contamination may have occurred on sites. Otherwise the ST71 was described by Hayward et al. (2016) and associated with turkey meat.

Our work is the first large survey at the serotype level taking into account the different sectors with which *Salmonella* Derby has been associated and that clearly demonstrate the polyphyletic nature of the Derby serotype by whole genome analysis. It further demonstrates through a large national sampling the strong host-association of the four identified lineages. Within our panel the lineages ST40, ST39, and ST682 are associated with pork while ST71 is associated with poultry. Our results were obtained on a large national collection presenting a wide diversity of sampling origins. It can, however, be considered limited due to the time scale (2 years) and the restricted national area and, of course, it does not preclude that additional, uncharacterized lineages could express the Derby antigens. However, our collection brings together in a coherent way, all the genetic diversity that was previously described in a series of separate studies (Hauser et al., 2011; Wang et al., 2011; Cai et al., 2016; Hayward et al., 2016; Zheng et al., 2017).

The phylogenetic analysis carried out on ST40 and ST71, regrouping 87% of the strains analyzed, allowed us to further discriminate four, and statistically supported, clades. The ST40 can be divided into two clades. Clade 1 is highly heterogeneous regarding the antimicrobial profile; it contains susceptible strains and a diversity of antimicrobial-resistance profiles, while Clade 2 is characterized by strains showing a highly homogeneous AMR pattern, STR-SSS-TET (streptomycin, sulfonamides, and tetracycline). The strains clustered in Clade 2 are genetically less distant than the strains of Clade 1, with an average of 68 (SD of 16) and 192 (SD of 83) SNPs, respectively. ST71 can also be divided into two statistically different clades with no differences concerning AMR. The strains belonging to these two clades are essentially pansusceptible.

The ST40 lineage is found throughout France and strains from Clades 1 and 2 were isolated in various links of the food chain: slaughterhouses, secondary processors, retail sale, and food. In light of our results, it would be interesting to acquire information on the genetic diversity of *S.* Derby at the farm level, which is currently lacking. There is a widely held hypothesis that pork becomes contaminated at slaughterhouses supported by several publications (Rostagno et al., 2003; Botteldoorn et al., 2004; Valdezate et al., 2005; Buncic and Sofos, 2012). A recent study by Fois et al. (2017), on the occurrence and the characterization of *Salmonella* strains in slaughtered pigs in Italy, shows self-contamination for 71.5% of *Salmonella*-positive carcasses. Contaminated tools used for slaughtering can participate in the dissemination of this pathogen between carcasses. Contamination can also occur by the use of contaminated tools and work surfaces in the meat-cutting workshops. Above all, an understanding of the modes of transmission within the food production chains and the monitoring of the different lineages throughout the pork and poultry sectors would help national and international health, food, and agricultural authorities to establish suitable hygiene practices against the spread of this pathogen.

The main geographical origin of ST71 is from Brittany and Pays-de-la-Loire, where 76% of the poultry industry in France and 80% of the slaughterhouses are concentrated.

Our phylogenetic analysis does not differentiate Derby strains isolated from turkey from those isolated from *Gallus gallus*. The similar prevalence of Derby in the two animal species, 12.6% for *Gallus gallus* and 14.2% for turkey (DGAl, 2014, 2015), suggests that the ST71 strains are well adapted to both *Gallus gallus* and turkey. On the other hand, these data underline the disproportion observed in France between the prevalence of Derby upstream of the food chain (in animals) and the contamination recorded downstream (in food). In our panel representative of the food production chain in France, two strains were isolated from the turkey food sectors and 54 from the *Gallus gallus* food sector. This discrepancy could be explained by the fact that 81% of the production in the French poultry sector concerns *Gallus gallus* and only 5% turkey, and that the farm environment is often common for the two species, so that cross contamination should not be excluded.

Antimicrobial Susceptibility

Antimicrobial resistance is strongly associated with lineage ST40 in this study, as only 14% of the strains belonging to this lineage are pansusceptible. Most of ST71, ST39, and ST682 strains have no known resistance gene. Clade 2 of ST40 includes the majority of the strains presenting the STR-SSS-TET profile (n = 46/48). The STR-SSS-TET profile was first described in Spain (Valdezate et al., 2005). It has also been described in France by Kerouanton et al. (2013) for strains isolated from pig, pork, and humans (accounting for 66.7% of the isolates collected in 2006–2007). Aminoglycosides, sulfonamides, and tetracyclines are the most used antimicrobial classes in the pig sector in France (Binh et al., 2009; Méheust et al., 2016). The exploration of the genome of the ST40 Clade 2 strains revealed the presence between the *trmE* and *yidY* genes of a *Salmonella* genomic island (SGI-1) specific to this clade. This SGI-1 contained a class1 integron delimited by *intI1* and IS1326 insertion sequences. This integron is similar to the class1 integron containing the *aadA2* and *sul1* genes of the SGI-1C described by Beutlich et al. (2011). The SGI-1 also included the *tetA* gene and a cluster of mercuric resistance genes (*merA*, *merC*, *merP*, *merT*, and *merR*) located in a Tn7 transposon.

Interestingly, several studies have shown that the soil plays an important role in the dissemination within microorganisms of IncN, IncW, IncP-1, and pHHV216-like plasmids carrying resistance genes (Binh et al., 2008). The practice of field application of piggery manure, which harbors a substantial reservoir of broad-host-range plasmids conferring multiple antibiotic resistance genes, has been demonstrated to be responsible for this dissemination into agricultural soils, favoring horizontal gene transfer (Binh et al., 2009). Clinically relevant Class 1 integrons are also introduced into soil via similar practices (Aminov, 2011; Colavecchio et al., 2017; Pornsukarom and Thakur, 2017).

Finally, 98% (56/57) of the experimental AMR phenotypes were in agreement with the predictions made by ResFinder. One strain (2014LSAL04065) contains a gene coding for trimethoprim resistance (*dfrA12*), as trimethoprim has only been tested in combination with sulphonamides (SXT), the expression of this trimethoprim resistance phenotype hasn't been validated. Our results highlight the complementarity of these two analyses, the conjunction of WGS and phenotypic approaches could indeed be able to detect new AMR mechanisms. These affirmations are concordant with the conclusions exposed in the EUCAST review for the antimicrobial susceptibility of *Salmonella* (Ellington et al., 2017). A WGS approach can provide rapid identification for well-known and characterized AMR mechanisms. However, no WGS approach is so far able to predict antimicrobial susceptibility.

Virulence Factors and Host Specificity

The results concerning the *Salmonella* pathogenicity island -1, -2, -3, -4, -5 showed that all these genomic regions were detected within the panel of *S.* Derby strains analyzed, with an average sequence identity of 97%. The observed sequence differences between the lineages for the genes *sseB*, *sseC,* and *sseD* coding for the SPI-2 translocon that influences the capacity of type III secretion system (T3SS) protein to invade host cell cytoplasm, could impact the ability of the four *S.* Derby lineages to efficiently multiply in host cells and to survive in macrophages (Wilson et al., 2007; Reynolds et al., 2014). Even if the substitutions were conserved inside each lineage and could be used to discriminate between the different lineages of *S.* Derby, only cellular test experiments could confirm this hypothesis. The SPI-3 was the most variable SPI within the genomes analyzed. We observed a deletion of the left-terminal genes *rhuM* and *sugR* that had been previously reported in several studies as a common deletion in several serotypes such as Derby, Infantis, Virchow, Havana, Newport, and Albany (Beutlich et al., 2011; Figueiredo et al., 2015; McWhorter and Chousalkar, 2015). The *rhuM* gene was present only in the lineage ST682. It is known that *rhuM* gene deletion causes a significant decrease of the epithelial cell invasion capacity (Tenor et al., 2004). The *sugR* gene was found partial or fragmented in all the lineages. ST40, ST39, and ST71 possessed an integron in their SPI-3 containing seven genes related to the adhesion structures, pili and fimbriae genes that were reported as characteristic of *S.* Derby (Amavisit et al., 2003). These results highlight the genomic difference between ST682 and the other *S.* Derby lineages, but as a whole it could be concluded that the genetic distinction between the four Derby lineages does not reside in these pathogenicity islands.

The *Salmonella* pathogenicity island 23 (SPI-23) described by Hayward et al. (2013) has been shown to play a role in adhesion and invasion of porcine tissues (Hayward et al., 2014). The presence of this particular SPI helps explain the host pig specificity of the strains belonging to the lineages ST39 and ST40. The SPI-23 is absent in the strains belonging to the lineage ST71 associated with poultry and in the lineage ST682 related to pig. It has been shown that in the United Kingdom two distinct lineages of *S.* Derby coexist and these two lineages seem to be adapted to distinct sources (pig and turkey), being distinguished by the presence and absence of SPI-23 and the ability to more

efficiently invade the porcine jejunum derived cell line IPEC-J2 (Hayward et al., 2016). The results of this study support the hypothesis that the differences in host ranges of *S.* Derby are adaptations to pathogenesis, environmental persistence, and the use of metabolites abundant in their respective host environments.

FimH adhesin, located on type 1 fimbriae, mediates the adhesion to gut tissues and colonization of the alimentary tract of the host, an important stage in the pathogenesis of *Salmonella* (Grzymajlo et al., 2013). A previous study (Yue et al., 2015) suggested a possible correlation between the allelic variation of the amino acid sequence of the FimH protein and bacterial host specificity. FimH protein is involved in regulation of length and mediation of adhesion of type 1 fimbriae and consists of a peptide signal of 22 residues followed by a lectin domain (residues 23–196) and a pilin domain (residues 199–335) connected by a 3-residue link (Yue et al., 2015). In the whole dataset, 19 different missense substitutions have been detected in the lectin domain and 8 different missense substitutions in the pilin domain compared to *S.* Typhimurium's SL1344 sequence. Although most of the substitutions were reported previously by Yue et al. (2015), 5 new substitutions (residues 91, 112, 113, 252, and 278) were identified in this study. The substitution of the residue 91 in the lectin domain was notably specific to the lineages ST40 and ST39. This new substitution was the only variation that differentiates the FimH protein sequence of the *S.* Derby lineages ST40 and ST39 from that of the *S.* Typhimurium SL1344. The impact of this mutation on host specificity could be investigated by invasion test on epithelial cells. The lineage ST682 shared four identical substitutions (Q89R, L126R, Y131S, and I317N) with *S.* Typhisuis and *S.* Choleraesuis, two serotypes well known for their pathogenesis in pig (Barrow and Methner, 2013). *S.* Typhisuis and *S.* Derby FimH protein sequences differed by only one missense mutation in the peptide signal (K2L) located before the lectin domain (**Figure 5**). FimH adhesin carrying these three substitutions have been shown to present a higher specificity for porcine enterocyte IPEC-J2 than those of *S.* Typhimurium (Yue et al., 2015). The mutations in the ST71 FimH alleles did not relate to any host-specific profiles that have been described in previous study (Yue et al., 2015).

The differences observed between the *fimH* alleles within the different *S.* Derby lineages and other significant SPI substitutions described above, could be used for developing specific real time PCR probe assays aimed at identifying and following the spread of these four *S.* Derby lineages throughout the pork and poultry sectors, from farm to fork.

CONCLUSION

The results of this study show that *Salmonella* Derby serotype in France is polyphyletic and can be divided into four distinct lineages distinguished by 14,866 SNPs on average (SD of 12,961). Our results indicate that it should be possible to develop specific and sensitive molecular markers for each lineage. The lineages correspond to the four MLST profiles that have been independently described for *S.* Derby. The dominant lineage in France corresponds to ST40, which is associated with the pork sector. Two other lineages associated with the pork sector are ST39 and ST682. The strains belonging to ST71 are associated with the poultry sector and were isolated from *Gallus gallus*, turkey, but also from duck and guinea fowl food products. These four lineages differ principally by sequence differences in part of the SPI-3, by different allelic composition of the SPI-2 and -5, and by the presence of the pork invasion-related SPI-23 that characterizes the lineages ST40 and ST39. The *S.* Derby strains belonging to ST40 also harbor several AMR genes that are absent from the other STs. Clade 2 of this lineage is characterized by the AMR pattern STR-SSS-TET and likely corresponds to the major PFGE profile described previously by Kerouanton et al. (2013) (profil SDX01). The STR-SSS-TET profile is associated with the presence of an SGI-1 containing a class 1 integron sequence and a mercury resistance gene cluster. Mutations in the lectin and pilin domain of the FimH adhesion protein clearly characterize the different *S.* Derby lineages, and the ST682 presents the same protein lectin and pilin domain sequences than *S.* Typhisuis, causing acute infections in pig.

In this study, the whole-genome-sequencing approach was carried out to provide a high-resolution molecular typing of *S.* Derby strains isolated from the pork and poultry sectors, which can be used to further investigate the host specificity of the four lineages identified. This study constitutes the baseline for identifying over time which genetic lineages were and are present in the livestock and farm environment in France, and will contribute to our understanding of how these lineages can be transmitted to the food industries. Considering source attribution, our data constitute a strong basis for determining which *S.* Derby lineages are responsible for human contamination in France. In that last perspective, we are currently analyzing the genomes of a panel of *Salmonella* Derby isolated from human clinical cases during 2014 and 2015.

AVAILABILITY OF DATA

Genomics sequence assemblies used in this project are available online on the NCBI network under the accession: PRJNA391404 (available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA391404). Details for each genomic assembly are summarized in **Supplementary Table S2**, including the accession codes for each genome.

AUTHOR CONTRIBUTIONS

SC-S piloted and administered the project. SC-S and YS designed and developed the experiments. YS, SG, and M-LV carried out the experiments and the analyses. SC-S, M-YM, RL, and CF provided acquisitions. YS, SC-S, SG, and M-YM drafted the manuscript. SLH and CF participated in the discussion and reviewed the report.

FUNDING

This study was supported by funding by the French Ministry of Agriculture, Food and Forestry, by the *Salmonella* Network, part of the ANSES-Laboratory for Food Safety (France), and by INAPORC. YS is the recipient of a doctoral fellowship (DGER-ANSES) cofinanced by AgroParisTech and the French Agency for Food, Environmental and Occupational Health & Safety (ANSES).

ACKNOWLEDGMENTS

We would like to thank Nicolas Randomski and Arnaud Felten for advice on data interpretation, Laurent Guillier for advice on collection selection and statistical analysis. We thank Neelanzaly Garounnamourdy for technical assistance and Véronique Noel for prevalence data for *S.* Derby in the food sector.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.00891/full#supplementary-material

FIGURE S1 | Geographical selection plan for the strains isolated from the pork and poultry sectors.

FIGURE S2 | Results of the *in silico* PCR for the SPI-23 into *S.* Derby collection. The *docB* gene corresponds to the coding sequence located immediately after the SPI-23. Only 2014LSAL05133 present small fragments of the SPI-23 (not corresponding to the primers highlighted in this figure) in the ST71.

TABLE S1 | Strains information. Sheet 1: Epidata, sample ID, NCBI accession number, genome deep coverage, and breadth. Sheet 2: repartition of the selected strains in each French region. Sheets 3 and 4: selection plan for the strains isolated from the pork and poultry sectors respectively.

TABLE S2 | List of the genomes from different serotypes of *Salmonella enterica* subsp. *enterica* used for the FimH sequence analysis. Epidata and sample ID.

TABLE S3 | List of strains presenting antimicrobial resistance genes.

REFERENCES

Achtman, M., Wain, J., Weill, F. X., Nair, S., Zhou, Z., Sangal, V., et al. (2012). Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathog.* 8:e1002776. doi: 10.1371/journal.ppat.1002776

Amavisit, P., Lightfoot, D., Browning, G. F., and Markham, P. F. (2003). Variation between pathogenic serovars within *Salmonella* pathogenicity islands. *J. Bacteriol.* 185, 3624–3635. doi: 10.1128/jb.185.12.3624-3635.2003

Aminov, R. I. (2011). Horizontal gene exchange in environmental microbiota. *Front. Microbiol.* 2:158. doi: 10.3389/fmicb.2011.00158

Barrow, P., and Methner, U. (2013). *Salmonella in Domestic Animals*. Wallingford: CAB International. doi: 10.1079/9781845939021.0000

Beutlich, J., Jahn, S., Malorny, B., Hauser, E., Huhn, S., Schroeter, A., et al. (2011). Antimicrobial resistance and virulence determinants in European *Salmonella* genomic island 1-positive *Salmonella enterica* isolates from different origins. *Appl. Environ. Microbiol.* 77, 5655–5664. doi: 10.1128/aem.00425-11

Binh, C. T., Heuer, H., Kaupenjohann, M., and Smalla, K. (2009). Diverse *aadA* gene cassettes on class 1 integrons introduced into soil via spread manure. *Res. Microbiol.* 160, 427–433. doi: 10.1016/j.resmic.2009.06.005

Binh, C. T., Heuer, H., Kaupenjohann, M., and Smalla, K. (2008). Piggery manure used for soil fertilization is a reservoir for transferable antibiotic resistance plasmids. *FEMS Microbiol. Ecol.* 66, 25–37. doi: 10.1111/j.1574-6941.2008.00526.x

Botteldoorn, N., Herman, L., Rijpens, N., and Heyndrickx, M. (2004). Phenotypic and molecular typing of *Salmonella* strains reveals different contamination sources in two commercial pig slaughterhouses. *Appl. Environ. Microbiol.* 70, 5305–5314. doi: 10.1128/AEM.70.9.5305-5314.2004

Buncic, S., and Sofos, J. (2012). Interventions to control *Salmonella* contamination during poultry, cattle and pig slaughter. *Food Res. Int.* 45, 641–655. doi: 10.1016/j.foodres.2011.10.018

Cai, Y., Tao, J., Jiao, Y., Fei, X., Zhou, L., Wang, Y., et al. (2016). Phenotypic characteristics and genotypic correlation between *Salmonella* isolates from a slaughterhouse and retail markets in Yangzhou, China. *Int. J. Food Microbiol.* 222, 56–64. doi: 10.1016/j.ijfoodmicro.2016.01.020

CLSI (2015). *Performance Standards for Antimicrobial Disk Susceptibility Tests; Approved Standard. CLSI Document M02-A12*. 12th Edn. Wayne, PA: Clinical and Laboratory Standards Institute.

CLSI (2016). *Performance Standards for Antimicrobial Susceptibility Testing; CLSI Supplement M100S*. 26th Edn. Wayne, PA: Clinical and Laboratory Standards Institute.

Colavecchio, A., Cadieux, B., Lo, A., and Goodridge, L. D. (2017). Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the *Enterobacteriaceae* family – a review. *Front. Microbiol.* 8:1108. doi: 10.3389/fmicb.2017.01108

DGAl (2014). *Bilan 2014 des Plans de Surveillance et de Contrôle. Surveillance Sanitaire des Denrées Animales et Végétales en France*. Boulder, CO: DGAI, 69–74.

DGAl (2015). *Bilan 2015 des Plans de Surveillance et de Contrôle. Surveillance Sanitaire des denrées Animales et Végétales en France*. Boulder, CO: DGAI, 112–125.

Dunn, J. R. (2016). Whole-genome sequencing: opportunities and challenges for public health, food-borne outbreak investigations, and the global food supply. *J. Infect. Dis.* 213, 499–501. doi: 10.1093/infdis/jiv298

EFSA (2009). Analysis of the baseline survey on the prevalence of *Salmonella* in holdings with breeding pigs in the EU, 2008 - Part A: *Salmonella* prevalence estimates. *EFSA J.* 7:1377. doi: 10.2903/j.efsa.2009.1377

EFSA (2016). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in2015. *EFSA J.* 14:e04634. doi: 10.2903/j.efsa.2016.4634

Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., et al. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clin. Microbiol. Infect.* 23, 2–22. doi: 10.1016/j.cmi.2016.11.012

Felten, A., Vila Nova, M., Durimel, K., Guillier, L., Mistou, M.-Y., and Radomski, N. (2017). First gene-ontology enrichment analysis based on bacterial core genome variants: insights into adaptations of *Salmonella* serovars to mammalian- and avian-hosts. *BMC Microbiol.* 17:222. doi: 10.1186/s12866-017-1132-1

Ferrari, R. G., Panzenhagen, P. H. N., and Conte-Junior, C. A. (2017). Phenotypic and genotypic eligible methods for *Salmonella* typhimurium source tracking. *Front. Microbiol.* 8:2587. doi: 10.3389/fmicb.2017.02587

Figueiredo, R., Card, R., Nunes, C., AbuOun, M., Bagnall, M. C., Nunez, J., et al. (2015). Virulence characterization of *Salmonella enterica* by a new microarray: detection and evaluation of the cytolethal distending toxin gene activity in the unusual host *S.* Typhimurium. *PLoS One* 10:e0135010. doi: 10.1371/journal.pone.0135010

Fois, F., Piras, F., Torpdahl, M., Mazza, R., Consolati, S. G., Spanu, C., et al. (2017). Occurrence, characterization, and antimicrobial susceptibility of *Salmonella enterica* in slaughtered pigs in Sardinia. *J. Food Sci.* 82, 969–976. doi: 10.1111/1750-3841.13657

Gilchrist, C. A., Turner, S. D., Riley, M. F., Petri, W. A. Jr., and Hewlett, E. L. (2015). Whole-genome sequencing in outbreak analysis. *Clin. Microbiol. Rev.* 28, 541–563. doi: 10.1128/cmr.00075-13

Grimont, P., and Weill, F.-X. (2007). *Antigenic Formulae of the Salmonella serovars. [Online]. Paris: WHO Collaborating Center for Reference and Research on*

Salmonella, Institut Pasteur. 9th Edn. Available at: https://www.pasteur.fr/sites/default/files/veng_0.pdf [accessed October 25, 2016].

Grzymajlo, K., Ugorski, M., Kolenda, R., Kedzierska, A., Kuzminska-Bajor, M., and Wieliczko, A. (2013). FimH adhesin from host unrestricted *Salmonella* Enteritidis binds to different glycoprotein ligands expressed by enterocytes from sheep, pig and cattle than FimH adhesins from host restricted *Salmonella* Abortus-ovis, *Salmonella* Choleraesuis and *Salmonella* Dublin. *Vet. Microbiol.* 166, 550–557. doi: 10.1016/j.vetmic.2013.07.004

Hauser, E., Hebner, F., Tietze, E., Helmuth, R., Junker, E., Prager, R., et al. (2011). Diversity of *Salmonella enterica* serovar Derby isolated from pig, pork and humans in Germany. *Int. J. Food Microbiol.* 151, 141–149. doi: 10.1016/j.ijfoodmicro.2011.08.020

Hayward, M. R., AbuOun, M., La Ragione, R. M., Tchorzewska, M. A., Cooley, W. A., Everest, D. J., et al. (2014). SPI-23 of *S.* Derby: role in adherence and invasion of porcine tissues. *PLoS One* 9:e107857. doi: 10.1371/journal.pone.0107857

Hayward, M. R., Jansen, V., and Woodward, M. J. (2013). Comparative genomics of *Salmonella enterica* serovars Derby and Mbandaka, two prevalent serovars associated with different livestock species in the UK. *BMC Genomics* 14:365. doi: 10.1186/1471-2164-14-365

Hayward, M. R., Petrovska, L., Jansen, V. A., and Woodward, M. J. (2016). Population structure and associated phenotypes of *Salmonella enterica* serovars Derby and Mbandaka overlap with host range. *BMC Microbiol.* 16:15. doi: 10.1186/s12866-016-0628-4

Huang, X., Huang, Q., Dun, Z., Huang, W., Wu, S., Liang, J., et al. (2016). Nontyphoidal *Salmonella* infection, Guangdong Province, China, 2012. *Emerg. Infect. Dis.* 22, 726–729. doi: 10.3201/eid2204.151372

Keelara, S., Scott, H. M., Morrow, W. M., Gebreyes, W. A., Correa, M., Nayak, R., et al. (2013). Longitudinal study of distributions of similar antimicrobial-resistant *Salmonella* serovars in pigs and their environment in two distinct swine production systems. *Appl. Environ. Microbiol.* 79, 5167–5178. doi: 10.1128/AEM.01419-13

Kerouanton, A., Hirchaud, E., Rose, V., Esnault, E., Naquin, D., and Denis, M. (2015). First complete genome sequence of a *Salmonella enterica* subsp. *enterica* serovar derby strain associated with pork in France. *Genome Announc.* 3:e853-15. doi: 10.1128/genomeA.00853-15

Kerouanton, A., Marault, M., Lailler, R., Weill, F. X., Feurer, C., Espie, E., et al. (2007). Pulsed-field gel electrophoresis subtyping database for foodborne *Salmonella enterica* serotype discrimination. *Foodborne Pathog. Dis.* 4, 293–303. doi: 10.1089/fpd.2007.0090

Kerouanton, A., Rose, V., Weill, F. X., Granier, S. A., and Denis, M. (2013). Genetic diversity and antimicrobial resistance profiles of *Salmonella enterica* serotype derby isolated from pigs, pork, and humans in France. *Foodborne Pathog. Dis.* 10, 977–984. doi: 10.1089/fpd.2013.1537

Leclerc, V., Moury, F., Noel, V., Berta-Vanrullen, I., Cadel-Six, S., and Lailler, R. (2016). Le réseau *Salmonella*, un dispositif de surveillance des salmonelles sur la chaîne alimentaire: bilan 2015. *Bull. Épidémiol. Santé Anim. Aliment.* 77, 75–81.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Markowski, C. A., and Markowski, E. P. (1990). Conditions for the effectiveness of a preliminary test of variance. *Am. Stat.* 44, 322–326. doi: 10.2307/2684360

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

McWhorter, A. R., and Chousalkar, K. K. (2015). Comparative phenotypic and genotypic virulence of *Salmonella* strains isolated from Australian layer farms. *Front. Microbiol.* 6:12. doi: 10.3389/fmicb.2015.00012

Méheust, D., Chevance, A., and Moulin, G. (2016). *Suivi des Ventes de Médicaments Vétérinaires Contenant des Antibiotiques en France en 2015*. Buenos Aires: Anses.

Menard, J.-N., Nil, A., and Pierre, T. (2015). *Bilan Diagnostic des Bassins de Production de Volailles de Chair*. Paris: CGAAER.

Pornsukarom, S., and Thakur, S. (2017). Horizontal dissemination of antimicrobial resistance determinants in multiple *Salmonella* serotypes following isolation from the commercial swine operation environment after manure application. *Appl. Environ. Microbiol.* 83:e1503-17. doi: 10.1128/aem.01503-17

Reynolds, C. J., Jones, C., Blohmke, C. J., Darton, T. C., Goudet, A., Sergeant, R., et al. (2014). The serodominant secreted effector protein of *Salmonella*, SseB, is a strong CD4 antigen containing an immunodominant epitope presented by diverse HLA class II alleles. *Immunology* 143, 438–446. doi: 10.1111/imm.12327

Rostagno, M. H., Hurd, H. S., McKean, J. D., Ziemer, C. J., Gailey, J. K., and Leite, R. C. (2003). Preslaughter holding environment in pork plants is highly contaminated with *Salmonella enterica*. *Appl. Environ. Microbiol.* 69, 4489–4494. doi: 10.1128/AEM.69.8.4489-4494.2003

Royston, P. (1995). Remark as R94: a remark on algorithm as 181: the W-test for normality. *J. R. Stat. Soc. Series C* 44, 547–551. doi: 10.2307/2986146

Simon, S., Trost, E., Bender, J., Fuchs, S., Malorny, B., Rabsch, W., et al. (2017). Evaluation of WGS based approaches for investigating a food-borne outbreak caused by *Salmonella enterica* serovar Derby in Germany. *Food Microbiol.* 71, 46–54. doi: 10.1016/j.fm.2017.08.017

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Tenor, J. L., McCormick, B. A., Ausubel, F. M., and Aballay, A. (2004). Caenorhabditis elegans-based screen identifies *Salmonella* virulence factors required for conserved host-pathogen interactions. *Curr. Biol.* 14, 1018–1024. doi: 10.1016/j.cub.2004.05.050

Timme, R. E., Pettengill, J. B., Allard, M. W., Strain, E., Barrangou, R., Wehnes, C., et al. (2013). Phylogenetic diversity of the enteric pathogen *Salmonella enterica* subsp. *enterica* inferred from genome-wide reference-free SNP characters. *Genome Biol. Evol.* 5, 2109–2123. doi: 10.1093/gbe/evt159

Valdezate, S., Vidal, A., Herrera-Leon, S., Pozo, J., Rubio, P., Usera, M. A., et al. (2005). *Salmonella* derby clonal spread from pork. *Emerg. Infect. Dis.* 11, 694–698. doi: 10.3201/eid1105.041042

Wang, B., Wang, C., McKean, J. D., Logue, C. M., Gebreyes, W. A., Tivendale, K. A., et al. (2011). *Salmonella enterica* in swine production: assessing the association between amplified fragment length polymorphism and epidemiological units of concern. *Appl. Environ. Microbiol.* 77, 8080–8087. doi: 10.1128/aem.00064-11

Weill, F., and Le Hello, S. (2014). *Bilan des Activités 2013 du Centre National de Référence des Esherichia coli, Shigella et Salmonella [Online]. Centre National de Référence des Escherichia coli, Shigella et Salmonella*. Available at: https://www.pasteur.fr/fr/sante-publique/CNR/les-cnr/escherichia-coli-shigella-salmonella/rapports-d-activite [accessed December 12, 2016].

Wilson, J. W., Coleman, C., and Nickerson, C. A. (2007). Cloning and transfer of the *Salmonella* pathogenicity island 2 type III secretion system for studies of a range of gram-negative genera. *Appl. Environ. Microbiol.* 73, 5911–5918. doi: 10.1128/aem.00952-07

Wilson, M. R., Brown, E., Keys, C., Strain, E., Luo, Y., Muruvanda, T., et al. (2016). Whole genome DNA sequence analysis of *Salmonella* subspecies *enterica* serotype Tennessee obtained from related peanut butter foodborne outbreaks. *PLoS One* 11:e0146929. doi: 10.1371/journal.pone.0146929

Yue, M., Han, X., Masi, L. D., Zhu, C., Ma, X., Zhang, J., et al. (2015). Allelic variation contributes to bacterial host specificity. *Nat. Commun.* 6:8754. doi: 10.1038/ncomms9754

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

Zheng, H., Hu, Y., Li, Q., Tao, J., Cai, Y., Wang, Y., et al. (2017). Subtyping *Salmonella enterica* serovar Derby with multilocus sequence typing (MLST) and clustered regularly interspaced short palindromic repeats (CRISPRs). *Food Control* 73(Part B), 474–484. doi: 10.1016/j.foodcont.2016.08.051

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

2.2.1 Supplementary material

Supplementary table 1: Strains information: Epidata, sample ID, NCBI accession number, genome deep coverage, and breadth.

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL00238	07/01/2014	Rhone-Alpes	45.6486	5.2817	Pork	Industrial transformation	Meat from pig - fresh	SAMN07267519	273	91.92%
2014LSAL00274	09/01/2014	Lorraine	49.0163	6.6165	Pork	Retail sale	Meat from pig - meat products - raw but intended to be eaten cooked	SAMN07734874	113	91.4%
2014LSAL00337	14/01/2014	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig - meat preparation intended to be eaten cooked	SAMN07734875	129	91.38%
2014LSAL00455	15/01/2014	Poitou-Charentes	46.4719	0.7745	Pork	Slaughter	Meat from pig - fresh	SAMN07734876	75	91.41%
2014LSAL00717	11/02/2014	Midi-Pyrenees	43.4888	1.8842	Pork	Slaughter	Meat from pig - carcass	SAMN07734877	235	91.0%
2014LSAL00762	18/02/2014	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig - fresh	SAMN07734878	282	91.28%
2014LSAL00814	28/01/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734879	307	91.42%
2014LSAL00873	17/02/2014	Bretagne	48.0083	-1.2387	Pork	Unspecified	Meat from pig - fresh	SAMN07734880	221	91.34%
2014LSAL00980	05/03/2014	Bretagne	48.0083	-1.2387	Pork	Industrial transformation	Meat from pig - fresh	SAMN07734881	99	91.4%
2014LSAL01158	14/03/2014	Pays de la Loire	47.1094	-1.1124	Pork	Retail sale	Meat from pig - offal	SAMN07734882	142	91.56%
2014LSAL01220	17/03/2014	Rhone-Alpes	45.6486	5.2817	Pork	Slaughter	Meat from pig - fresh	SAMN07734883	244	91.43%
2014LSAL01256	13/03/2014	Provence-Alpes-Cote d'Azur	43.1811	5.809	Pork	Slaughter	Meat from pig - carcass	SAMN07734884	139	91.36%
2014LSAL01306	24/03/2014	Limousin	45.7388	1.6891	Pork	Retail sale	Meat from pig - offal	SAMN07734885	51	91.28%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL01311	16/03/2014	Pays de la Loire	47.1094	-1.1124	Poultry	Slaughter	Meat from duck - carcass	SAMN07734886	164	92.55%
2014LSAL01383	26/03/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734887	376	92.55%
2014LSAL01467	04/04/2014	Auvergne	45.6716	3.5211	Pork	Slaughter	Meat from pig - carcass	SAMN07734888	227	91.46%
2014LSAL01499	01/04/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734889	309	92.53%
2014LSAL01651	09/04/2014	Pays de la Loire	47.1094	-1.1124	Pork	Retail sale	Meat from pig - offal	SAMN07734890	140	91.4%
2014LSAL01779	24/04/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734891	236	92.58%
2014LSAL01780	24/04/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734892	251	92.57%
2014LSAL01781	24/04/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734893	190	92.51%
2014LSAL01894	28/04/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig	SAMN07734894	252	91.43%
2014LSAL02051	21/05/2014	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734895	229	91.4%
2014LSAL02073	13/05/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734896	101	92.39%
2014LSAL02163	20/05/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734897	208	92.47%
2014LSAL02205	04/06/2014	Poitou-Charentes	46.4719	0.7745	Poultry	Livestock farming	Gallus gallus - fowl - broilers	SAMN07734898	247	91.4%
2014LSAL02278	06/01/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - fresh	SAMN07734899	178	91.42%
2014LSAL02285	21/05/2014	Poitou-Charentes	46.4719	0.7745	Pork	Slaughter	Meat from pig - carcass	SAMN07734900	238	91.36%
2014LSAL02325	02/06/2014	Rhone-Alpes	45.6486	5.2817	Poultry	Slaughter	Meat from turkey - carcass	SAMN07734901	275	92.55%
2014LSAL02431	18/06/2014	Nord-Pas-de-Calais	50.5361	3.5019	Pork	Industrial transformation	Meat from pig - carcass	SAMN07734902	182	91.48%
2014LSAL02441	18/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734903	281	92.53%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL02463	20/06/2014	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - fresh	SAMN07734904	257	91.38%
2014LSAL02547	18/06/2014	Poitou-Charentes	46.4719	0.7745	Pork	Slaughter	Meat from pig - fresh	SAMN07734905	240	91.5%
2014LSAL02838	11/07/2014	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734906	250	91.37%
2014LSAL03150	15/07/2014	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - fresh	SAMN07734907	239	92.51%
2014LSAL03218	30/07/2014	Pays de la Loire	47.1094	-1.1124	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - mechanically separated meat - MSM	SAMN07734908	217	92.96%
2014LSAL03309	02/04/2014	Pays de la Loire	47.1094	-1.1124	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734909	152	92.51%
2014LSAL03340	05/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734910	250	92.54%
2014LSAL03342	04/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734911	229	92.51%
2014LSAL03344	05/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734912	238	92.55%
2014LSAL03348	11/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734913	232	92.53%
2014LSAL03350	11/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734914	226	92.52%
2014LSAL03355	11/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734915	245	92.5%
2014LSAL03356	11/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734916	240	92.53%
2014LSAL03383	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734917	251	92.53%
2014LSAL03388	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734918	238	92.53%
2014LSAL03390	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734919	214	92.54%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL03393	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734920	233	92.52%
2014LSAL03396	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734921	225	92.53%
2014LSAL03398	17/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734922	205	92.53%
2014LSAL03401	24/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734923	238	92.53%
2014LSAL03404	24/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734924	229	92.67%
2014LSAL03407	24/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734926	233	92.53%
2014LSAL03409	24/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734927	253	92.53%
2014LSAL03410	24/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734928	205	92.56%
2014LSAL03414	02/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734929	228	92.56%
2014LSAL03416	02/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734930	193	92.55%
2014LSAL03420	02/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734931	265	92.52%
2014LSAL03463	09/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734932	251	92.54%
2014LSAL03465	09/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734933	278	92.55%
2014LSAL03470	15/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734934	251	92.51%
2014LSAL03478	15/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734935	256	92.55%
2014LSAL03510	30/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734936	279	92.55%
2014LSAL03521	22/07/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734937	236	92.54%
2014LSAL03524	30/06/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - carcass	SAMN07734938	227	92.5%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL03616	13/08/2014	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - meat preparation	SAMN07734939	214	91.37%
2014LSAL03694	08/08/2014	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from guinea fowl	SAMN07734940	297	92.55%
2014LSAL03785	27/08/2014	Pays de la Loire	47.1094	-1.1124	Pork	Slaughter	Meat from pig - offal	SAMN07734941	271	91.92%
2014LSAL03787	27/08/2014	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig - offal	SAMN07734942	350	91.92%
2014LSAL03926	27/08/2014	Bretagne	48.0083	-1.2387	Pork	Industrial transformation	Meat from pig - fresh	SAMN07734943	329	97.39%
2014LSAL03973	07/08/2014	Aquitaine	44.7253	-0.1401	Pork	Industrial transformation	Meat from pig - meat preparation	SAMN07734944	248	91.34%
2014LSAL03974	07/08/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734945	269	91.41%
2014LSAL03975	07/08/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734946	194	91.42%
2014LSAL03976	07/08/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734947	210	91.41%
2014LSAL03985	15/09/2014	Rhone-Alpes	45.6486	5.2817	Pork	Slaughter	Meat from pig - carcass	SAMN07734948	119	92.3%
2014LSAL04065	17/09/2014	Nord-Pas-de-Calais	50.5361	3.5019	Pork	Industrial transformation	Meat from pig - meat products	SAMN07734949	247	91.28%
2014LSAL04466	01/10/2014	Pays de la Loire	47.1094	-1.1124	Pork	Retail sale	Meat from pig - meat preparation intended to be cooked	SAMN07734950	174	91.41%
2014LSAL04673	06/10/2014	Rhone-Alpes	45.6486	5.2817	Pork	Industrial transformation	Meat from pig	SAMN07734951	258	92.45%
2014LSAL04833	17/09/2014	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - minced meat - intended to be eaten cooked	SAMN07734952	135	91.38%
2014LSAL05133	05/11/2014	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from turkey - carcass	SAMN07734953	206	93.35%

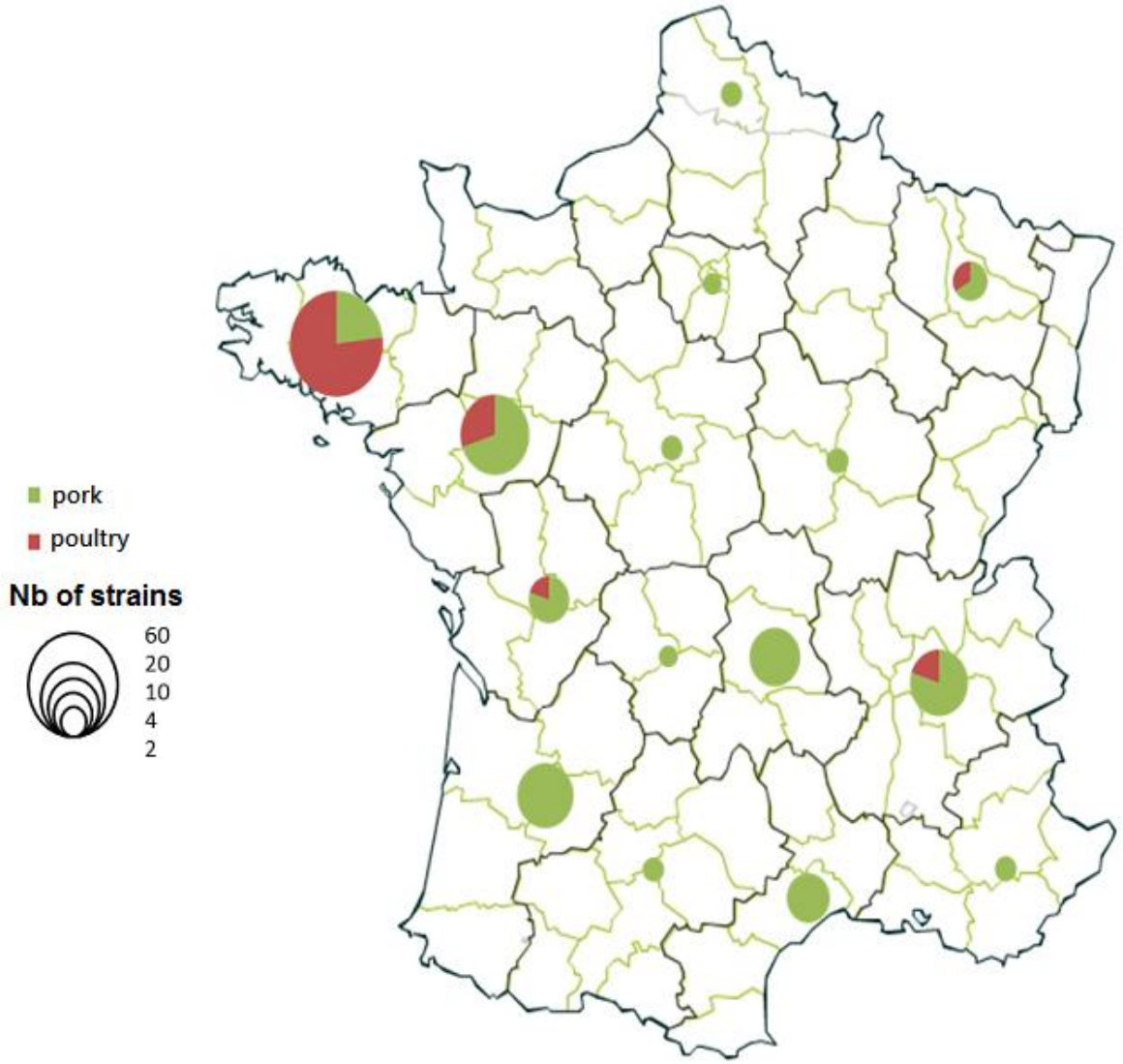
Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2014LSAL05213	12/11/2014	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - carcass	SAMN07734954	193	92.58%
2014LSAL05495	29/10/2014	Languedoc-Roussillon	43.4968	4.3176	Pork	Industrial transformation	Meat from pig - minced meat - intended to be eaten cooked	SAMN07734955	211	91.43%
2014LSAL05538	01/12/2014	Pays de la Loire	47.1094	-1.1124	Pork	livestock farming	Meat from pig - fresh	SAMN07734956	290	91.33%
2014LSAL05735	06/11/2014	Ile-de-France	48.7508	2.7905	Pork	Retail sale	Meat from pig - fresh	SAMN07734957	244	92.5%
2015LSAL00134	27/10/2014	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - fresh	SAMN07734958	265	91.47%
2015LSAL00137	17/12/2014	Poitou-Charentes	46.4719	0.7745	Pork	Slaughter	Meat from pig - fresh	SAMN07734959	227	91.38%
2015LSAL00190	18/12/2014	Lorraine	49.0163	6.6165	Pork	Retail sale	Meat from pig - meat products - raw but intended to be eaten cooked	SAMN07734960	286	91.45%
2015LSAL00219	30/12/2014	Pays de la Loire	47.1094	-1.1124	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - fresh	SAMN07734961	257	92.54%
2015LSAL00314	09/10/2014	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734962	251	91.52%
2015LSAL00428	02/06/2014	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734963	236	91.42%
2015LSAL00429	02/06/2014	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734964	243	91.41%
2015LSAL00452	10/06/2014	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734965	241	91.4%
2015LSAL00465	10/06/2014	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734966	189	91.39%
2015LSAL00477	21/01/2015	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - fresh	SAMN07734967	216	92.5%
2015LSAL00641	30/01/2015	Centre	47.7947	2.3483	Pork	Industrial transformation	Meat from pig - fresh	SAMN07734968	254	93.62%
2015LSAL00691	03/02/2015	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - carcass	SAMN07734969	282	91.32%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2015LSAL00728	13/02/2015	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig - meat preparation destinee a etre consommee cuite	SAMN07734970	237	91.34%
2015LSAL00981	05/03/2015	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - carcass	SAMN07734971	257	91.34%
2015LSAL00989	02/03/2015	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734972	262	91.43%
2015LSAL00990	16/02/2015	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734973	205	91.4%
2015LSAL01029	10/03/2015	Pays de la Loire	47.1094	-1.1124	Pork	Retail sale	Meat from pig - meat preparation destinee a etre consommee cuite	SAMN07734974	233	91.42%
2015LSAL01058	17/03/2015	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734975	212	91.4%
2015LSAL01075	17/03/2015	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - fresh	SAMN07734976	156	92.53%
2015LSAL01179	25/03/2015	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - fresh	SAMN07734977	130	92.56%
2015LSAL01314	23/03/2015	Centre	47.7947	2.3483	Pork	Slaughter	Meat from pig - carcass	SAMN07734978	224	91.4%
2015LSAL01334	24/03/2015	Rhone-Alpes	45.6486	5.2817	Pork	Slaughter	Meat from pig - carcass	SAMN07734979	220	91.38%
2015LSAL01341	01/04/2015	Lorraine	49.0163	6.6165	Poultry	Livestock farming	Gallus gallus - fowl - broilers	SAMN07734980	307	91.42%
2015LSAL01474	15/04/2015	Bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from broilers - Gallus gallus - mechanically separated meat - MSM	SAMN07734981	321	92.53%
2015LSAL01578	22/04/2015	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig	SAMN07734982	215	91.39%
2015LSAL01591	28/04/2015	Pays de la Loire	47.1094	-1.1124	Poultry	Slaughter	Meat from broilers - Gallus gallus - fresh	SAMN07734983	229	92.53%
2015LSAL01697	13/05/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig – carcass	SAMN07734984	192	91.38%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2015LSAL01713	17/03/2015	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig	SAMN07734985	221	91.41%
2015LSAL01716	24/03/2015	Auvergne	45.6716	3.5211	Pork	Slaughter	Meat from pig - carcass	SAMN07734986	176	92.49%
2015LSAL01748	20/04/2015	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734987	224	91.29%
2015LSAL01817	28/04/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734988	186	91.28%
2015LSAL01821	19/05/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07734989	214	91.2%
2015LSAL01829	27/05/2015	Bretagne	48.0083	-1.2387	Poultry	Slaughter	Meat from broilers - Gallus gallus - fresh	SAMN07734990	211	92.51%
2015LSAL01973	30/04/2015	Midi-Pyrenees	43.4888	1.8842	Pork	Industrial transformation	Meat from pig - meat preparation	SAMN07734992	246	91.92%
2015LSAL02005	04/06/2015	bretagne	48.0083	-1.2387	Poultry	Industrial transformation	Meat from duck - fresh	SAMN07734993	238	92.53%
2015LSAL02462	07/07/2015	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07734994	221	91.4%
2015LSAL02588	23/06/2015	Aquitaine	44.7253	-0.1401	Pork	Industrial transformation	Meat from pig - minced meat - intended to be eaten cooked	SAMN07734995	234	92.52%
2015LSAL02717	28/07/2015	Rhone-Alpes	45.6486	5.2817	Pork	Slaughter	Meat from pig - carcass	SAMN07734997	205	91.36%
2015LSAL03121	17/08/2015	Aquitaine	44.7253	-0.1401	Pork	Slaughter	Meat from pig - carcass	SAMN07734998	238	91.35%
2015LSAL03260	02/09/2015	Rhone-Alpes	45.6486	5.2817	Pork	Slaughter	Meat from pig - carcass	SAMN07734999	117	91.25%
2015LSAL03284	11/08/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07735000	226	91.15%
2015LSAL03285	17/08/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07735001	209	91.13%
2015LSAL03548	09/06/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07735002	129	91.25%
2015LSAL03550	08/06/2015	Bretagne	48.0083	-1.2387	Pork	Slaughter	Meat from pig - carcass	SAMN07735003	138	91.21%
2015LSAL03563	21/09/2015	Auvergne	45.6716	3.5211	Pork	Slaughter	Meat from pig - carcass	SAMN07735004	175	92.91%
2015LSAL03564	14/09/2015	Rhone-Alpes	45.6486	5.2817	Pork	Industrial transformation	Meat from pig - fresh	SAMN07735005	231	91.42%

Key	Sampling date	Region	Lattitude	Longitude	Sector	Sampling stage	Matrix	NCBI accession	Deep coverage	Deep breadth
2015LSAL03879	14/10/2015	Languedoc-Roussillon	43.4968	4.3176	Pork	Slaughter	Meat from pig - carcass	SAMN07735006	265	91.43%
2015LSAL03891	02/10/2015	Auvergne	45.6716	3.5211	Pork	Industrial transformation	Meat from pig - carcass	SAMN07735007	204	91.47%
2015LSAL04018	22/10/2015	Rhone-Alpes	45.6486	5.2817	Poultry	Livestock farming	Gallus gallus - fowl - broilers	SAMN07735008	212	91.42%
2015LSAL04158	28/10/2015	Auvergne	45.6716	3.5211	Pork	Slaughter	Meat from pig - carcass	SAMN07735009	232	91.52%
2015LSAL04225	03/11/2015	Bourgogne	47.214	5.4822	Pork	Retail sale	Meat from pig - fresh	SAMN07735010	140	91.39%
2015LSAL04227	26/10/2015	Bourgogne	47.214	5.4822	Pork	Slaughter	Meat from pig - carcass	SAMN07735011	233	91.37%
2015LSAL04332	03/11/2015	Pays de la Loire	47.1094	-1.1124	Poultry	Slaughter	Meat from broilers - Gallus gallus - fresh	SAMN07735013	222	92.52%
2015LSAL04562	31/08/2015	Pays de la Loire	47.1094	-1.1124	Pork	Slaughter	Meat from pig - carcass	SAMN07735014	107	91.22%
2016LSAL00028	22/12/2015	Pays de la Loire	47.1094	-1.1124	Pork	Industrial transformation	Meat from pig - fresh	SAMN07735015	274	91.34%
2016LSAL00056	30/11/2015	Aquitaine	44.7253	-0.1401	Pork	Industrial transformation	Meat from pig - offal	SAMN07735016	165	91.33%

Supplementary figure 1 : Geographical selection plan for the strains isolated from the pork and poultry sectors.



Supplementary table 2: List of the genomes from different serotypes of Salmonella enterica subsp. enterica used for the FimH sequence analysis. Epidata and sample ID.

Genome name	Strain	Serotype	Sample ID	Country	Collection Year	Source Type
Agona 12A06	12A06	Agona	SAMEA1889081	Ireland	2006	Poultry
Anatum str. CVM_N44699	CVM_N44699	Anatum	SAMN03894116	United States	2013	Poultry
Derby 2014LSAL00238	2014LSAL00238	Derby	SAMN07267519	France	2014	swine
Derby 2014LSAL00814	2014LSAL00814	Derby	SAMN07734879	France	2014	Swine
Derby 2014LSAL02431	2014LSAL02431	Derby	SAMN07734902	France	2014	Swine
Derby 2014LSAL03350	2014LSAL03350	Derby	SAMN07734914	France	2014	Poultry
Derby 2014LSAL03785	2014LSAL03785	Derby	SAMN07734941	France	2014	swine
Derby 2014LSAL03926	2014LSAL03926	Derby	SAMN07734943	France	2014	Swine
Derby 2014LSAL05133	2014LSAL05133	Derby	SAMN07734953	France	2014	Poultry
Derby 2015LSAL00614	2014LSAL00614	Derby	SAMN07734968	France	2015	Swine
Derby 2015LSAL01829	2015LSAL01829	Derby	SAMN07734990	France	2015	Poultry
Derby 2015LSAL01973	2015LSAL01973	Derby	SAMN07734992	France	2015	swine
Derby 2015LSAL02717	2014LSAL02717	Derby	SAMN07734997	France	2015	Swine
Derby 2015LSAL04158	2014LSAL04158	Derby	SAMN07735009	France	2015	Swine
Dublin str. NY_IDR1200021877-01	NY_IDR1200021877-01	Dublin	SAMN01902415	United States	2012	Bovine
Enteritidis str. EC20120005	str. EC20120005	Enteriditis	SAMN02384189	Canada	2011	Poultry
Hadar CVM-N46859	CVM N46859	Hadar	SAMN03894238	United States	2013	Poultry
Havana BCW_2696	BCW 2696	Havana	SAMN02368573	Nigeria	2009	Poultry
Indiana C629	C629	Indiana	SAMN04961653	China	2014	Poultry
Infantis CVM-N45946	CVM-N45946	Infantis	SAMN03894177	United States	2013	Swine
Kentucky CVM N51982	CVM N51982	Kentucky	SAMN03894408	United States	2013	Poultry
Livingstone str. MDH-2014-00079	MDH-2014-00079	Livingstone	SAMN02646853	United States	2003	Swine
Mbandaka str. WAPHL_SAL-A00124	WAPHL_SAL-A00124	Mbandaka	SAMN02182958	United States	2012	Giblets not specified
Montevideo str. 19N	str. 19N	Montevideo	SAMN00710598	nd	nd	Poultry

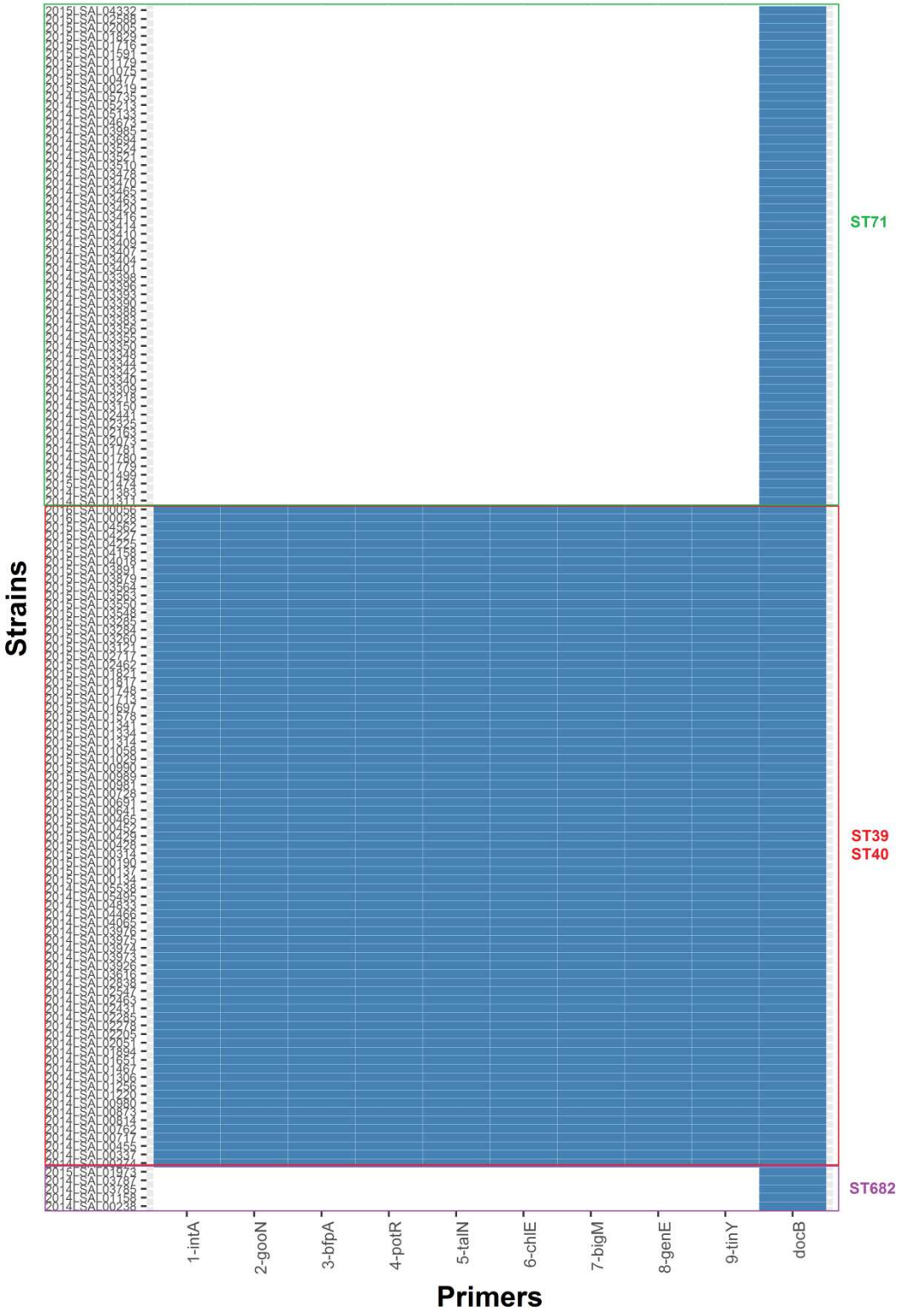
Genome name	Strain	Serotype	Sample ID	Country	Collection Year	Source Type
Muenster VA-WGS-str00091	VA WGS str. 00091	Muenster	SAMN02403312	United States	2006	Bovine
Newport CVM21554	CVM21554	Newport	SAMN01813483	United States	2001	Swine
Oranienburg BCW-2047	BCW-2047	Oranienburg	SAMN02367859	United States	2010	Poultry
Panama FSW0120	FSW0120	Panama	SAMN02678847	Vietnam	2013	Aquatic
Poona FSF0055	FSF0055	Poona	SAMN02344819	China	2010	Invertebrates
Rissen FAR0099	FAR0099	Rissen	SAMN02345539	Mexico	2012	Plant
Saintpaul NY-IDR1000004536	NY IDR1000004536	Saintpaul	SAMN01902359	United States	2010	Poultry
Senftenberg CVm-N51312	CVM N51312	Senftenberg	SAMN03894403	United States	2013	Poultry
Tennessee FSF0063	FSF0063	Tennessee	SAMN02344793	China	2010	Plant
Typhi CR0044	CR0044	Typhi	SAMN02415239	Malaysia	2007	Human
Typhimurium SL1344 (reference)	SL1344	Typhimurium	SAMEA3138382	United Kindoms	nd	nd
Typhimurium str. 798	str. 798	Typhimurium	SAMN02604223	France	nd	Swine
Virchow str. FAR0119	FAR0119	Virchow	SAMN02678442	India	2013	Plant
Choleraesuis SC-B67	SC-B67	Choleraesuis	SAMN02603109	Taiwan	2002	Human
Typhisuis 8719-97	8719-97	Typhisuis	KJ095885.1	ND	ND	Swine

Supplementary table 3: List of strains presenting antimicrobial resistance genes.

Key	ST profile	Antimicrobial resistance profile (in red = prediction)	Aminoglycosides	Beta-lactams	Tetracyclines	Sulphonamides	Trimetroprim
2014LSAL00274	ST40	TET			tet(B)		
2014LSAL00455	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL00717	ST40	TET			tet(B)		
2014LSAL00762	ST40	STR AMP SMX SXT TET	aadA1	bla TEM-1C	tet(A)	sul1	dfrA1
2014LSAL00873	ST40	AMP SMX SXT		blaTEM-1B		sul2	dfrA1
2014LSAL00980	ST40	STR SMX	aadA2			sul1	
2014LSAL01158	ST682	STR KAN TET	aph (3')-IC + strA + strB		tet(B)		
2014LSAL01220	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL01256	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL01306	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL01651	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL01894	ST40	STR SMX SXT TET	aadA2 + ΔstrA + strB		tet(A)	sul1 + sul2	dfrA14
2014LSAL02051	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02205	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02278	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02285	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02325	ST71	AMP SMX SXT TET	ΔstrA + strB	blaTEM-1B	tet(A)	sul2	dfrA14
2014LSAL02431	ST40	AMP CEF CTX CAZ		blaCTX-M-1			
2014LSAL02463	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02547	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL02838	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL03616	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL03973	ST40	STR SMX TET	aadA2		tet(A)	sul1	

Key	ST profile	Antimicrobial resistance profile (in red = prediction)	Aminoglycosides	Beta-lactams	Tetracyclines	Sulphonamides	Trimetroprim
2014LSAL03974	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL03975	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL03976	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL04065	ST40	STR SMX TET	aadA2		tet(A)	sul3	dfrA12
2014LSAL04466	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL04833	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2014LSAL05495	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00134	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00137	ST40	STR SMX TET	aadA2		tet(A)	sul2	
2015LSAL00190	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00428	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00429	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00452	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00465	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00728	ST40	TET			tet(C)		
2015LSAL00989	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL00990	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01029	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01058	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01314	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01334	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01341	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL01697	ST40	STR KAN SMX TET	aadA2 + aadA5		tet(A)	sul1	
2015LSAL01713	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL02717	ST40	STR SMX TET	aadA2		tet(A)	sul1	
2015LSAL03260	ST40	STR SMX TET	aadA2		tet(A)	sul1	

Supplementary figure 2: Results of the in silico PCR for the SPI-23 into S. Derby collection.



The docB gene corresponds to the coding sequence located immediately after the SPI-23. Only 2014LSAL05133 present small fragments of the SPI-23 (not corresponding to the primers highlighted in this figure) in the ST71.